



**HAL**  
open science

# Les outils de Business Intelligence dans le secteur des études et de la communication

Tom Blachon

► **To cite this version:**

Tom Blachon. Les outils de Business Intelligence dans le secteur des études et de la communication. domain\_shs.info.docu. 2022. mem\_03847446

**HAL Id: mem\_03847446**

**[https://memic.ccsd.cnrs.fr/mem\\_03847446v1](https://memic.ccsd.cnrs.fr/mem_03847446v1)**

Submitted on 10 Nov 2022

**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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



le **cnam**  
intd

Mémoire pour l'obtention du  
**Master Sciences humaines et sociales**  
**mention humanités numériques**

**Parcours Mégadonnées et analyse sociale (MEDAS)**

Les outils de *Business Intelligence* dans le  
secteur des études et de la communication

**Tom BLACHON**

Date et lieu de la soutenance

- Mercredi 6 juillet 2022
- CFA CNAM IDF - La Plaine Saint-Denis

Membres du jury

- Ghislaine CHARTRON, directrice

Promotion (2020-2022)



Paternité Pas d'Utilisation Commerciale - Pas de Modification

**BLACHON Tom.** Les outils de *Business Intelligence* dans le secteur des études et de la communication. Mémoire professionnel INTD, Titre I, Master Méga Données et Analyse Sociale. Conservatoire national des arts et métiers, 2021, 101p.

Ce mémoire aborde sous différents aspects les techniques et les outils de l'informatique décisionnelle (ou *Business Intelligence/BI*) dans le domaine des études et sondages. Il propose en premier lieu une approche théorique du sujet visant à présenter les grands concepts de la *BI* et ses possibilités d'applications au sein des instituts de sondages. Dans un second temps, il présente un projet de déploiement d'outils *BI* en situation réelle au sein du cabinet d'étude *Occurrence* sur plus de 18 mois. Cette partie est notamment consacrée au progiciel *Microsoft Power BI* dont l'intégration à fait l'objet de nombreux tests itératifs avant de proposer un service de création de tableaux de bord aux clients de l'entreprise.

### Descripteurs

Données  
Informatique décisionnelle  
Sondages  
Analyse de données  
Système d'information  
Intégration  
Déploiement  
Datavisualisation  
Tableaux de bord  
Indicateurs de performance  
Pilotage

This master thesis is an approach to different aspects of *Business Intelligence (BI)* techniques and tools in the field of studies and surveys. It first proposes a theoretical approach of the subject aiming to present the main concepts of *BI* and its possible applications within survey institutes. In the second part, he presents a project of *BI* tools deployment in real situation within the research firm *Occurrence* over 18 months. his section is dedicated to the *Microsoft Power BI* software package, the integration of which was the subject of numerous iterative tests before offering a *Dashboard* creation service to the company's customers.

### Keywords

Data  
*Business Intelligence*  
*Datawarehouse*  
*KPI*  
*Dashboard*  
*Power BI*  
Data analysis  
Survey  
Extract – Transform – Load  
Information system

## **Remerciements :**

*Je tiens à remercier chaleureusement Madame Ghislaine Chartron qui a accepté de diriger cette étude. Elle a su montrer de l'intérêt pour mes travaux, m'apporter son soutien et ses conseils tout au long de l'année.*

*Merci également à Thomas Skorucak pour m'avoir encadré durant mes recherches en entreprise, ainsi qu'à Johann Damen pour avoir supervisé ce projet au sein d'Occurrence. Sans ses conseils avisés, je ne serais pas parvenu à progresser aussi rapidement sur le sujet.*

*Enfin, un grand merci à Chloé et mes proches pour leurs encouragements.*

# Sommaire

Introduction .....	7
<b>I. Présentation de l'entreprise et problématique.....</b>	<b>10</b>
<b>A. Présentation de l'entreprise.....</b>	<b>10</b>
a. Historique de l'entreprise .....	10
b. Secteur, contexte et domaine d'activité métier .....	11
c. Organisation générale .....	13
d. Focus sur le service « Marque et Transformation ».....	15
e. Présentation de mes missions et projets de <i>Business Intelligence</i> .....	17
1. <i>Power BI</i> et datavisualisation .....	17
2. <i>Travail sur les standards</i> .....	18
3. <i>Veille sur les outils de collecte</i> .....	18
4. <i>Participation aux études clients</i> .....	19
<b>B. Développement de la problématique .....</b>	<b>20</b>
a. Contexte <i>data</i> de l'entreprise.....	20
1. Type de données exploitées .....	21
2. Outils et pilotage .....	23
3. Compétences et culture <i>data</i> .....	24
4. Qualité des données .....	26
5. Constats et objectifs.....	28
b. Enjeux des données et de la <i>Business Intelligence</i> .....	29
c. Définition de la problématique d'étude .....	31
<b>II. Cadre théorique : Etude des concepts <i>BI</i> .....</b>	<b>33</b>
<b>A. État de l'art et historiographie.....</b>	<b>33</b>
a. Le traitement des données de sondage et de communication : Approche théorique.....	33
1. Ouvrages et auteurs de référence.....	33
2. Grands principes d'analyse .....	35
3. Biais d'analyse à éviter.....	37
4. Analyse des données de communication.....	39
b. Emergence et généralisation des outils de <i>Business Intelligence</i> .....	40
1. Etat de l'art .....	40
2. Histoire de la <i>Business Intelligence</i> .....	42
3. Rôles et fonctionnement général des outils <i>BI</i> .....	44
<b>B. Concepts clés et terminologie du sujet.....</b>	<b>46</b>
a. Décryptage des concepts phares de l'informatique décisionnelle et du secteur.....	47

1.	Automatisation.....	47
2.	Indicateurs de performance ( <i>KPI</i> ) .....	48
3.	Tableaux de bord.....	49
4.	Datavisualisation .....	50
5.	Concept de « <i>no-code</i> ».....	50
b.	Mise en perspective avec les projets menés par les instituts d'étude.....	52
III.	Cas pratique : Mise en place et exploitation des outils <i>BI</i> au sein d'un cabinet d'étude et de conseils en communication .....	57
A.	Présentation du projet.....	57
a.	Contexte de mise en place.....	57
1.	Besoins concrets.....	57
2.	Equipe.....	59
3.	Objectifs projet.....	60
b.	Approche méthodologique.....	61
B.	Mise en œuvre .....	62
a.	Déploiement de <i>Microsoft Power BI</i> .....	62
1.	Avantages de <i>Power BI</i> .....	63
2.	Phase 1 : Formation à l'outil sur les standards en interne .....	64
3.	Phase 2 : Tests de substitution des outils actuels ( <i>Excel</i> ) .....	67
4.	Phase 2.5 : Formation à l'automatisation.....	69
5.	Phase 3 : Déploiement en interne dans un cas d'usage concret.....	71
6.	Phase 4 : Déploiement lors de missions clients.....	72
7.	Limites de <i>Power BI</i> et obstacles au déploiement .....	74
b.	Visualiser les données spatiales : SIG et cartes de réseaux.....	75
1.	Échantillonnage .....	75
2.	Mieux comprendre la répartition des données à l'aide des cartes .....	76
3.	Cas spécifique de la spatialisation des réseaux ( <i>Géphi</i> ).....	78
c.	Exploration de la suite <i>Microsoft</i> : Anticiper la <i>BI</i> en amont et multiplier les supports d'interprétation .....	80
1.	Joindre les outils de collecte et les tableaux de bord : <i>Power Apps</i> .....	80
2.	Automatiser la collecte et le nettoyage en amont : <i>Power Automate</i> et <i>Power Query</i> ....	83
3.	Assouplir les supports de tableaux de bord : <i>Sharepoint Online</i> et <i>Microsoft Teams</i> ....	84
C.	Pour aller plus loin : Recommandations et mises en garde .....	85
a.	Limites et axes de vigilances .....	86
b.	Bonnes pratiques pour assurer une pérennisation de l'utilisation des outils <i>BI</i> .....	87
	Conclusion.....	90

<b>Bibliographie</b> .....	92
<b>Annexes</b> .....	97

## **Table des abréviations :**

**ACM** : Analyse des Correspondances multiples

**AGD** : Analyse Géométrique des Données

**AO** : Appel d'Offre

**BI** : *Business Intelligence*

**CRM** : *Customer Relationship Management*

**CSP** : Catégories Socioprofessionnelles

**DAX** : *Data Analysis Expressions*

**ETL** : « *Extract, Transform, Load* »

**INSEE** : Institut National de la Statistique et des Etudes Economiques

**KPI** : *Key Performance Indicator*

**PME** : Petite et Moyenne Entreprise

**R&D** : Recherche et Développement

**RGPD** : Règlement Général sur la Protection des Données

**SAAS** : *Software As A Service*

**SI** : Système d'Information

**SIG** : Système d'Information Géographique

**SQL** : *Structured Query Language*

# Introduction

Depuis la fin du XXème siècle, l'exploitation et la collecte des données sont devenus des enjeux centraux dans de nombreux secteurs d'activités. En effet, leur analyse permet par exemple à une entreprise d'identifier de façon claire ses indicateurs de performance clés (*KPI*) et ainsi l'aider à améliorer son expertise. De même, les données peuvent aider une marque ou une institution à mieux identifier son public et donc d'adapter son offre à ce dernier. De ce fait, les données constituent aujourd'hui un formidable outil d'aide à la prise de décision, aux retombées économiques avérées.

Le secteur de la communication n'échappe pas à cette règle. De prime abord, il pourrait pourtant sembler difficile d'exploiter les données au service des actions de communications. En effet, la perception humaine face à une marque, une publicité ou un événement est subjective et donc l'efficacité de cette dernière peut paraître difficilement quantifiable. Mais il est possible, grâce à de nombreuses méthodes, notamment des études quantitatives et qualitatives, d'identifier les *KPI* propres à une campagne de communication.

« Pour améliorer une performance, il faut s'avoir la mesurer. »<sup>1</sup>. Telle est la vision qu'a choisi d'adopter l'entreprise *Occurrence* dans ses études depuis sa création en 1995. Il s'agit d'un cabinet d'étude et conseil en communication que j'ai rejoint en contrat d'apprentissage depuis le 4 janvier 2021. Composé de 24 membres, ce cabinet a pour but d'aider les entreprises et les institutions à évaluer la performance de leur stratégie de communication. Il peut s'agir de communication interne, externe, événementielle ou encore publicitaire, en France et à l'international (50% des études réalisées). L'objectif est ainsi d'éclairer les décisions stratégiques prises par les directions de la communication afin d'optimiser le management et le pilotage au sein de leur organisation.

L'entreprise *Occurrence* possède par ailleurs trois filiales agissant dans des domaines d'activités spécifiques :

- ***Occurrence Healthcare*** : Il s'agit d'un institut d'études spécialisé sur les problématiques de santé, de bien-être et de qualité de vie.
- ***Start Cities*** : Il s'agit d'une filiale de conseil dédiée à la concertation lors de grands projets portés par des entreprises ou des métropoles.

---

<sup>1</sup> Site d'Occurrence, cabinet d'études et conseil en communication. URL : <http://Occurrence.fr/>

- **Deep Opinion** : Filiale spécialisée dans le domaine la communication sur les réseaux sociaux, de l'influence et de l'e-réputation.

*Occurrence* se démarque de nombreux autres cabinets d'études dans ce domaine par son approche et sa méthodologie de travail. En effet, l'entreprise a pour volonté de faire de la communication une discipline quantifiable et mesurable. Pour cela, la collecte et l'analyse de données se place comme un enjeu central de ce cabinet.

Cette approche m'a ainsi décidé à candidater, puis à réaliser mon alternance dans cette entreprise, en accord avec mon projet professionnel. Ayant réalisé un premier master dans les métiers du livre et travaillé au sein d'une bibliothèque durant deux ans, je souhaite à plus long terme me rapprocher à nouveau de la filière culturelle. Encore sous-utilisées aujourd'hui, je souhaite pouvoir accompagner les institutions culturelles dans l'exploitation de leurs données afin de les aider à gagner en efficacité, identifier leurs publics et optimiser leur prise de décision. Le travail d'*Occurrence* auprès de ses clients présente ainsi de nombreuses similarités avec mon projet.

Etant junior en matière de programmation, je désirais par ailleurs pouvoir mettre à profit dans mon entreprise ma double expertise, en analyse de données et en Sciences Humaines et Sociales. J'ai ainsi été séduit par l'approche quasi sociologique adoptée par *Occurrence* lors de ses études. Celle-ci combine à la fois la collecte de données de terrain, avec un traitement de ces dernières à l'aide d'outils informatique, ainsi que la restitution des résultats sous forme de tableaux de bords.

J'occupe ainsi le poste d'assistant chargé de mission pour l'entreprise, bien que mon travail soit en réalité plus proche de la *data* analyse. Je réalise beaucoup de chantiers de Recherche & Développement que nous détaillerons plus tard. Mon rôle est ainsi d'étudier de nouvelles solutions de traitement des données, de la collecte à la visualisation, afin de pouvoir, à terme, les proposer lors de missions clients.

L'objectif de l'étude que nous allons réaliser est dans un premier temps de mieux comprendre les missions d'*Occurrence*, ainsi que ma place au sein de cette entreprise. Mais elle ne se limite pas une approche centrée sur l'entreprise : la finalité de ce projet est de s'intéresser

à un sujet en relation avec mon travail de terrain, tout en gardant la thématique des données comme point central.

Pour réaliser notre étude, nous présenterons dans un premier temps l'entreprise *Occurrence* de façon plus approfondie. Nous verrons en effet l'historique du cabinet, son secteur d'activité, son organisation interne, ainsi que ma place au sein de l'entreprise. Puis nous déterminerons une problématique d'étude concrète répondant aux besoins du cabinet. Dans une seconde partie, nous aborderons cette problématique de façon théorique, en définissant les concepts clés ainsi que les précédents travaux scientifiques conduits sur le sujet. Enfin, la troisième partie de notre étude viendra illustrer les éléments théoriques vu précédemment dans le cadre d'un cas d'application concret en entreprise.

# I. Présentation de l'entreprise et problématique

## **A. Présentation de l'entreprise**

En premier lieu, il convient de présenter de manière plus détaillée le fonctionnement d'*Occurrence*, ses missions et sa place dans son secteur d'activité. Ainsi nous pourrions mieux comprendre mon rôle au sein de cette entreprise et les problématiques qui en découlent.

### **a. Historique de l'entreprise**

Commençons par revenir brièvement sur l'Histoire de cette entreprise, depuis sa création jusqu'à aujourd'hui. Comme mentionné plus tôt, *Occurrence* est un cabinet d'étude et de conseil indépendant, spécialisé dans l'évaluation de la communication. Il est co-fondé par Assaël Adary, philosophe de formation et Benoit Volatier en 1995. L'objectif était alors le même qu'aujourd'hui : faire de la communication une discipline mesurable et quantifiable pour évaluer sa performance.

En 2004, le cabinet reçoit la certification ISO 9001. Il s'agit d'une norme remise par l'Organisation Internationale de Normalisation (ISO) relative au management de la qualité. Elle garantit notamment que l'entreprise respecte un certain nombre d'engagements afin d'améliorer continuellement la qualité de ses services et assure une satisfaction client élevée. Cette certification, gage de qualité, est détenue par *Occurrence* depuis 16 ans et a toujours été renouvelée depuis sa première attribution.

En 2016, Benoit Volatier, co-fondateur d'*Occurrence*, quitte finalement le cabinet pour entreprendre dans de nombreux autres domaines (distribution de lombricompost, accompagnement de randonnées pédestres et fluviales...). Assaël Adary reste alors seul, à la tête de la présidence de l'entreprise. Il est toutefois accompagné de Céline Mas, directrice générale, avant que celle-ci ne parte également en 2018. Assaël Adary est aujourd'hui seul à la tête d'*Occurrence*, en tant que Président Directeur Général.

Le cabinet d'étude se fait ensuite remarquer en 2017 grâce à la mise en place d'une méthode de comptage de foule qui se voulant objective. En effet, outre la communication,

*Occurrence* propose, depuis 2007, des services de comptage de foule. Entre 2007 et 2017, ce service n'a que peu de visibilité et est surtout utilisé par des magasins et centres commerciaux pour le décompte de ses clients. Mais à partir de 2017, le cabinet est engagé par un regroupement de médias (presse, radio, TV) pour effectuer du comptage de foule dans les manifestations. Les chiffres donnés par *Occurrence* se veulent « objectifs » afin de proposer un équilibre entre le comptage annoncé par la police et celui des organisations syndicales.

La technique utilisée vise à placer une caméra filmant le passage de la manifestation à un point donné. Un algorithme trace alors une ligne virtuelle et est capable de compter chaque personne franchissant cette ligne dans le sens de la manifestation. Une marge d'erreur est alors prise en considération et est rectifiée manuellement. Cependant, une polémique a éclaté quant à l'objectivité de la méthode utilisée. Le cabinet a notamment été accusé par certaines organisations de falsifier les chiffres annoncés. Ces accusations ont toutefois été démenties et la véracité des chiffres annoncé a été vérifiée suite à un recomptage manuel lors de certaines manifestations.

En 2021 enfin, *Occurrence* continue de voir croître son chiffre d'affaires, s'élevant à plus de 3,6 millions d'euros. À cette date, le cabinet vient de dépasser la barre symbolique des 1 millions de retombées médiatiques analysées, après 26 ans d'existence.

## **b. Secteur, contexte et domaine d'activité métier**

Intéressons-nous désormais plus concrètement à ce qu'est la société *Occurrence* en tant qu'entreprise, dans un secteur concurrentiel et dans un marché où la demande est forte.

Étant une société de conseil en communication, elle propose à ses clients une double expertise, à la fois en matière d'Études et en matière de communication. La majorité des membres de l'équipe sont en effet des spécialistes de la fonction communication et du marketing, ayant déjà occupé des postes dans ce domaine par le passé.

Le marché du conseil en communication est un secteur particulièrement concurrentiel, si bien que l'on dénombre à Paris plusieurs dizaines de sociétés proposant des services dans ce domaine. Une part importante du travail réalisé chez *Occurrence* est donc la veille et la réponse à des appels d'offre. Dans la plupart des cas, les entreprises souhaitant recevoir une expertise

mettent en concurrence les différents prestataires et opèrent à une sélection après l'étude des différents dossiers de candidature.

Cependant, *Occurrence*, en plus des réponses aux appels d'offre, dispose de contrats avec de gros clients réguliers. Le cabinet est alors engagé pour mener une série d'études sur plusieurs années, sans avoir à repasser par une mise en concurrence. Ces contrats sont ainsi très importants et très précieux car ils assurent une partie du chiffre d'affaires annuel, sans perdre de temps dans la réalisation de dossiers de candidatures.

Pour tenter de se démarquer dans ce marché, *Occurrence* essaye ainsi de proposer une vision différente de ses concurrents, comme mentionné plus haut. Le cabinet désire montrer aux directions d'entreprise que la communication est un investissement sérieux et non un poste de dépense ou une charge. De plus, l'entreprise mise sur la contribution des données dans la stratégie des organisations. Elle estime qu'il est notamment important d'opérer une sélection fine des données à exploiter, en distinguant les données « utiles » et « superflues ». Prendre en compte trop de données est en effet susceptible de complexifier la décision plutôt que de la simplifier.

Une autre part de la légitimité d'*Occurrence* dans ce marché réside dans ses standards. Il s'agit de données de référence, se basant sur toutes les études antérieures réalisées. Elles permettent, lors d'une nouvelle mission auprès d'un client, de lui fournir des éléments de comparaison vis-à-vis des autres entreprises de son secteur d'activité. Ces standards constituent un argument commercial fort. Ils sont en effet un gage de légitimité pour *Occurrence*, prouvant qu'ils sont implantés depuis longtemps dans le marché du conseil en communication. De plus, proposer à une entreprise de comparer ses performances à celles des sociétés concurrentes va lui permettre de mieux identifier ses axes d'amélioration.

Enfin, la certification ISO 9001, évoquée plus tôt et détenue par *Occurrence* depuis 2004 est un dernier gage de qualité permettant au cabinet de se démarquer sur le marché des études. Elle est le signe d'une forte satisfaction client et d'une volonté continue de s'améliorer et d'optimiser ses process en interne. Un audit a ainsi lieu tous les trois ans afin de s'assurer que la société respecte ses engagements en termes de management de la qualité.

### c. Organisation générale

Allons désormais plus en profondeur en nous intéressant à l'organisation générale d'*Occurrence*, ses services, ses métiers, et les missions de chacun.

Notons tout d'abord que malgré ses 26 ans d'ancienneté, la société *Occurrence* reste de taille modeste, puisqu'elle ne compte que 24 membres, stagiaires et alternants compris. Lorsque l'on regarde l'organigramme de l'entreprise<sup>2</sup>, nous pouvons constater qu'elle se divise en trois grands pôles, dont les missions sont distinctes :

- Le pôle « **Marques et Transformation** » : Également appelé pôle « *Corporate* », il est le plus gros service d'*Occurrence*. Il s'agit du service dans lequel je travaille en tant qu'assistant chargé de mission, et dont nous reviendrons sur le fonctionnement détaillé plus tard. Ce pôle est spécialisé dans l'image de Marque auprès des entreprises et des institutions. Il réalise beaucoup d'études de communication interne, de bien-être au travail, mais également de communication externe (publicité, magazine...). Ce pôle est ainsi composé d'un directeur de pôle, de son adjoint, ainsi que d'un directeur des missions. Les clients de ce service sont nombreux et variés, tel que la SNCF, la FDJ, ou encore le ministère de l'Enseignement Supérieur et de la Recherche.
- Le pôle « **Médias et Influences** » : Ce service, composé de 4 membres, est spécialisé dans la stratégie média des entreprises. En clair, il va aider ses clients à améliorer leur stratégie de communication envers les médias et leur image publique. Les solutions proposées sont, par exemple, des bilans d'e-réputation, la création de *Dashboard* médias, des cartographies d'influenceurs ou encore la réalisation de baromètres d'image presse. Aux résultats constatés s'ajoutent des conseils et un accompagnement au pilotage de stratégies médias.
- Le pôle « **Événementiel *Public Policies and Campaigns*** » : En réalité, ce service peut être scindé en deux sous-service distincts. D'un côté la partie « Événementiel », dont le rôle est d'évaluer les actions de communications relatives à la mise en place d'évènements ponctuels. Ce secteur, fortement impacté par la crise sanitaire du COVID19, regagne progressivement son ampleur d'origine. D'autre part, la section

---

<sup>2</sup> Annexe 1 : Organigramme de la société *Occurrence*.

« Europe et International » Est en charge des missions hors de France et est composé d'experts en communication internationale.

Au sein de chacun de ces services, les chargés de mission et leurs assistants sont chargés de collecter et d'analyser les données de l'étude, avant de les restituer sous forme d'un *PowerPoint*. Ils sont constamment pilotés et cadrés par les responsables de service, qui se chargent régulièrement de la restitution des résultats auprès des clients.

Outre ces trois grands pôles, *Occurrence* dispose de trois filiales déjà évoquées. *Occurrence Healthcare* en charge des problématiques santé et bien-être, *Deep Opinion* sur les questions d'Influences en ligne et sur les réseaux sociaux. Ainsi que *Start Cities*, spécialisé dans la concertation lors de grands projets.

Par ailleurs, *Occurrence* fait appel à de la sous-traitance lors de certaines missions, pour répondre à des besoins spécifiques, tel que de la traduction ou encore la réalisation d'infographies. De même, des contrats sont établis avec des « supers sous-traitants ». Il s'agit de travailleurs en freelance, mais travaillant en étroite collaboration et presque à temps plein avec le cabinet d'étude.

Enfin, *Occurrence* dispose d'une secrétaire comptable, une responsable des Ressources Humaines, ainsi qu'un responsable systèmes (Informatique et Qualité). Ce dernier est le seul employé ayant des compétences avancées en informatique et occupe des fonctions diverses. Il est en charge de la gestion et maintenance de l'Intranet, il intervient lors des missions clients afin de programmer des questionnaires en ligne et il effectue également la maintenance informatique générale. Comme nous l'avons mentionné, le cabinet ne dispose pas de *Data analysts* pour exploiter les données. Ce travail est à la charge des chargés de missions, qui utilisent un logiciel de traitement appelé « *Modalisa* ».

Ainsi, *Occurrence* étant une petite structure et ne disposant pas d'un service informatique à proprement parlé, la structure du Système d'Information du cabinet est relativement simple. En somme, l'ensemble des données de l'entreprise sont regroupées sur un seul serveur central. Celui-ci est toutefois réparti en plusieurs lecteurs :

- **Missions en Cours** : Tous les documents relatifs aux études en cours.
- **Clients Archives** : Tous les documents relatifs aux études passées.
- **Communication Occurrence** : Documents liés à la communication externe d'*Occurrence*.

- **Développement** : Il recouvre la promotion d'*Occurrence*, les présentations produits...
- **Ressources Administratives** : Regroupe les documents relatifs aux fonctions administratives (accès restreint).
- **Ressource Production** : Tous les documents nécessaires pour le travail quotidien en interne. Il peut s'agir de formations, de modèles de documents, de notices d'explications...
- **Qualité** : Tous les documents nécessaires pour appliquer la démarche qualité.

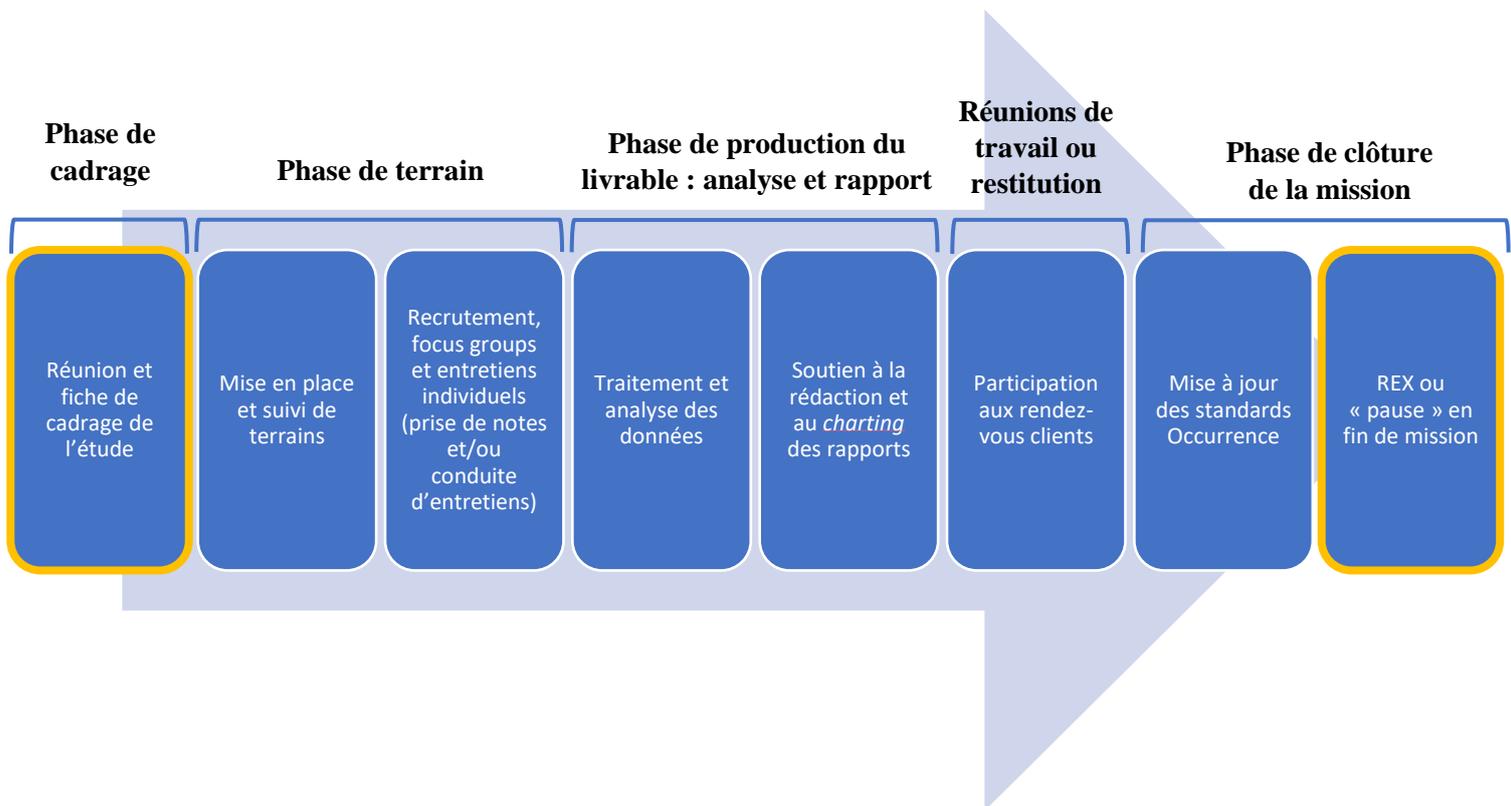
Chaque poste local est ainsi rattaché à ce serveur et l'utilisation de VPN a été mise en place pour l'accès aux documents à distance, en télétravail par exemple.

#### **d. Focus sur le service « Marque et Transformation »**

Afin de mieux comprendre mes missions et mon rôle au sein de la société, commençons par présenter succinctement le service « **Marque et Transformation** » (M&T), dans lequel je travaille.

Ce pôle est spécialisé dans la communication dite *corporate*, c'est-à-dire des études et baromètres de notoriété et d'image, de lectorat, des audits de dispositifs de communication externe, des post-tests publicitaires, etc. Mais il est également spécialisé dans la communication interne, grâce à des études et baromètres d'opinion interne, de climat social, de bien-être au travail...

Comme indiqué, le travail consiste donc majoritairement à réaliser des études pour des clients. Une étude se déroule donc en plusieurs phases, de la façon suivante :



Plus concrètement, les apprentis chargés de mission sont en charge de la partie opérationnelle de l'étude. Ils interviennent notamment lors de la phase de production, en analysant les données collectées et réalisent le *charting*, c'est-à-dire la charte graphique du rapport final, ainsi que l'intégration des données. Les chargés de missions et les responsables du pôle s'occupent quant à eux de l'ensemble de la relation client, de la mise en place du terrain à la restitution finale.

D'un point de vue organisationnel, le pôle Marques et Transformation dispose de deux outils indispensables à son fonctionnement interne. En premier lieu, un plan de charge est mis en place, commun à l'ensemble du service. Il s'agit d'un tableur *Excel* permettant d'avoir une vision à long terme de la charge de travail des membres de l'équipe. L'objectif étant de lisser cette charge afin d'éviter une surcharge à un moment donné. Il s'apparente à un planning de la production assez sommaire où l'on y renseigne le client pour lequel on travaille à un moment donné, ainsi que les jours de congés, etc.

Le second outil indispensable est le *Trello* du pôle. Il s'agit d'un outil de gestion de projet en ligne, basé sur une organisation des phases de production en planches listant des cartes, chacune représentant des études en cours. Chaque mission est alors rangée dans la phase

de production correspondant à son état d'avancement : développement commercial, cadrage, production, création du livrable, en *standby*...

Ces outils sont utilisés et mis à jour lors de la réunion de service hebdomadaire, ayant lieu tous les mardis matin. Cette réunion est l'occasion de faire le point sur l'état d'avancement de chaque membre de l'équipe, et de faire parvenir les nouvelles à l'ensemble du pôle, tels que les nouveaux chantiers « gagnés » par exemple.

## **e. Présentation de mes missions et projets de *Business Intelligence***

Nous avons ainsi pu mieux comprendre le fonctionnement d'*Occurrence*, et plus particulièrement du pôle Marques et Transformation auquel j'appartiens. Nous avons également vu les missions qu'occupent traditionnellement les apprentis chargés de mission, dont j'occupe la fonction. Cependant, mes missions au sein d'*Occurrence* ne se limitent pas à la réalisation d'études pour des clients. En effet, ayant été recruté pour ma formation en analyse de données, j'occupe un rôle bien spécifique au sein de ce pôle.

En somme, notamment encadré par Thomas Skorucak, directeur adjoint du pôle M&T et par Johann Damen, responsable informatique d'*Occurrence*, je réalise principalement des missions de Recherche et Développement (R&D). Mon travail consiste alors à explorer de nouvelles solutions en matière de collecte, d'analyse et de visualisation des données afin de mettre en place, à terme, de nouveaux outils aux services des missions clients.

### **1. *Power BI* et datavisualisation**

Ma première mission de longue durée portait sur le logiciel *Microsoft* « *Power BI* ». Il s'agit d'un outil de création de rapport et de tableaux de bords dynamiques et interactifs particulièrement puissant. Depuis longtemps, *Occurrence* cherchait à pouvoir exploiter ses données à travers ce logiciel et en explorer les possibilités.

Je me suis donc formé à son utilisation, en réalisant une première série de tableaux de bord « tests ». Mais je me suis vite aperçu que *Power BI* n'est pas un simple outil de datavisualisation. Il permet, en outre, de manipuler des jeux de données complexes à l'aide de requêtes sur l'outil *Power Query*. Il permet également de gérer des relations entre les tables

d'une base de données relationnelle, ou encore de réaliser des calculs conditionnels sur les données à l'aide du langage *DAX*.

Puis je me suis ensuite intéressé, durant plusieurs semaines, aux solutions de partage et d'actualisation automatique des données. En effet, dans l'idée d'une mise en application auprès de clients, il est nécessaire de proposer des tableaux de bords s'actualisant automatiquement. Ainsi, il suffit à l'utilisateur de renseigner ses données depuis la source de son choix (*SQL*, *Excel*, *Excel Online*, *Sharepoint*...) pour que les graphiques s'actualisent.

Encore en développement, la principale limite aujourd'hui rencontrée est la question de la sécurité des données. Si le client ne dispose pas d'un abonnement *Microsoft* adapté, ses données seront publiques et non sécurisées.

## 2. *Travail sur les standards*

Déjà présentés plus hauts, les standards représentent des données de références, calculés à partir des résultats des études antérieures réalisées. Ils constituent ainsi une source d'éléments comparatifs pour les clients futurs, quelque soient le type d'étude désiré. Ces standards posent toutefois des problèmes en interne, car chaque pôle dispose de ses propres normes, modalités et modes de saisie. Ainsi, ces données manquent de cohérence et sont ainsi difficiles à croiser et, parfois, à exploiter.

Depuis début 2021, un chantier global sur les standards a été ouvert, afin d'unifier les données, de réfléchir à leur harmonisation, à de nouveaux modes de saisie (actuellement stockés sur des tableaux *Excel*...). Prenant part à ce chantier, je participe à ce mouvement de réflexion collectif. Je me suis par exemple formé à un autre outil *Microsoft*, intitulé *Power Apps*, permettant de développer, grâce à un logiciel *low code*, des applications de formulaires.

## 3. *Veille sur les outils de collecte*

De nouveau dans l'optique d'explorer des solutions permettant d'ajouter une plus-value aux études d'*Occurrence*, je réalise depuis peu une veille sur les outils de collecte et d'analyse de la donnée. Notamment intéressé par les outils de *webscraping* et de *scraping* sur les réseaux sociaux, je m'entretiens avec les différents pôles afin d'identifier clairement leurs besoins et

attentes dans ce domaine. À l'avenir, j'effectuerais des tests sur les outils et packages identifiés, afin de lister les possibilités techniques de chacun, avant d'envisager une mise en application concrète.

#### 4. Participation aux études clients

Bien qu'occupant un poste spécifique, je participe également aux tâches plus traditionnelles des assistants chargés de mission. En effet, j'ai déjà réalisé plusieurs *charting*, c'est-à-dire la mise en forme du rapport qui sera restitué au client à la fin de l'étude. Il s'agit d'une mission nécessitant de la créativité, car les employés sont laissés libre de valoriser les données comme ils le souhaitent, du moment que la charte graphique est respectée. De plus, ce travail requiert une réflexion plus « technique » sur la datavisualisation en général, afin de déterminer quels sont les graphiques les plus adaptés en fonction de la nature des données collectées.

Je vais par ailleurs prochainement être formé à l'outil *Modalisa*, utilisé chez *Occurrence* pour l'analyse des données issues de questionnaires. Ce logiciel permet notamment de croiser les données et les différentes modalités de réponse en fonction des profils de répondant, et de mettre en valeur les informations significatives.

## **B. Développement de la problématique**

Nous venons de voir quelles sont les missions de la société *Occurrence* ainsi que le rôle que j’y exerce. Nous avons ainsi pu identifier plus précisément les besoins et les enjeux de cette entreprise. Aussi, nous allons aborder dans cette seconde partie le processus de développement de notre problématique de recherche et les angles de recherche envisagés.

S’inscrivant dans le cadre d’un travail de recherche autour d’une problématique métier, cette étude doit ainsi répondre à des critères précis. D’une part, elle doit être en relation directe avec mes activités chez *Occurrence* et doit répondre à un besoin de la part de cette dernière. Par ailleurs, la question des données doit être au cœur de notre étude car cette dernière s’intègre dans le champ plus large de l’analyse *data* à des fins sociales, c’est-à-dire pour une exploitation humaine directe. Au sens large, il peut alors s’agir d’une problématique de collecte, de traitement ou encore d’automatisation des données de l’entreprise.

Afin de déterminer notre sujet d’étude, identifions les besoins de l’entreprise en matière d’exploitation des données, les moyens et les supports déjà existants, ainsi que les enjeux de déploiement de nouveaux outils.

### **a. Contexte *data* de l’entreprise**

Intéressons-nous en premier lieu au niveau de maturité d’*Occurrence* en matière de traitement et de gestion des données. Tel qu’employé, le terme de « maturité » désigne selon la *Digital League* d’Auvergne-Rhône-Alpes « les moyens, connaissances et actions visant à améliorer la performance globale de l’entreprise à l’aide des données. »<sup>3</sup>. Au sein de notre entreprise, nous allons élargir cette définition afin d’y inclure les outils et méthodologies de traitement des données dans le cadre des missions d’études et de conseils réalisés pour les clients.

Ainsi, dressons un bilan de l’existant afin d’évaluer, suivant différents critères choisis, le niveau de maturité *data* d’*Occurrence*. D’une part, au regard des données à disposition et

---

<sup>3</sup> Digital League Rhône-Alpes, « La maturité Data, c’est quoi ? ». URL : <https://www.digital-league.org/blog/2021/05/26/la-maturite-data-cest-quoi/#:~:text=La%20maturit%C3%A9%20data%20d'une,m%C3%AAme%20en%20fonction%20du%20secteur>

leur exploitation, les outils utilisés, mais également en matière de compétences, de culture *data* et de gestion de la qualité des données en interne.

### 1. Type de données exploitées

Les données exploitées chez *Occurrence* sont variées, car il s'agit à la fois d'éléments internes servant au pilotage de l'activité et de données externes collectées et analysées dans le cadre de ses études clients. N'oublions pas que le parti pris de l'entreprise est d'évaluer l'impact des actions de communications de ses clients à l'aide des données. Plus encore, il s'agit même d'un argument de vente important afin de se démarquer de ses concurrents dans le secteur des études et du conseil. Par conséquent, observons les différents types de données exploitées par le cabinet :

- Les **données d'enquête** : Il s'agit de données collectées lors d'une enquête interne ou externe réalisée pour un client. Elles se divisent en deux catégories :
  - Les données d'**études quantitatives** : Il s'agit d'enquêtes classiques, où les données résultent de la diffusion d'un questionnaire auprès d'un groupe d'individus. Ces dernières sont alors relativement structurées, puisqu'elles sont cadrées par les modalités de réponses proposées aux répondants. Il peut s'agir de questions à choix unique (QCU), à choix multiples (QCM), ou encore de questions ouvertes (QO) qui peuvent être numériques ou textuelles. Par conséquent, le traitement de ces données est réalisé à l'aide d'outils numériques dédiés sur lesquels nous reviendrons.
  - Les données d'**études qualitatives** : Il s'agit d'enquêtes sur un nombre restreint de répondants. Celles-ci sont réalisées sous forme d'interviews pouvant aller jusqu'à plusieurs heures et sont cadrées par un guide d'entretien. Les échanges sont alors retranscrits sous différentes formes : prise de note, vidéo... De fait, aucun outil informatique n'est employé pour accompagner leur analyse. La variation des supports de l'information et l'interprétation sémantique des propos retranscrits sont très difficilement assimilables par une machine. Les professionnels du conseil chez *Occurrence* se chargent donc d'analyser et de synthétiser manuellement ces données.

- Les **données de réseaux sociaux** : Le pôle « Média et Influence » d'*Occurrence* et sa filiale *Deep Opinion* traitent les données de réseau sociaux pour observer l'e-réputation de nos clients à l'aide d'outils de *reporting* et de requêtage.
- Les **données du web** (observatoires) : *Occurrence* réalise également des observatoires. Il s'agit d'outils de veille et de collectes sur des sujets spécifiques, à l'image de l'observatoire des pétitions, dédié à la surveillance quotidienne de tous les grands sites de pétitions en ligne<sup>4</sup>. Pour cela, nous disposons d'algorithmes de *scrapping* afin de collecter les données issues du web sur le sujet et de les stocker dans des bases internes, tout en veillant au bon respect du Règlement Général sur la Protection des Données (RGPD).
- Les **données clients** : Afin d'assurer ses services de conseil en communication, l'entreprise analyse un certain nombre de données produites par les organismes clients. Il peut s'agir de bases de données comptables, de documents de travail, de pilotage, etc. Ces données ne font généralement pas l'objet d'une analyse statistique approfondie et sont donc explorées manuellement par les experts conseils.
- Les **bases de données de contacts client** : Lors de certaines études, nous devons disposer de bases de contact afin de joindre les répondants ciblés par l'enquête. Pour cela, le client fournit parfois une base de contact dont il dispose déjà. Cela est notamment le cas lors de campagnes de phoning (appels téléphoniques) ou d'e-mailing par exemple. Ces données sont exploitées mais ne servent uniquement à joindre les répondants, elles ne font donc pas l'objet d'une analyse spécifique.
- Les **données internes** : Cela désigne toutes les données produites par *Occurrence* et permettant le pilotage de l'activité. Il peut s'agir de données comptables, administratives, mais également de suivi de l'activité interne. Ces dernières données font l'objet de traitements à l'aide d'outils dédiés.

Paradoxalement, nous pouvons constater que si, comme tout organisme, *Occurrence* génère nécessairement de la donnée, ce cabinet est majoritairement porté sur le traitement de données externes. En effet, s'il met en place une évaluation pointue de la communication de ses clients, à l'aide d'indicateurs clés (*KPI*), les mêmes mécanismes ne sont pas appliqués au sein de l'institut. Certains indicateurs et outils de pilotage sont certes mis en place afin d'assurer le

---

<sup>4</sup> Observatoire des pétitions par *Occurrence*. URL : <https://observatoirepétitions.fr/>

suivi de l'activité, mais ceux-ci visent plutôt à garantir le respect des objectifs annuels que d'améliorer concrètement les performances.

## 2. Outils et pilotage

Afin de poursuivre l'évaluation du contexte *data* de l'entreprise, intéressons-nous désormais aux outils de pilotage et de traitement des données à disposition. Nous venons en effet de souligner le fait qu'*Occurrence* dispose de solutions d'évaluation de son activité. Bien que parfois sommaire, cherche des moyens d'exploiter les données internes à disposition afin d'en générer de la valeur pour l'entreprise. Ainsi, passons en revue les principales méthodes de pilotage et de gestion des données :

- **L'intranet** : Il s'agit probablement de l'outil de suivi le plus utilisé et le plus automatisé. Ce site web interne a été développé par le Directeur du Système d'Information (DSI) en langage PHP. Il n'a pas vocation à être une plateforme de communication interne et rempli uniquement un rôle fonctionnel. Entre autres, il permet à chaque salarié d'indiquer chaque semaine le temps qu'il a consacré à ses différentes missions. Les données saisies sont alors stockées dans une base de données relationnelle et tous les temps inscrits pour une même tâche sont alors combinés. A cela s'ajoute le montant facturé au client pour cette mission afin de générer un indicateur « CVP », soit un indice de rentabilité en fonction du temps consacré. L'intranet permet par ailleurs de gérer le taux de sous-traitance de l'entreprise.
- **Trello** : Déjà évoqué plus haut, cet outil exclusivement utilisé par le pôle « marque et transformation » permet un suivi des missions en cours, des échéances et des différentes tâches à réalisées. Les nouvelles données et les mises à jour se font de façons hebdomadaires lors de la réunion de pôle. Bien connu et utilisé par de nombreuses entreprises, il s'agit d'un outil de pilotage professionnel plébiscité pour sa flexibilité et sa simplicité. Cependant, il ne s'agit pas réellement d'un outil d'analyse ou de traitement de données et son utilisation chez *Occurrence* se limite aux fonctionnalités de base. Notons par ailleurs que chaque pôle dispose d'un outil de pilotage similaire à celui-ci, mais la pratique n'est pas harmonisée au sein de l'entreprise.

- Le **tableau des missions en cours** : Base de données *Excel* regroupant le *listing* des missions en cours de réalisation. Composé de nombreuses variables, ce tableau permet notamment de piloter les échéances de facturation par pôle et de dresser une prévision mensuelle du chiffre d'affaires réalisé.
- Le **tableau de bord du développement** : Également sur *Excel*, cette base ressemble à celle des missions en cours, bien que plus globale. Elle permet le recensement de toutes les missions passées et en cours par pôle, y compris les missions ayant fait l'objet d'un Appel d'Offre (AO) qui a été perdu par *Occurrence*. Ce tableau est un outil central au service des managers et permet le calcul de nombreux *KPI* de l'entreprise, tel que l'espérance en chiffre d'affaires de l'année en cours. Les objectifs annuels fixés en début d'activité sont, pour la plupart, déterminés à l'aide des indicateurs issus de ce tableau.

Nous pouvons ainsi constater qu'*Occurrence* dispose bien d'un certain nombre d'outils de pilotage de son activité sous divers aspects, à travers des données saisies et des indicateurs concrets. Cependant, ceux-ci ne sont pas toujours harmonisés au sein de l'entreprise, à l'image du logiciel *Trello*, utilisé par un seul pôle. De même, ils ne sont pas interconnectés et il n'existe pas de flux de données entre les outils au sein du Système d'Information (SI). Ceci est un signe fort d'un niveau de maturité *data* encore jeune au sein du groupe. En effet, les sources de données, notamment *Excel*, font également office d'outils de pilotage, ne facilitant pas une approche globale de la performance.

### 3. Compétences et culture data

Intéressons-nous à un nouveau critère d'évaluation de notre maturité *data* : les compétences *datas* et la culture *data* en interne. Si les compétences évaluent la capacité des salariés à traiter et analyser des données, la notion de « culture *data* » désigne la conscience des enjeux liés à une bonne exploitation des données et de ses bénéfices pour les performances de l'entreprise. Les équipes de développement du logiciel Tableau, outil majeur en matière d'informatique décisionnelle, en donne la définition suivante : « La culture des données, c'est la somme des valeurs et des comportements communs qui encouragent l'utilisation des données pour améliorer la prise de décisions. »<sup>5</sup>. Il y a donc une véritable idée de conscience collective,

---

<sup>5</sup> Tableau, « Culture des données ». URL : <https://www.tableau.com/fr-fr/why-tableau/data-culture>

la culture des données n'est pas uniquement l'affaire des décideurs ou des responsables SI, mais bien de tous les acteurs de l'entreprise.

En matière de compétences tout d'abord, nous avons déjà souligné le fait que les prestations proposées par *Occurrence* impliquent, pour la plupart, du traitement et de l'analyse de données :

- Les **études quantitatives** impliquent le traitement des bases de données d'enquête, puis leur interprétation au sein d'un rapport.
- Les **tableaux de bords** réalisés pour les clients (généralement sur *Excel*), impliquent la mise en place d'espace de saisie des *KPI* et leur projection sur une *scorecard* dynamique.
- Certains pôles, comme le pôle « média et Influence » et la filiale *Deep Opinion* réalisent également de l'**analyse média**, issu de la presse ou de contenus numériques tels que les données de réseaux sociaux.

Par conséquent, en termes de compétence *data*, tous les collaborateurs sont en mesure de réaliser des analyses sur des données de communication après nettoyage d'une base de données. De plus, presque tous savent manipuler un ou plusieurs outils spécialisés de traitement des données : *Modalisa* pour le traitement de données d'études quantitatives, ou encore *Talkwalker* pour le traitement de données de réseaux sociaux. Enfin, quelques collaborateurs disposent de compétences avancées sur *Excel* et tous savent l'utiliser à minima de façon sommaire. Nous ne sommes donc pas au niveau zéro en termes de compétence *data* chez *Occurrence* et les salariés sont conscient des enjeux et du potentiel d'une bonne exploitation des données.

Cependant, il s'agit majoritairement de compétences spécifiques dédiées à la réalisation des missions à destination des clients de l'entreprise. Nous l'avons vu plus tôt, ces outils ne sont pas employés ni adaptés en interne à des fins d'amélioration des performances. Ainsi, il s'agit de compétences *datas* opérationnelles mais ces pratiques ont du mal à être exploitées en interne pour diverses raisons, à commencer par la difficulté à déployer une culture de la donnée commune.

En somme, en termes de culture *data*, si l'intégralité des employés est sensible à l'impact de la donnée et des apports pour l'entreprise, cette culture ainsi que les pratiques ne sont pas

harmonisées. Nous l'avons dit, les différents pôles utilisent leurs propres outils de traitement et leurs propres process de réalisation des missions qui leur sont confiés. Bien que des tentatives soient mises en place, telle qu'une formation interne à destination de tous les nouveaux arrivants, les bonnes pratiques en matière de qualité des données sont parfois négligées ou inégales par manque de sensibilisation collective. Toutefois, soulignons le fait qu'il ne s'agit pas d'une difficulté de gouvernance des données. Si les process peuvent encore être perfectibles, les grands principes de disponibilité, d'usabilité, d'intégrité et de sécurité des données<sup>6</sup> sont respectés et institués de longue date. Il s'agit donc plutôt d'un travail à long terme de communication et de partage des bonnes pratiques à l'ensemble des collaborateurs.

#### **4. Qualité des données**

Pour terminer notre audit de maturité *data*, abordons un dernier axe d'évaluation : la qualité des données. Il s'agit d'une notion très large englobant, d'une part la qualité des bases de données de sorte à ce que celles-ci soient le plus simplement exploitables et robustes afin d'éviter de potentielles erreurs de traitement. D'autre part, le concept de qualité des données englobe également l'ensemble des dispositifs et des bonnes pratiques mises en œuvre afin d'assurer la fiabilité des bases de façon pérenne.

Aussi, pour comprendre la question globale de la qualité des données chez *Occurrence*, il faut prendre en compte le contexte de l'entreprise. Il s'agit d'une PME, ne disposant pas des moyens techniques et financiers pour déployer un logiciel de gestion de la qualité des données tel que *Talend*<sup>7</sup> par exemple. De plus, nous l'avons dit, les équipes sont conscientes des enjeux et des apports de l'exploitation des données sur la performance, mais elles ne maîtrisent pas véritablement les besoins en matière de qualité que cela implique.

Notons également que nous disposons d'un SI « à l'ancienne », composé de 8 serveurs locaux, sans processus automatisé d'acheminement des données, d'*ETL (Extract – Transform – Load)* ou de *data cleaning*. Il s'agit en effet du contrecoup d'une transition numérique de l'activité réalisés relativement tôt. L'architecture du système d'information n'avait alors pas été imaginé selon les problématiques et les technologies actuelles, notamment en matière de *Cloud*

---

<sup>6</sup> R. Gaetan, « Data Governance ou gouvernance des données : Qu'est-ce que c'est ? », Lebigata.fr. URL : <https://www.lebigdata.fr/data-governance>

<sup>7</sup> Site de Talend : <https://www.talend.com/fr/>

*computing*. Cette architecture est désormais ancrée dans pratiques collectives de l'entreprise, si bien qu'il est extrêmement délicat d'envisager aujourd'hui une restructuration globale.

Cependant, *Occurrence* dispose d'une démarche qualité globale bien aboutie et régulièrement renouvelée. Le cabinet est détenteur de la certification ISO9001 depuis 2004. Il s'agit d'une norme relative au management de la qualité globale de l'entreprise. Pour la renouveler, cette dernière se soumet donc à un audit de qualité tous les 3 ans. Bien sûr, la « Qualité » au sens large englobe de nombreux champs, mais elle inclut celui des données. Un référent qualité a donc été nommé au sein de l'institut, à qui l'on peut se référer pour toute question à ce sujet. Celui-ci étant également responsable SI d'*Occurrence*, il accorde une grande importance à promouvoir une démarche qualité concernant les données.

Toutefois, outre le contexte historique de l'entreprise limitant les possibilités techniques, d'autres obstacles rendent difficile l'instauration d'un bon processus de qualité des données. D'une part, car comme nous l'avons mentionné plus tôt, il n'y a pas de processus de traitement réellement uniformisé entre les pôles, mais également au sein d'un même pôle. Ainsi, chaque chargé de mission a ses propres méthodes de traitement, entraînant *in fine* des disparités de forme entre les bases de données. D'autant qu'une même étude est parfois assurée par deux ou trois chargés de missions en fonction des différentes étapes de production. L'un peut réaliser le nettoyage initial de la base, un second les traitements statistiques et le dernier procéder à l'analyse des résultats. De fait, il est nécessaire d'assurer la qualité malgré ces changements de main, ce qui est parfois complexe en l'absence d'un processus uniformisé.

D'autre part, si comme expliqué plus haut, l'ensemble des collaborateurs dispose des bases de la bureautique et savent utiliser certains logiciels spécialisés, la *pool* de compétences informatiques générales reste assez faible. En effet, la majorité des salariés sont en premier lieu des spécialistes de la communication, de la sociologie et des études, qui se sont adaptés et formés aux outils nécessaires pour leur activité. Ainsi, *Occurrence* ne dispose pas de véritables spécialistes de la donnée, en mesure de déployer des modèles statistiques avancés ou de manipuler des jeux de données à l'aide de langages de programmation dédiés tels que *R* ou *Python*. Par conséquent, les nettoyages et les traitements ne sont pas automatisés, pouvant entraîner une perte de qualité et de temps non négligeable.

Enfin, bien qu'il existe des outils à disposition des collaborateurs, tels qu'un guide des bonnes pratiques, des processus qualité déployés, ou encore des espaces de partage des savoirs, ces solutions ne sont que faiblement utilisées. La faible communication sur l'existence de ces

ressources accessibles à tous a en effet entraîné cette sous-exploitation. De plus, la période COVID entre 2020 et 2021 a fortement impacté les temps d'échanges et de partage des compétences déployées jusqu'alors. Or, la question de la qualité des données est l'affaire de tous et non uniquement des équipes informatiques, comme publié par *Gartner* en 2007<sup>8</sup>.

## 5. Constats et objectifs

Maintenant que nous avons réalisé un état des lieux général du niveau de maturité de l'entreprise en matière de données selon plusieurs critères, dressons le bilan et les conclusions de notre évaluation. L'Observatoire de la Maturité *Data* des Entreprises<sup>9</sup> a établi une échelle numérotée de 0 à 6 afin de graduer le niveau de maturité d'une entreprise. Le grade « zéro » indique l'absence d'outil numérique pour la gestion des données de l'entreprise. À l'inverse, le grade le plus haut – six – signifie qu'une entreprise dispose de spécialistes de la donnée, d'une Gouvernance *Data* solide, d'un système d'information entièrement interconnecté et multidirectionnel, ainsi que d'outils de traitement avancés telles que des solutions de *machine* et *deep learning*. Bien sûr, toutes les entreprises n'ont pas les moyens ni le besoin d'atteindre ce dernier palier, s'adressant principalement aux grands groupes. Mais l'objectif reste toutefois de s'en rapprocher car tout organisme est générateur de données exploitables dans son pilotage et ses performances.

Ainsi, sur cette échelle, nous pouvons placer *Occurrence* entre les niveaux 1 et 2. En effet, bien que l'entreprise dispose de quelques outils de traitement de la donnée, le pilotage est majoritairement réalisé à l'aide de bases *Excel* indépendantes. De même, une culture *data* commune et de bonnes pratiques en matière de qualité des données sont encore en voie de développement et peu institutionnalisées. Toutefois, un pilotage informatisé par processus, ainsi que la présence d'un Directeur SI viennent rehausser le bilan général.

Au regard de ce constat, l'objectif de mon étude serait de parvenir à rehausser le niveau de maturité d'*Occurrence* et de l'orienter vers un « Pilotage évolué (*BI*) ». Il s'agit du niveau 3 au sein du barème mis en place par l'Observatoire<sup>10</sup>. Les niveaux supérieurs impliqueraient un remaniement complet de la structure du SI de l'entreprise dont cette dernière n'a ni les moyens

---

<sup>8</sup> Gartner, « « Dirty data » is a business problem, not an IT problem, says Gartner », 2007. URL : <https://www.gartner.com/newsroom/id/501733>

<sup>9</sup> Site de l'Observatoire de la Maturité *Data*. URL : <https://observatoire-data.fr/>

<sup>10</sup> *Op cit.* Site de l'Observatoire de la Maturité *Data*.

techniques et financiers, ni le besoin à court terme. Ainsi, ce troisième niveau consiste à déployer et démocratiser l'usage d'outils de *Business Intelligence (BI)*, tant pour le pilotage interne que lors des missions clients. En somme, il s'agit d'outils facilitant la compréhension des données à disposition et permettant d'éclairer les prises de décision stratégiques des décideurs de l'entreprise.

Pour *Occurrence*, le déploiement d'outils *BI* ne se limiterait d'ailleurs pas à un usage interne. En effet, l'un des grands enjeux du cabinet est d'assurer la gestion de données d'organismes bien plus gros et structurés que lui en matière de *data*, en particulier pour les missions de suivi et d'accompagnement de l'activité. L'instauration et la démocratisation de la *BI* en entreprise nécessite également de parvenir à adapter ces nouveaux outils aux données clients afin de générer un gain de performance et une plus-value auprès des partenaires.

### **b. Enjeux des données et de la *Business Intelligence***

Essayons désormais de comprendre quels sont les enjeux d'une bonne exploitation des données et du déploiement des outils de *Business Intelligence*, notamment au regard du secteur d'activité de l'entreprise. Le monde des études et de la communication possède en effet ses propres besoins et contraintes métiers que nous détaillerons dans un second temps.

En premier lieu, notons que, toujours selon l'Observatoire de la Maturité *Data* des Entreprises, « 80% des entreprises interrogées pensent pouvoir améliorer leur performance grâce aux données. »<sup>11</sup>. Ce chiffre illustre bien le changement de paradigme s'étant propagé ces 15 dernières années. Les données sont génératrices de valeur, dont la bonne exploitation peut se traduire par une croissance économique, un gain de temps sur de nombreux processus, ainsi que des décisions plus avisées en matière de pilotage.

La *BI*, ou « informatique décisionnelle » en français est donc la discipline visant à encadrer la collecte, le traitement, l'analyse et la restitution des données afin d'en extraire de la valeur. Dans leur article « *Business Intelligence & analytics in management accounting research : Status and future focus* », publié en 2018 dans l'*International Journal of Accounting Information System*, P. Rikhardsson et O. Yigitbasioglu la définissent comme : « la technologie qui facilite la collection des données, son analyse et la transmission de l'information, et qui la

---

<sup>11</sup> *Ibid.*

met en forme pour être un support à la prise de décision. »<sup>12</sup>. L'un des grands enjeux de la *BI*, quel que soit le secteur d'activité, est donc de déployer des outils d'interprétation. C'est-à-dire, de constituer une passerelle entre les sources de données et les décideurs de l'entreprise, afin de les rendre compréhensibles, génératrices de nouveaux questionnements et de nouvelles décisions.

Pour *Occurrence*, un premier enjeu en matière de performances est celui de la centralisation des données internes. Nous l'avons dit, l'entreprise ne dispose pas d'un *Datawarehouse* regroupant l'ensemble des bases opérationnelles, notamment constituées sur *Excel*. Or, hormis une restructuration complète de son système d'information, l'entreprise n'est pas en mesure de déployer rapidement un tel dispositif. Les outils *BI*, peuvent, par nature, apporter une solution à ce problème. En effet, nous y reviendrons, ceux-ci sont conçus pour pouvoir se connecter à de nombreuses sources externes et d'importer les données dans un même espace afin de pouvoir procéder à une analyse multi sources.

Plus qu'un simple regroupement, les outils d'informatique décisionnelle sur le marché permettent également d'harmoniser les traitements et d'établir des relations entre les bases. Des *pipelines* et des processus *ETL* peuvent être paramétrés afin de consolider la qualité lors de l'import des données, de sorte que celles-ci soient directement exploitables suite à leur acheminement.

Par ailleurs, rappelons que les outils *BI* sont également des logiciels de restitution des données. Cela passe notamment par la construction de tableaux de bord et de rapports permettant le suivi des indicateurs de performance les plus impactant à travers un ensemble de graphiques et de chiffres clés. Or *Occurrence* réalise également des missions de création de tableaux de bord de la communication pour ses clients. Une bonne datavisualisation est un élément primordial afin de mettre en valeur les informations les plus importantes et d'engendrer de la satisfaction client. Plus flexibles et interactifs que des tableaux de bord *Excel*, le déploiement d'outils *BI* constitue donc une nouvelle offre pour le cabinet. Tout en restant dans son domaine de spécialité (l'évaluation de la communication), ces solutions sont susceptibles d'attirer une nouvelle clientèle et, par conséquent, de générer du chiffre d'affaires.

---

<sup>12</sup> RIKHARDSSON Pall et YIGITBASIOGLU Ogan, « *Business Intelligence & analytics in management accounting research: Status and future focus* », dans *International Journal of Accounting Information Systems*, vol. 29, 2018, pp. 37-58 [en ligne]. URL : [https://econpapers.repec.org/article/eeeijjoais/v\\_3a29\\_3ay\\_3a2018\\_3ai\\_3ac\\_3ap\\_3a37-58.htm](https://econpapers.repec.org/article/eeeijjoais/v_3a29_3ay_3a2018_3ai_3ac_3ap_3a37-58.htm)

De façon pragmatique, la génération de profit est en effet l'enjeu principal du déploiement de ces solutions par l'entreprise. Une PME ne dispose pas des fonds nécessaires pour investir dans la Recherche et Développement (R&D) sans espérer un retour sur investissement. En l'*Occurrence*, les outils de *Business Intelligence* sont au goût du jour et une demande commence à naître progressivement sur le marché des études. Il s'agit alors de pouvoir rester concurrentiel lors de réponses à des appels d'offre, bien que la demande du client ne reflète pas toujours ses besoins réels qui ne nécessiteraient pas l'usage d'outils *BI*.

### **c. Définition de la problématique d'étude**

Nous cernons désormais mieux les enjeux relatifs au déploiement de l'informatique décisionnelles, tant en matière de gouvernance des données que pour les avantages financiers qui en découlent. *Occurrence* avait d'ailleurs déjà envisagée avant mon arrivée au sein de l'entreprise l'utilisation d'un outil *BI*. Certains tests, bien que peu concluants dans un premier temps, avaient même déjà réalisés avec l'envie de creuser le sujet plus en profondeur. Il s'agit ainsi d'un besoin explicitement exprimé par le cabinet et pour lequel le recrutement d'un apprenti *Data analyst* en janvier 2021 fait partie intégrante du processus. C'est donc naturellement que j'ai entrepris de porter mon étude sur cette thématique.

Toutefois, la *BI* est un sujet vaste qu'il nous faut délimiter et intégrer au cadre spécifique de son exploitation par une société telle qu'*Occurrence*. Ainsi, bien que nous reviendront brièvement sur les origines de l'informatique décisionnelle au sein d'une partie plus théorique, ne nous attardons pas sur des bornes chronologiques ou géographiques. Nous allons plutôt restreindre notre analyse à un secteur d'activité spécifique. En effet, nous allons nous concentrer sur l'exploitation et l'apport de la *Business Intelligence* dans le secteur des études et de la communication. Précisons que si nous présenterons un cas de déploiement concret au sein de la société *Occurrence*, notre étude se veut plus englobante. Nous souhaitons ainsi mettre en lumière les besoins et les usages spécifiques à ce secteur, mais également les sources de données privilégiées et les principaux *KPI* de communication.

Il existe une multitude d'outils d'informatique décisionnelle car cette discipline englobe un champ relativement large aux frontières parfois floues : logiciels de datavisualisation divers, de manipulation des données, de question de la qualité, logiciels hybrides, etc. De fait, nous ne pourrions pas dans cette étude nous montrer exhaustif sur les spécificités techniques de chacun,

d'autant que cela présenterait un intérêt scientifique relativement faible. Ainsi, pour des raisons pratiques, nous évoquerons un certain nombre d'outils *BI* déployés chez *Occurrence* durant un an et demi d'étude, mais nous nous concentrerons sur le logiciel *Microsoft Power BI*. Leader du marché dans ce domaine, cette solution hybride a fait l'objet d'une implémentation longue et en plusieurs phases de façon incrémentale. Nous reviendrons plus tard sur le choix de cet outil, mais soulignons qu'il est à l'origine de cette étude, ayant éveillé l'intérêt d'*Occurrence* pour les problématiques *BI*.

Ainsi, notre étude visera à nous demander **en quelle mesure les nouveaux outils de *Business Intelligence* sont une précieuse aide à la prise de décision dans un secteur hautement concurrentiel comme celui des études, et plus spécifiquement dans le domaine de la communication.**

Pour répondre à cette question, la seconde partie de notre travail constituera une approche théorique des grands concepts de la *BI* et de l'exploitation des données dans le secteur des études et de la communication. Nous réaliserons ainsi un double état de l'art de ces deux disciplines avant de spécifier les besoins et les cas d'application de la *BI* dans notre secteur d'activité. Enfin, la troisième partie sera consacrée à la présentation d'un cas concret : le déploiement progressif des outils *BI* au sein du cabinet d'étude *Occurrence*. Nous y exposerons les différents logiciels utilisés, à commencer par *Microsoft Power BI*, les phases de tests réalisées ainsi que les obstacles et les limites rencontrés.

## II. Cadre théorique : Etude des concepts *BI*

Dans cette seconde partie, nous allons nous concentrer sur une approche théorique de la *Business Intelligence* dans le secteur des études et sondage en communication. Pour cela, nous décrypterons les grands travaux ayant déjà été réalisés par le passé sur ces thématiques et dont nous nous sommes servis pour notre étude. Puis, nous mettrons en lumière les concepts phares de la *BI* et les contextualiserons au sein de notre secteur d'étude.

### **A. État de l'art et historiographie**

#### **a. Le traitement des données de sondage et de communication : Approche théorique**

Tout d'abord, nous souhaitons dissocier la *Business Intelligence* à proprement parler du secteur des sondages et de la communication. Il nous a semblé en effet intéressant de faire le point de façon plus générale sur l'histoire et les grands principes de l'analyse des données dans ces deux disciplines.

En réalité, nous allons principalement nous concentrer sur le domaine des études et sondage, tant ce secteur est intimement lié à la collecte et le traitement des données. Le monde de la communication n'est pas en reste, comme nous le verrons plus tard, mais l'analyse des données de ce secteur n'as pas encore véritablement l'objet de publication scientifique sérieuse. De ce fait, nous disposions de moins de matériel d'étude pour en livrer une analyse approfondie.

#### **1. Ouvrages et auteurs de référence**

Commençons par réaliser un état de l'art des publications de référence dans ce domaine. Notons que tant pour les sondages que la communication, il s'agit de disciplines ayant une forte dimension sociologique. De fait, nous nous sommes intéressés aux grands travaux scientifiques et universitaires de ce secteur. Précisons également que nous avons privilégié les ouvrages

récents ou les rééditions, car il s'agit d'une discipline jeune (datant de la seconde moitié du XXème siècle) et encore en mouvement, impliquant des changements dans sa pratique.

Ainsi, avons principalement basé notre étude sur cinq publications de référence :

- ***L'Opinion publique n'existe pas*, Pierre Bourdieu, 1973<sup>13</sup>** : Et plus généralement, tous ses travaux sur l'analyse des données. Sociologue français iconique ayant révolutionné la pratique de sa discipline, il est cité en référence par tous ses successeurs. Cet article de 1973 cristallise son point de vue sur la pratique des sondages qui tend à se démocratiser en sciences humaines à cette époque. Relativement réfractaire à leur utilisation, il en souligne notamment les biais et les limites.

- ***Sociologie des Opinions*, Claude Dargent, 2011<sup>14</sup>** : Cet ouvrage dédié à un public universitaire centralise toutes les avancées de la recherche en matière de collecte et de traitement de l'opinion publique. L'auteur y présente les méthodes et les bonnes pratiques de la collecte des données de sondage, et fait le point sur leurs utilisations à des fins politiques ou commerciales.

- ***La méthodologie de Pierre Bourdieu en action*, Frédéric Lebaron et Brigitte Le Roux (dir.), 2015<sup>15</sup>** : Travail collectif de plusieurs chercheurs en sociologie et statistiques, il s'agit d'un ouvrage de référence en matière d'analyse des données en sociologie. Sa première partie est consacrée aux théories statistiques entourant ce champ d'étude, en partant du travail réalisé par Bourdieu dans les années 70. Sa seconde partie présente plusieurs cas concrets d'application à l'échelle internationale.

- ***L'analyse quantitative des données*, Olivier Martin, 2020<sup>16</sup>** : Dans sa 5<sup>ème</sup> réédition, ce livre porte également sur la manière de traiter des données quantitatives dans le cadre d'enquêtes et de sondages. Cependant, il ne s'agit pas d'un ouvrage mathématique ou théorique sur le fonctionnement statistique détaillé des analyses. Il se présente plutôt comme un guide fonctionnel des bonnes pratiques à adopter afin d'éviter les biais de traitement.

- **« L'analyse quantitative des médias sociaux, une alternative aux enquêtes déclaratives ? La mesure de la popularité des personnalités politiques sur Twitter »,**

---

<sup>13</sup> BOURDIEU Pierre, « L'opinion publique n'existe pas », dans Les Temps modernes. URL : [https://www.acrimed.org/IMG/article\\_PDF/article\\_a3938.pdf](https://www.acrimed.org/IMG/article_PDF/article_a3938.pdf)

<sup>14</sup> DARGENT Claude, *Sociologie des opinions*, 2011, 240 p.

<sup>15</sup> LE ROUX Brigitte et LEBARON Frédéric (dir.), *La méthodologie de Pierre Bourdieu en action : espace culturel, espace social et analyse des données*, 2015, 262 p.

<sup>16</sup> MARTIN Olivier, *L'analyse quantitative des données*, 2020, 128 p.

***Questions de communication, Julien Boyadjian et Julien Velcin, 2017***<sup>17</sup> : Ce dernier article, bien moins théorique que les précédents ouvrages, présente un cas concret de traitement de données de la communication politique sur Twitter. Les auteurs y décrivent leur méthodologie d'analyse quantitative d'un panel, représentatif à l'aide des nouvelles technologies de traitement des données : Modèles de régression, outils de traitement du langage naturel... Il s'agit de la seule publication que nous ayons trouvée abordant à la fois les thématiques de la communication et des sondages.

Notons que nous nous sommes penchés sur d'autres ouvrages et articles sur le sujet (cf bibliographie), mais ces cinq publications centralisent déjà largement les connaissances théoriques et techniques sur cette discipline.

## 2. Grands principes d'analyse

Dans sa *Sociologie des Opinions*, Claude Dargent nous rappelle que le concept d'« opinion publique » émerge dans la seconde moitié du XVIII<sup>ème</sup> siècle. Mais les premiers sondages de qualité sont arrivés bien plus tard, aux Etats-Unis durant l'entre-deux-guerres. Avant, cela, toujours aux Etats-Unis, la presse invente le concept de sondage en allant demander les intentions de vote des hommes d'environ 45 ans lors d'une élection présidentielle. A ce moment-là, il n'y avait encore aucune notion de représentativité. En 1935, un premier institut de sondage est créé afin de réaliser de la recherche en marketing. Cela constitue une rupture car on commence à se préoccuper de la représentativité des données collectées, et donc de la qualité de l'échantillon. En France, la pratique se développe notamment en 1938, lors de la création de l'Institut Français d'Opinion Publique (IFOP) par le philosophe Jean Stoezel à son retour des Etats-Unis<sup>18</sup>.

Dans leur article, J. Boyadjian et J. Velcin définissent les sondages comme « une technique d'enquête, parmi d'autres, d'analyse des opinions, avec ses avantages et ses limites. » et également « une des seules techniques permettant de mesurer quantitativement des

---

<sup>17</sup> BOYADJIAN Julien et VELCIN Julien, « L'analyse quantitative des médias sociaux, une alternative aux enquêtes déclaratives ? La mesure de la popularité des personnalités politiques sur Twitter », *Questions de communication*, 2017, pp. 111-135

<sup>18</sup> *Op. cit.* DARGENT, *Sociologie des opinions*, pp. 50-65.

opinions. »<sup>19</sup>. Cela nous montre bien que sondage et analyse de données sont intimement liées, il n'y a pas de sondage sans analyse. C'est pourquoi nous allons aborder ici les grands principes de ces méthodes d'analyses.

Largement abordée dans *La méthodologie de Pierre Bourdieu en action*, mais également dans *L'analyse quantitative des données*<sup>20</sup> d'Olivier Matrin, la principale méthode statistique d'analyse de données de sondage est l'Analyse Géométrique des Données (AGD), également appelée Analyse des Correspondances Multiples (ACM). En effet, dans les années 1970, Pierre Bourdieu a cherché très tôt à « étudier les effets globaux d'une structure complexe d'interrelations »<sup>21</sup> que l'on ne peut pas limiter à l'interprétation combinée des effets de variables indépendantes. Pour lui, les modèles statistiques en sciences sociales ont leur limite, c'est pourquoi il se tourne rapidement vers l'ACM.

Il s'agit d'une méthode d'analyse principalement française, impulsée par Bourdieu dans le champ de la sociologie. Originellement développée par J-P Benzécri dans les années 60, elle consiste à placer les individus sur un axe [x,y] en fonction de ses données. Un groupe d'individu constituera alors un nuage de point exploitable (avec des méthodes de *clustering* par exemple) pour identifier les axes principaux de ce nuage de point et comprendre le profil des individus<sup>22</sup>.

Elle constitue la méthode de traitement privilégiée pour le traitement des données de sondages, encore aujourd'hui, car elle permet l'analyse des correspondances entre des individus et des variables catégorielles. Comme les sondages sont largement composés de questions à choix multiples (qui sont bien des variables catégorielles), cette méthode est parfaitement adaptée à l'analyse de ces données. Elle explose en France après les années 80, bien que jugée initialement purement descriptive et non modélisable. Si bien qu'au début des années 90 s'oppose une « école française descriptive » et une « école anglo-saxonne modélisante »<sup>23</sup>.

D'autre part et de façon plus pratique, la création d'un sondage passe par la définition d'une cible d'enquête, puis de la création d'un échantillon provenant de cette population cible

---

<sup>19</sup> *Op.cit.* BOYADJIAN et VELCIN, « L'analyse quantitative des médias sociaux, une alternative aux enquêtes déclaratives ? La mesure de la popularité des personnalités politiques sur Twitter », *Questions de communication*, pp. 111-135

<sup>20</sup> *Op. cit.* MARTIN, *L'analyse quantitative des données*, 128 p.

<sup>21</sup> LEBARON Frédéric (dir.), *La méthodologie de Pierre Bourdieu en action : espace culturel, espace social et analyse des données*, Chap. 3, 2015, p. 45.

<sup>22</sup> LE ROUX Brigitte et LEBARON Frédéric (dir.), *La méthodologie de Pierre Bourdieu en action : espace culturel, espace social et analyse des données*, Chap. 1, 2015, 262 p.

<sup>23</sup> BONNET Philippe, *La méthodologie de Pierre Bourdieu en action : espace culturel, espace social et analyse des données*, Chap. 2, 2015, p. 37

et qui sera interrogé lors du sondage. Il s'agit du principe d'échantillonnage et est un aspect déterminant de la qualité des données obtenues. Olivier Martin, dans son ouvrage, présente plusieurs méthodes d'échantillonnage : échantillon aléatoire, empirique, de convenance<sup>24</sup>... Mais la méthode la plus répandue, et développée par Claude Dargent, est la « méthode des quotas »<sup>25</sup>. Celle-ci consiste se baser sur des quotas caractéristiques de la population étudiée (genre, CSP, âge...) afin de s'assurer que l'échantillon interrogé soit représentatif de la cible visée. Nous aurons l'occasion de présenter cette méthode d'échantillonnage dans un cas concret chez *Occurrence* dans la troisième partie.

Puis, une fois la collecte des données effectuée, la base est nettoyée et enfin, analysée. Une fois de plus présentées par Olivier Martin et Claude Dargent, les deux méthodes d'analyses privilégiées sont les « tris à plat » et les « tris croisés » :

- **Tri à plat** : Il « constitue la première étape de toute analyse de données »<sup>26</sup>. Il permet d'observer la distribution des réponses sur les différentes modalités de chaque question posée. On étudie donc les modalités au sein d'une même variables, qui peuvent être recodées à cette occasion. Le recodage des modalités consiste à regrouper ensemble plusieurs modalités de réponses pour en observer la distribution commune. Cela permet de mettre en avant une modalité de réponse qui se démarque de toute les autres, ou, à l'inverse, la prise en compte d'une opinion minoritaire dans un avis plus global.
- **Tris croisés** : Ils visent à croiser deux variables afin d'en observer la corrélation. Par exemple, l'opinion des répondants sur un sujet en fonction de leur genre ou de leur tranche d'âge. Cela permet de constituer un maillage de l'opinion sur un sujet donné et d'en dégager les différences en fonction des profils de répondants.

### 3. **Biais d'analyse à éviter**

Les différents auteurs s'accordent sur un point : si la science des sondages est intimement liée à l'analyse des données, elle est également soumise à de nombreux biais, dont certains sont inévitables. En effet, dès 1973, Bourdieu publie son texte emblématique intitulé *L'Opinion*

---

<sup>24</sup> *Op. cit.* MARTIN, *L'analyse quantitative des données*, pp. 10-46.

<sup>25</sup> *Op. cit.* DARGENT, *Sociologie des opinions*, p. 101

<sup>26</sup> *Ibid.* p. 148

*publique n'existe pas*. Dedans, il met en garde sur trois éléments inhérents à la pratique des sondages<sup>27</sup> :

- **Tout le monde n'a pas une opinion** : Or la neutralité n'est souvent pas prise en compte dans les sondages.
- **Toutes les opinions ne se valent pas** : Le poids de l'opinion de toutes les catégories et classe sociale n'a pas la même valeur, dépendamment du sujet.
- **Les sondages sont toujours motivés par un intérêt politique ou économique** : Donc, par définition, ils seront biaisés pour servir leurs intérêts.

Par conséquent, il se montre très critique vis-à-vis d'une analyse statistique des données d'opinion. Selon lui, un sondage implique nécessairement des biais d'analyse, qu'il faut au mieux parvenir à éviter. C'est également pour cela qu'il privilégie une approche descriptive qu'une approche modélisante : il est plus intéressant de parvenir à établir des profils types que d'essayer de dégager des tendances.

En effet, appelés « biais de collecte » il s'agit de la principale limite des données de sondage. Selon la tournure de la question posée aux répondants, celle-ci sera interprétée différemment et les données collectées présenteront des variations. Même sans cela, J. Boyadjian et J. Velcin expliquent qu'il peut y avoir un décalage entre le répondant et le sondeur quant à la façon dont il faut répondre. Pour les auteurs, il n'y a donc « pas de formulation idéale »<sup>28</sup>, ces biais sont inévitables et il est nécessaire de les considérer lors de l'interprétation des données collectées.

De même, un autre biais soulevé par Claude Dargent est le biais d'échantillonnage<sup>29</sup>. Nous l'avons vu, il s'agit d'une étape cruciale de préparation d'une enquête, mais les résultats peuvent être faussés si l'échantillon n'est pas représentatif. Malgré les méthodes évoquées plus haut, il est possible que celui-ci ne soit pas représentatif par manque de significativité. Cela signifie que l'échantillon est trop petit pour la population étudiée. En effet, plus la population cible est conséquente, plus le nombre de répondant doit être important. C'est pourquoi malgré le respect des quotas, les résultats peuvent être biaisés si les principes de base de l'échantillonnage ne sont pas respectés scrupuleusement.

---

<sup>27</sup> *Op. cit.* BOURDIEU, « L'opinion publique n'existe pas », dans *Les Temps modernes*.

<sup>28</sup> *Op.cit.* BOYADJIAN et VELCIN, « L'analyse quantitative des médias sociaux, une alternative aux enquêtes déclaratives ? La mesure de la popularité des personnalités politiques sur Twitter », *Questions de communication*, pp. 111-135

<sup>29</sup> *Op. cit.* DARGENT, *Sociologie des opinions*, 240 p.

#### 4. Analyse des données de communication

Observons également les spécificités de la communication. Il s'agit peut-être de l'un des secteurs où l'on dispose du plus grand nombre d'outils et de données pour mesurer ses performances. En effet, il s'agit d'un domaine très large englobant de nombreuses sous-catégories : les réseaux sociaux, l'événementiel, la presse, les sites web, les partenaires commerciaux, etc. Tous ces éléments sont quantifiables et disposent de *KPI* très bien définis.

Les données d'évaluation de la communication d'une entreprise peuvent notamment être collectées et analysées par le prisme des sondages, comme le fait *Occurrence*. Aujourd'hui, Claude Dargent nous rappelle que tous les domaines d'activités mettent en place des sondages d'opinion. Les visées économiques de ces sondages ont largement dépassé leur dimension sociologique ou politique<sup>30</sup>. Ceux-ci permettent en effet d'interroger directement le consommateur et de collecter son point de vue sur différents sujets générateurs de profits pour les entreprises. La communication s'inscrit donc dans cette démarche et cherche également à collecter l'opinion des citoyens : sur une publicité, un événement, une image de marque ou encore sur le climat interne en entreprise.

Mais d'autres moyens de collectes et d'analyse des données sont possibles. Par exemple, l'attractivité d'un site internet peut être quantifiée à l'aide de *Google Analytics*. De manière générale, les outils digitaux ont en effet ouvert de nouvelles possibilités dans l'analyse des données de communication. Les réseaux sociaux en sont le plus bel exemple, offrant aux entreprises des plateformes de collecte d'opinions communiquées publiquement. Cette communication sur certains sujets peut ainsi être évaluée, comme le démontrent J. Boyadjan et J. Velcin dans leur publication sur l'analyse de la communication politique sur Twitter<sup>31</sup>.

---

<sup>30</sup> *Ibid.*

<sup>31</sup> *Op.cit.* BOYADJIAN et VELCIN, « L'analyse quantitative des médias sociaux, une alternative aux enquêtes déclaratives ? La mesure de la popularité des personnalités politiques sur Twitter », *Questions de communication*, pp. 111-135

## **b. Emergence et généralisation des outils de *Business Intelligence***

Dans cette seconde section, intéressons-nous au deuxième volet de notre étude théorique : la *Business Intelligence*. Pour rappel, il s'agit d'une discipline informatique également appelée « informatique décisionnelle », visant à collecter, analyser et restituer de la donnée afin d'assister les décideurs en entreprise dans leurs prises de décisions stratégiques. Ainsi, nous allons réaliser, comme précédemment, une historiographie de ce sujet, en présentant les ouvrages de références de ce domaine et dont nous nous sommes servis, ainsi que certaines publications traitant d'une thématique précise. Dans un second temps, nous reviendrons brièvement sur l'histoire et le fonctionnement théorique des outils *BI* afin de mieux cerner les enjeux de cette discipline.

### **1. Etat de l'art**

Tout d'abord, présentons les ouvrages et auteurs de référence en matière d'informatique décisionnelle, nous ayant permis de mieux appréhender cette vaste discipline. Notons que cette pratique a commencée à se développer aux portes du XXIème siècle et a rapidement explosé en apportant un nouvel éclairage sur les données et leurs usages. De ce fait, la recherche documentaire sur ce sujet a été particulièrement prolifique et nous avons dû procéder à une sélection des publications retenues pour notre étude.

En premier lieu, à quelques exceptions sur des éléments spécifiques, nous avons exclus les ouvrages trop anciens. Nous le savons, l'informatique a connu des évolutions drastiques au cours du XXIème siècle. A l'an 2000, le terme de *Big data* était encore inconnu de tous et les premiers outils de *Business Intelligence* étaient loin de détenir la puissance de calcul et la flexibilité dont nous disposons aujourd'hui. De même, les technologies du *Cloud* et les logiciels SAAS, pourtant si importantes pour la *BI* moderne, n'existaient pas encore. De ce fait, nous avons entrepris de ne conserver que les publications ayant moins de 20 ans – sauf exception – afin de conserver une vision actuelle de la discipline.

Par ailleurs, remarquons que si l'informatique décisionnelle est un sujet ayant fait l'objet de nombreuses publications ces dernières années, très peu d'entre elles présentent une approche théorique de celui-ci. En effet, la plupart des ouvrages abordent la *BI* de façon fonctionnelle,

dans des cas d'application concrets de son déploiement et à destination des professionnels en entreprise. Nous avons ainsi essayé de nous pencher à la fois sur des publications universitaires théoriques et sur son approche pratique afin de mieux identifier les enjeux et les obstacles à son implantation.

De fait, voici les quatre publications de références qui constituent les piliers de notre compréhension de la *BI* :

- ***Le Projet Décisionnel : Enjeux, Modèles, Architectures du Data Warehouse, Jean-Marie Gouarné, 1997***<sup>32</sup> : Plus ancien ouvrage de notre bibliographie, il s'agit d'une référence qui met en lumière le rôle du *Data Warehouse* dans le développement des systèmes d'informations contemporains. Il présente ses enjeux et son impact lors des projets d'informatique décisionnelle, dont il constitue les fondements. Très technique, nous avons principalement utilisé ce document pour comprendre le fonctionnement d'un *Datawarehouse* et sa complémentarité avec les outils *BI*.

- ***Les nouveaux tableaux de bord des managers, Alain Fernandez, 2013***<sup>33</sup> : Ayant rédigé plusieurs ouvrages portant sur le déploiement des outils de *Business Intelligence* en entreprise, Alain Fernandez signe ici son guide de référence en matière de gestion de projet dans ce domaine, réédité pour la sixième fois. Il y explique comment réussir son projet d'intégration *BI*, comment exploiter les données et réaliser des tableaux de bords impactant. Dans une moindre mesure, il s'intéresse également aux indicateurs de mesure de la performance.

- ***Construire des tableaux de bord vraiment utiles, Marie-Hélène Millie-Timbal, 2020***<sup>34</sup> : Il s'agit également d'un guide, plus synthétique et plus actuel que le précédent, dédié à la conception de tableaux de bord de pilotage en entreprise. Celui-ci donne notamment des conseils pratiques sur la mise en place, l'identification et la documentation des indicateurs de la performance.

---

<sup>32</sup> GOUARNE Jean-Marie, *Le Projet décisionnel : Enjeux, Modèles, Architectures du Data Warehouse*, 1997. 164 p.

<sup>33</sup> FERNANDEZ Alain, *Les Nouveaux Tableaux de bord des managers, Le Projet Business Intelligence clés en main*, 2013, 662 p.

<sup>34</sup> MILLIE-TIMBAL Marie-Hélène, *Construire des tableaux de bord vraiment utiles : associer les indicateurs à la stratégie de l'entreprise et à son management*, 2020, 190 p.

- « **L'évolution des rôles du contrôleur de gestion à l'ère de la *Business Intelligence*** », ACCRA, par **Élodie Allain, Patrice Landagaray et Jérôme Reutter, 2021**<sup>35</sup> : Cet article, au sujet spécifique peut sembler de prime abord moins pertinent que les précédents pour estimer qu'il constitue une « publication de référence ». Toutefois, il se distingue par sa grande clarté sur un sujet si dense et parfois relativement opaque. Ainsi, ce travail de quasi-vulgarisation de la *BI* à travers un cas d'implantation concret auprès des contrôleurs de gestion nous a apporté un éclairage nouveau sur la discipline et de nombreux éléments de réponses. Si tant est qu'il ne puisse être considéré comme une publication majeure sur ce sujet, il représente à minima un écrit central pour notre étude.

Notons que si ces publications constituent la base de notre étude, nous avons également consulté d'autres ressources pour étayer notre compréhension de points spécifiques. Par exemple, l'article de Brahim Saber, publié de 2011 dans la *Revue internationale d'intelligence économique* et abordant le rôle des logiciels *open source* pour la *Business Intelligence*<sup>36</sup>. De même, Béa Arruabarrena publie, à l'occasion du colloque « Document numérique et société » de Rabat en 2015, un article portant sur les enjeux de la datavisualisation<sup>37</sup> qui nous a été très utile pour expliciter ce concept.

## **2. Histoire de la *Business Intelligence***

A travers les éléments issues de nos recherches et des ouvrages cités précédemment, procédons maintenant à une présentation succincte de la discipline qu'est l'informatique décisionnelle. Aussi commençons par un rappel historique afin de remonter aux origines de cette science nouvelle.

Les tableaux de bord, tout d'abord, constituent comme nous le verrons plus tard l'un des piliers de la *Business Intelligence*, à la fois outil de pilotage et support d'aide à la prise de

---

<sup>35</sup> REUTTER Jérôme, ALLAIN Élodie, LANDAGARAY Patrice, « L'évolution des rôles du contrôleur de gestion à l'ère de la *Business Intelligence* », dans ACCRA, 2021, pp. 85-107. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-accra-2021-2-page-85.htm>

<sup>36</sup> SABER Brahim, « Open source comme système d'informatique décisionnelle », dans *Revue internationale d'intelligence économique*, 2011, pp. 93-101. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-internationale-d-intelligence-economique-1-2011-1-page-93.htm>

<sup>37</sup> ARRUABARRENA Béa, « Datavisualisation : principes, enjeux et perspectives pour des utilisateurs non experts », dans BROUDOUX Évelyne, *Big Data - Open Data : Quelles valeurs ? Quels enjeux ? Actes du colloque « Document numérique et société »*, Rabat, 2015, pp. 151-163. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/--9782807300316-page-151.htm>

décision. Ceux-ci constituent sans nul doute les prémices les plus anciens de cette discipline. En effet, dans son ouvrage sobrement intitulé *Le Tableau de Bord*, publié en 2021, Pierre Maurin explique que leur première trace remonte aux Etats-Unis à la fin du XIX<sup>ème</sup> siècle, au sein des grandes entreprises américaines de l'époque<sup>38</sup>. Ceux-ci étaient réalisés à la main dans le but de synthétiser à l'aide de graphiques les finances du groupe.

Cependant, il faudra attendre près de 100 ans, jusque dans les 1970, pour que l'on dissocie les tableaux de bords de problématiques purement liées à la production, tel que la productivité des ouvriers ou les coûts de matières premières par exemple<sup>39</sup>. C'est ainsi le client qui devient le centre du *Dashboard*, avec l'émergence des premiers indicateurs qualitatifs, tels que la satisfaction de celui-ci. Ces tableaux, qui étaient jusque-là des outils de contrôle, deviennent progressivement des outils de pilotage et de performance, dont les champs d'évaluation s'étendent à tous les secteurs de l'entreprise : la comptabilité, l'administration, la communication, etc. Il s'agit là de la forme contemporaine des *Dashboards* tels que conçus aujourd'hui à travers les outils de d'informatique décisionnelle.

Concernant la *BI* plus directement, il semblerait que la première apparition du terme soit relativement ancienne, initiée par l'analyse Hans Peter Luhn dans un article rédigé en 1958 et intitulé « *A Business Intelligence System* »<sup>40</sup>. Il décrit ce système comme un outil automatique de diffusion de l'information aux décideurs d'une entreprise ou d'une organisation. En somme, nous nous rapprochons déjà de son acceptation contemporaine, bien qu'il ne prenne pas encore en compte la dimension informatique dans son système.

Puis, dans les années 70, le développement des technologies informatiques commencent à faire émerger les premiers systèmes de requêtes sur des bases de données. Encore dix ans plus tard, émergent les premiers « infocentres », c'est-à-dire, des espaces mutualisés stockant une copie des données destinée à leur analyse<sup>41</sup>. Il s'agit en effet des prémices de l'un des grands principes de l'informatique décisionnelle, visant à analyser une copie des données dans un espace sécurisé et non à la source directement.

Mais c'est réellement dans les années 90 que la *BI* prend réellement sa définition moderne, puisque l'analyse Howard Dresner reprend la terminologie de 1958 pour définir cette

---

<sup>38</sup> MAURIN Pierre, *Le tableau de bord*, 2021, p.7

<sup>39</sup> *Ibid*, p.7

<sup>40</sup> AURINE Guillaume, « Une brève histoire de la *Business Intelligence* », 7 juin 2016. URL : <https://www.salesforce.com/fr/blog/2016/06/une-breve-histoire-de-la-business-intelligence.html>

<sup>41</sup> *Ibid*.

discipline comme l'ensemble des méthodes d'analyse des données afin d'aider à la prise de décision<sup>42</sup>. En parallèle apparaissent les « *Data Warehouse* », ou entrepôts de données en Français, que Jean-Marie Gouarné qualifie dans son ouvrage de « clé de voute de ce que nous appelons l'information décisionnelle »<sup>43</sup>.

Enfin, la *BI* prend encore un nouvel essor dans le courant des années 2000 avec la naissance de la *Big data*, se développant à mesure que la puissance de calcul des ordinateurs augmente. Aujourd'hui, la *BI* est marquée le déploiement d'algorithmes prédictifs de pointe et un pilotage en temps réel alimenté par un flot discontinu de données.

### **3. Rôles et fonctionnement général des outils *BI***

Pour comprendre le fonctionnement général des outils de *Business Intelligence*, il nous faut en amont comprendre le fonctionnement des *datawarehouses* et son implantation au sein du système d'information (SI) de l'organisme. En effet, celui-ci a longtemps occupé une place prépondérante au sein des architectures SI et prédomine toujours malgré l'émergence de nouvelles solutions de substitutions, tels que les *datalakes*. Dans son article sur les systèmes d'aide à la décision, publiée en 2015, Christine Sybord les décrits comme des « systèmes [intégrant] des applications décisionnelles rassemblées dans une immense base de données, appelée « entrepôt » (Warehouse). Les données y sont thématiques [...]. »<sup>44</sup>.

Dans toute organisation, les données produites et générées proviennent de nombreuses sources différentes. Il peut s'agit de données issues du web, de bases relationnelles *SQL*, ou même de fichiers « plats » tels que des tableaux *Excel* par exemple. Les données dites « brutes » car elles n'ont pas fait l'objet de traitement prennent donc des formes différentes, et peuvent être structurées ou non structurées (stockées dans leur format d'origine). Par conséquent, il est impossible dans de telles dispositions de les faire communiquer entre elles et de procéder à leur analyse.

---

<sup>42</sup> *Ibid.*

<sup>43</sup> *Op. cit.* GOUARNE, *Le Projet décisionnel : Enjeux, Modèles, Architectures du Data Warehouse*, 1997. p. 4

<sup>44</sup> SYBORD Christine, « Intelligence économique et système d'aide à la décision : de l'opérationnel « big data » au stratégique « little knowledge » », dans *Revue internationale d'intelligence économique*, vol.7, 2015, p. 85. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-internationale-d-intelligence-economique-1-2015-1-page-83.htm>

Le *Datawarehouse* permet ainsi de palier à ce problème, comme l'explique avec précision Jean-Marie Gouarné dans *Le Projet Décisionnel*, consacré à la théorisation de l'architecture de ces entrepôts de données<sup>45</sup>. Au sein du système d'information de l'organisme, il constitue une base de données relationnel, conçue à des fins d'analyse de données. Il va donc regrouper une copie de l'ensemble des données fonctionnelles de l'entreprise et établir des relations entre elles de façon thématique, comme l'a mentionnée Christine Sybord dans son article. Pour cela, le SI de l'organisme doit comprendre des connexions entre les données sources et le *Datawarehouse*. En somme, voici comment s'articule le cheminement des données, ou *data flow* :

- **Écriture des données** dans la source : Les données sont stockées dans leur base source et y demeurent jusqu'à leur suppression.
- **Extraction des données** depuis la source et copie dans le *Datawarehouse* à l'aide d'un outil d'*ETL*. Nous en détaillerons le fonctionnement plus tard, mais celui-ci permet notamment de transformer les données entre leur extraction et leur chargement dans le *Datawarehouse*.
- **Intégration et transformation** : Il s'agit de l'étape cruciale du transfert des données. Les bases issues des différentes sources doivent ainsi être remodelées afin de toutes présenter la même structure et de pouvoir par conséquent établir des relations entre elles. Ce processus est automatisé, mais il requiert un long paramétrage et une très bonne connaissance des bases importées.
- **Procédure de traitement et d'analyse** : Certains *datawarehouses* proposent des solutions de traitement directement depuis celui-ci, mais la plupart d'entre eux ne disposent pas par exemple d'un outil de restitution des analyses. Par conséquent, les données peuvent ensuite être transférés vers d'autres outils d'informatique décisionnelle, qui ne modifieront pas directement les données stockées au sein du *Datawarehouse*.

Nous pouvons donc constater la place centrale qu'occupe cet entrepôt dans le système d'information d'une entreprise qui désire réaliser de la *Business Intelligence*. Cependant, tous

---

<sup>45</sup> *Op. cit.* GOUARNE, *Le Projet décisionnel : Enjeux, Modèles, Architectures du Data Warehouse*, 1997. 164 p.

les organismes ne disposent pas d'un SI aussi bien orchestré. La réalité du terrain englobe également de nombreux organismes n'ayant pas réellement opéré de transition numérique, et se limite à des bases *Excel* stockées sur des serveurs locaux. Mais des solutions existent aujourd'hui afin de palier à ce problème tout en réalisant de la *BI*.

Aujourd'hui, les outils *BI* sur le marché ont grandement évolué et certains se présentent donc comme des solutions complètes. De ce fait, elles peuvent partiellement remplir le rôle d'un *Datawarehouse*, sans avoir à reconstruire toute l'architecture du système d'information. En effet, ces outils, bien qu'ils soient conçus pour se connecter à un entrepôt de données, disposent de nombreux connecteurs vers d'autres sources de données. A l'aide d'un *ETL*, intégré ou non au logiciel, les données peuvent alors être transformées et copiées dans l'outil *BI*, à la manière d'un *Datawarehouse*, mais il ne dispose pas des mêmes capacités de stockage et de traitement.

En bout de chaîne, après le traitement des données et leur analyse par le biais de requête et de calculs, les résultats peuvent ensuite être restitués au sein d'un tableau de bord à des fins de pilotage. Bien qu'il ne s'agisse que de la dernière étape de l'ensemble complexe du processus de *Business Intelligence*, elle est également l'une des plus importante. En effet, tout l'enjeu de ce process est de rendre la donnée interprétable par les décideurs en entreprise afin d'accompagner leurs prises de décision. C'est pourquoi le tableau de bord constitue la finalité de la chaîne de traitement, et que les décideurs ne doivent pas être écartés du processus de création et d'analyse des données. M-H. Millie-Timbal le rappelle à bon escient dans son guide pratique : « [Les] objectifs doivent donc être clairement formulés : quelles sont les priorités et [...] quels résultats voulons-nous atteindre ? »<sup>46</sup>. Sans l'implication des décideurs pour répondre à ces questionnements et superviser la création, le projet *data* est voué à l'échec.

## **B. Concepts clés et terminologie du sujet**

Suite à l'étude théorique des deux disciplines qui encadrent notre projet, à savoir, le secteur des sondages et celui de l'informatique décisionnelles, nous pouvons progressivement préparer le terrain en prévision de notre projet concret d'intégration. Nous venons d'observer, dans les grandes lignes, les principaux axes de fonctionnement de ces secteurs. Dans cette

---

<sup>46</sup> *Op. Cit.* MILLIE-TIMBAL, *Construire des tableaux de bord vraiment utiles : associer les indicateurs à la stratégie de l'entreprise et à son management*, 2020, p. 49

seconde partie, nous allons donc nous rapprocher d'une vision pratique et de terrain des grandes notions observées. Nous conserverons cependant un angle d'approche généralisé et académique afin de ne pas encore évoquer notre cas spécifique.

Ainsi, nous allons aborder dans un premier temps les concepts phares de la *BI*, à l'image d'un dictionnaire pratique décrivant les grandes notions de cette discipline, susceptibles d'être utilisées dans le cadre d'un projet de *BI*. Dans un second temps, nous croiserons les disciplines afin de rattacher ces concepts à des cadres d'utilisation concrets au sein de cabinets d'études et de sondages.

### **a. Décryptage des concepts phares de l'informatique décisionnelle et du secteur**

La *BI* est un champ d'étude vaste, mais le champ sémantique l'accompagnant l'est tout autant. Les terminologies utilisées sont nombreuses et désignent parfois une notion complexe à travers un terme simple, où l'inverse. Les anglicismes sont également particulièrement présents dans cette discipline, d'où leur utilisation dans notre étude et le besoin de traduction de termes que nous effectuons par endroit.

Ici, nous allons donc nous pencher sur les principales notions qui régissent ce champ d'étude, en explicitant à quoi celles-ci se réfèrent précisément. Précisions que nous détaillerons moins en profondeur les éléments que nous avons déjà évoqués plusieurs fois précédemment, afin d'en alléger la lecture.

#### **1. Automatisation**

Le concept d'automatisation constitue un élément central de la pratique de la *Business Intelligence*. Elle n'est toutefois pas obligatoire et l'ensemble des processus peuvent continuer à être exécutés manuellement. Cependant, toutes les étapes de traitement, de l'import des données à leur restitution sous forme de tableaux de bord, ont été conçues et imaginées pour minimiser l'intervention humaine.

Nous l'avons vu, le processus d'intégration et la chaîne de traitement de la donnée une fois dans l'outil sont des éléments longs et complexes. Or les outils de *BI* ont justement été

conçus afin d'améliorer les performances opérationnelles. Il serait donc une perte de temps considérable, allant à l'encontre de la logique même d'une telle solution, de procéder à ces traitements de façon manuelle. Dans l'idéal, ces outils sont d'ailleurs conçus de sorte qu'une fois les données intégrées et le tableau de bord réalisé et diffusé, l'intervention humaine ne soit plus nécessaire. Même l'actualisation des données est automatisable de sorte qu'elle soit planifiée automatiquement à des dates précises.

Dans leur article intitulé « Automatisation des tableaux de bord et cohérence du contrôle de gestion : à propos de deux cas », Christine Marsal et Denis Travaillé soulignent que cette automatisation est rendue possible par la cohérence temporelle des données<sup>47</sup>. En somme, une fois une base de données source créée, celle-ci ne change plus de forme sur la durée. Les nouvelles données viennent ainsi s'incrémenter dans la base sans en modifier l'organisation. Cette stabilité des bases perdure ainsi lors de son import dans le logiciel *BI* à chaque nouvelle actualisation des données. Les instructions de traitement préconfigurées n'ont alors plus qu'à être répétées avec les nouvelles données pour mettre à jour un tableau de bord existant.

Il faut toutefois signaler deux contraintes à cette automatisation globalisée des processus. D'une part, elle implique une phase de paramétrage en amont, afin d'indiquer au logiciel les requêtes de traitement que l'on souhaite réaliser sur les données source afin de les mettre en forme et de procéder à leur analyse. Cette phase de paramétrage, dépendamment du nombre de bases importées, de la variété des sources et de la complexité du tableau de bord, peut être particulièrement complexe et chronophage. Elle implique également une bonne connaissance humaine des données traitées afin d'éviter que les résultats ne soient erronés une fois les paramètres d'automatisation configurés. Le second point de vigilance découle de ce fait : Si une erreur est réalisée lors du paramétrage, celle-ci sera reproduite constamment, sans vérification humaine de la validité des données.

## **2. Indicateurs de performance (KPI)**

Les indicateurs de performances, également appelé *KPI*, sont la concrétisation des objectifs de l'entreprise en matière de performance ou de pilotage, matérialisée par des éléments

---

<sup>47</sup> TRAVAILLÉ Denis, MARSAL Christine, « Automatisation des tableaux de bord et cohérence du contrôle de gestion : à propos de deux cas », dans *Comptabilité Contrôle Audit*, 2007, pp. 75-96. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-comptabilite-controle-audit-2007-2-page-75.htm>

quantifiables à l'aide des données. Plus concrètement, il s'agit des traitements et calculs réalisés sur les données une fois intégrées à l'outil afin d'apporter des éléments mesurables et interprétables aux décideurs d'entreprises. Ce sont ces *KPI* qui sont rapportés sur les tableaux de bord et qui constituent le pouvoir d'aide à la prise de décision des outils *BI*.

Attention toutefois, comme le rappellent A. Fernandez<sup>48</sup> et M-H Millie-Timbal<sup>49</sup> dans leurs ouvrages respectifs sur les tableaux de bord, le choix des indicateurs est une étape cruciale dans projet de pilotage. Les objectifs fixés dépendent d'une part du secteur d'activité, car tous les domaines disposent de leurs propres indicateurs de références, et d'autre part de la méthodologie de pilotage de l'utilisateur final, purement personnelle.

### 3. Tableaux de bord

Nous l'avons déjà largement évoqué, le tableau de bord constitue l'outil clé de pilotage à l'aide de la *Business Intelligence*. Il se place comme le support de l'interprétation des données, c'est-à-dire la passerelle entre données brutes et information. Il en existe différents types (Pilotage, suivi des performances internes, suivi de l'activité, suivi budgétaire, information client...) et son architecture évoluera selon ces types et les besoins spécifique des utilisateurs. Tout projet *BI* passe par l'élaboration d'un tableau de bord. Pour cela, il est nécessaire de consulter les utilisateurs finaux et de comprendre leurs besoins métiers.

Dans le domaine de l'informatique décisionnelle moderne, les tableaux de bord se démarquent de ceux créés à l'aide d'autres outils par leur interactivité. Ils sont en effet conçus de sorte à être consultés en ligne de façon dynamique. Ainsi, ces tableaux disposent de multiples fonctionnalités de filtres afin de pouvoir adapter les résultats en temps réels à de nombreuses contraintes métiers. Les filtres temporels sont notamment les plus répandus, car ils permettent d'explorer les données en 3 dimensions, ajoutant au pilotage des données d'évolution des indicateurs.

---

<sup>48</sup> Op. Cit. FERNANDEZ, *Les Nouveaux Tableaux de bord des managers, Le Projet Business Intelligence clés en main*, 2013, 662 p.

<sup>49</sup> Op. Cit. MILLIE-TIMBAL, *Construire des tableaux de bord vraiment utiles : associer les indicateurs à la stratégie de l'entreprise et à son management*, 2020, 190 p.

#### 4. Datavisualisation

La datavisualisation est une sous-discipline englobée dans le champ de la *Business Intelligence*. Dans un article sur ce sujet, publié en 2015 dans *Big Data – Open Data : Quelles valeurs ? Quels enjeux ?* par Béra Arruabarrena, l’auteur décrit cette science comme « un processus de traitement visuel des données à part entière qui comprend la collecte, la structuration visuelle jusqu’au rendu final de la visualisation. »<sup>50</sup>. En effet, cette discipline se différencie de la représentation artistique de la donnée, figée, par l’ensemble des traitements qui accompagnent le résultat final.

A l’origine réservée aux experts de la manipulation des données, la prolifération des outils *no-code* ces dernières années a ouvert cette pratique à tous, engendrant sa démocratisation. Son utilisation est d’ailleurs primordiale en matière de restitution de résultats de traitement des données. En effet, la datavisualisation simplifie grandement l’accès et la compréhension des données pour tous, jouant un rôle de médiation de l’interaction entre les hommes et celles-ci<sup>51</sup>.

Cependant, si les modes de représentation graphique des données se multiplient, il est nécessaire d’en comprendre les usages. En effet, toutes les visualisations ne sont pas adaptées à la restitution de toutes les données. Un camembert conviendra par exemple à une distribution en pourcentages, mais bien moins à des données temporelles. De même, un nuage de point conviendra parfaitement à l’étude combinatoire de deux variables sur de nombreuses données, mais n’est pas adapté à la comparaison entre deux éléments uniques. Dans le champ de la *BI* et de l’interprétation des données, il est donc important de distinguer ces éléments et de ne pas se perdre en complexité d’analyse afin de produire des tableaux de bord impactant.

#### 5. Concept de « no-code »

Les outils « *no-code* » désignent les solutions logicielles répondant à des problématiques informatiques de toutes sortes sans avoir besoin de maîtriser un langage de

---

<sup>50</sup> *Op. Cit.* ARRUBARRENA, « Datavisualisation : principes, enjeux et perspectives pour des utilisateurs non experts », dans *Big Data - Open Data : Quelles valeurs ? Quels enjeux ? Actes du colloque « Document numérique et société »*, 2015, p. 153.

<sup>51</sup> *Ibid.* p. 153

programmation. Ceux-ci ont explosés depuis ces dernières années : la programmation ne sert plus à exécuter des tâches, mais à développer des outils permettant à tous de le faire. Ces solutions font écho à un besoins émergents d'une prise en main plus rapide des outils, sans avoir recourt à une formation avancée en programmation.

Dans le domaine de la *BI*, les outils *no-code* sont la norme : il peut s'agir d'outils spécialisés dans le nettoyage de la donnée, la création de tableaux de bord, etc. La pratique de la *BI* à l'aide de langage de programmation n'est pas à la portée de tout le monde et semble très opaque et technique. Cependant, les outils *no-code* ont aussi leur lot de difficultés et leur maîtrise demande une formation rigoureuse.

En effet, nous pouvons constater une corrélation entre les possibilités techniques d'un outil et la difficulté de sa prise en main. En clair, plus un outil dispose de fonctionnalités, plus sa prise en main sera complexe. Chaque outil *BI* existant a son propre équilibre à ce niveau : certains vont être très technique (*Power BI* par exemple) et difficile à maîtriser. D'autres disposeront de possibilités limitées mais d'une prise en main intuitive (*Pocket Result* par exemple).

De plus en plus, l'analyse de données se détache de la programmation : les *Data analysts* ne sont pas des développeurs. Du moins, ils peuvent ne pas l'être, ce qui aurait été inenvisageable il y a quelques années. Il existe même des outils « *no-code* » destinés à la réalisation de modèles statistiques de régression, de prédiction, ou encore de *data mining*. Cependant, ces outils « *no-code* » posent également un problème de transparence. En programmation, les développeurs sont maîtres de leur code, ils savent ce qu'ils font et en comprennent le fonctionnement. En « *no-code* », ces logiciels « clique-bouton » nous indiquent à quoi servent chaque fonctionnalité, mais ne présentent pas la logique de fonctionnement. L'usage de certaines fonctionnalités peut alors paraître obscure et lorsqu'un problème ou un phénomène inattendu survient, il est parfois difficile d'en identifier la cause. Souvent sous licence propriétaire, ces outils ne s'inscrivent pas non plus dans une logique d'*open data* permettant à tout un chacun d'en parcourir le code source.

Cette explosion des logiciels *no-code* est à mettre en relation avec une réalité de terrain pour beaucoup d'entreprises et donc la tendance est à la hausse : lors du recrutement, on a tendance à privilégier l'expertise métier à la maîtrise technique des outils. Notamment au sein des PME et des start-ups, on considère qu'il est plus simple pour un employé de se former à des logiciels qu'il ne maîtrise pas plutôt que d'acquérir une compréhension métier relative à un

domaine d'activité spécifique. Suivant cette logique, il est alors plus simple d'acquérir les rudiments de la *BI* à l'aide d'outils « *no-code* » que d'acquérir la maîtrise d'un langage de programmation, dont la tâche est plus fastidieuse.

De plus, toutes les entreprises souhaitent valoriser leurs données à l'aide de l'informatique décisionnelle. Elles ont compris que les données avaient de la valeur, tout le monde veut pouvoir suivre ses indicateurs de performances, mieux comprendre ses clients, etc. Une fois de plus, pour des personnes non-initiées, intégrer ce domaine paraît bien plus simple à travers des outils sans programmation, d'apparence semblable à d'autres logiciels de leur quotidien, tel qu'*Excel*, *Power-Point*, etc. Mais cette simplicité est à double tranchant, l'apparente accessibilité de ces outils est attrayante, mais la compréhension technique des mécanismes de la *data* analyse n'est pas à la portée de tous. Il est facile de réaliser des erreurs de calcul ou de manipulation, et par conséquent, baser les décisions managériales sur des données erronées. Le *cleaning* des données en amont par exemple, est une étape primordiale pouvant être négligée. De la même façon, les graphiques choisis peuvent ne pas être pertinents et biaiser l'analyse. Globalement la logique qualité n'est pas toujours assurée et présente un risque en cas de mauvaise maîtrise de l'outil.

Toutefois, malgré les nombreux axes de vigilances que nous venons de souligner : les logiciels « *no-code* » présentent une avancée considérable permettant de rendre la *BI* et l'analyse de données de plus en plus accessible, alors qu'elle était jusqu'alors réservée à des experts de la programmation. Il faut cependant garder en tête que ces outils requièrent de la pratique pour en assurer la maîtrise, et qu'elles ne constituent pas toujours une solution de facilité.

## **b. Mise en perspective avec les projets menés par les instituts d'étude**

Après avoir étudié les bases théoriques de la *Business Intelligence* et de l'analyse des données de sondage, nous allons croiser dans cette partie ces deux domaines d'activités. L'objectif est de parvenir à déterminer les voies d'application de la *BI* au sein des instituts d'études, en identifier les apports et les contraintes.

Tout d'abord, notons que la *BI* peut avoir plusieurs voies d'utilisation, mais l'une de ses finalités est similaire aux sondages : la compréhension d'un publique ou d'une clientèle à l'aide des données. Dans le cas des sondages, on interroge le consommateur et on récolte son opinion. Dans l'analyse de données, on collecte et exploite de façon plus large les informations dont nous disposons en provenance de sources de données multiples. Comme l'explique Maria Mercanti-Guérin dans la *Revue internationale d'intelligence économique* en 2013, « L'objectif [est] de qualifier les consommateurs et de leur adjoindre des scores de performance »<sup>52</sup>.

Ainsi, la *BI* présente l'avantage, pour une entreprise, de porter un regard sans biais (du moins à l'origine) sur les consommateurs. La donnée expose des faits et non une opinion à paramètres variables : par exemple, les articles les plus achetés selon la saisonnalité, le profil des consommateurs, etc. Ainsi, on évite les biais d'interprétations propres aux sondages, tels que présentés plus tôt (cf *II.A.a.3. – Biais d'analyses à éviter*). Cependant, une telle analyse ne permet pas d'obtenir des données aussi précises et concrètes que les sondages en termes de ressenti du consommateur. Par exemple, l'extraction et le traitement de la donnée à l'aide de la *BI* ne saurait déterminer ce qui a motivé une personne à acheter un produit.

De ce fait, au sein des instituts d'études et de sondages, l'implémentation d'outils et de techniques d'informatique décisionnelle possède un rôle très complémentaire auprès des entreprises clientes. En somme, ces technologies ne viennent pas remplacer les méthodes d'étude d'opinion traditionnelle, mais elles proposent une nouvelle dimension de compréhension de la clientèle.

Cette utilisation est en effet en parfaite adéquation avec le fonctionnement des cabinets d'études tels qu'ils fonctionnent aujourd'hui. Comme, nous l'avons dit, les sondages sont désormais majoritairement réalisés à des fins économiques ou, politique, dans une moindre mesure. De fait, la majorité des études réalisés par les cabinets sont des missions de prestations auprès d'entreprises clientes, après avoir répondu et remporté un appel d'offre (AO). Or ces AO requièrent de plus en plus l'intégration d'outils de *BI* dans les études souhaitées, en complément ou non d'une étude par sondage. Ainsi, des experts de l'informatique décisionnelle au sein d'un institut peuvent se connecter, à l'aide de leurs outils, au *Data Warehouse* de l'entreprise, voire directement aux sources de données entrantes.

---

<sup>52</sup> MERCANTI-GUÉRIN Maria, « L'amélioration du reciblage par les Big Data : une aide à la décision qui menace l'image des marques ? », dans *Revue internationale d'intelligence économique*, 2013, pp. 153-165. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-internationale-d-intelligence-economique-1-2013-2-page-153.htm>

Attention toutefois : cela peut entraîner et entraîne déjà en partie la hausse du ciblage de la clientèle et la collecte d'informations sur celle-ci, notamment à travers le web. A terme, cela peut provoquer une dégradation de l'image de marque des entreprises et, par conséquent, un rejet collectif. C'est pourquoi il faut parvenir à modérer cette pratique, d'une part, pour ne pas en venir à porter atteinte à la propriété privée et aux données personnelles des individus. D'autre-part, pour ne pas provoquer de phénomène d'autodestruction de cette méthode d'analyse, comme évoquée plus haut, à travers un sur-ciblage et une collecte excessive des données.

Une autre voie d'intégration de la *BI* au sein des instituts d'études et de sondage est la réalisation de tableaux de bords de la performance et de pilotage stratégique. Nous l'avons vu plus tôt, l'un des axes majeurs de l'informatique décisionnelle est le reporting des données au sein de *Dashboards*. Il s'agit en effet de la principale méthode de restitution des données à l'aide de ces outils. Or notons que les instituts d'études et sondages ne réalisent pas uniquement des enquêtes externes, c'est pourquoi l'audit et le suivi des performances internes rentre également dans le champ de compétence de ce secteur. Les grands instituts, tels que l'IFOP ou *Ipsos*, disposent même de pôles dédiés à certains secteurs, tels que les banques et les assurances. Ceux-ci sont ainsi composés de spécialistes du domaine étudiés et sont par conséquent en mesure d'interpréter les indicateurs de performances de ces entreprises.

Dans un cas d'application de la sorte, l'outil *BI* est alors directement connecté sur les bases de production internes et non sur les CRM ou les bases clients. Le cabinet prestataire pourra alors construire un outil de suivi de la performance à l'aide des *KPI* clé du secteur et intégrer des éléments comparatifs sur la durée afin de mettre en lumière les évolutions. Encore faut-il déterminer quels indicateurs vont être calculés et intégrés au tableau de bord. En effet, la création d'un tel outil induit l'instauration d'un dialogue prestataire-client régulier et une adaptation aux sources de données privilégiées de l'entreprise. La construction et la validation du tableau de bord se fait de manière conjointe, comme nous le rappelle Alain Fernandez, indiquant qu'un tableau de bord n'est pas un instrument de mesure standard et figé, il doit être modulable. Pour lui, « chaque décideur a sa manière personnelle d'aborder le pilotage. [...] Un tableau de bord est un instrument personnel. »<sup>53</sup>.

---

<sup>53</sup> *Op. Cit.* FERNANDEZ, *Les Nouveaux Tableaux de bord des managers, Le Projet Business Intelligence clés en main*, 2013, 662 p.

Enfin, l'implémentation de la *BI* au sein d'un institut de sondage peut également servir au pilotage interne de ce dernier. En effet, de la même façon qu'il l'emploie pour ses clients, les outils peuvent collecter, transformer et harmoniser les données issues du Système d'information du cabinet. Qu'il dispose déjà de tableaux de bord du pilotage de son activité ou non, une gestion interne présente des avantages :

- **Le cabinet ne dispose pas encore d'outil *BI* pour assurer son pilotage :** Dans cette situation, le déploiement d'un outil est un véritable manque à gagner pour l'entreprise. Peut-être celle-ci exploite-t-elle ses données de façon éparses, sans que ces dernières ne soient réellement centralisées et interconnectées. Quoi qu'il en soit, la mise en place d'un système de suivi de l'activité automatisé et en temps réel constitue un tournant en termes de maturité *data* de l'entreprise. La mesure d'indicateurs propres au secteur des études viendra éclairer les prises de décisions et le management de l'institut : le taux de retour des questionnaires diffusés ou encore les tonalités de réponses sur certains sujets récurrents.
- **Le cabinet dispose d'un outil *BI* de pilotage externe assuré par un prestataire :** Il s'agit de la situation la plus fréquente, preuve en est que les instituts proposent eux-mêmes des services de production d'outils de pilotage externalisés. Cependant, le principal avantage d'une gestion interne des solutions de suivi des performances est d'en assurer la maintenance et y apporter des modifications. En effet, lors d'une gestion externe, le cabinet n'a pas la mainmise sur le fonctionnement de l'outil et le chargement des données. Si une modification, même minime, doit y être apportée, celle-ci ne peut pas être réalisée rapidement et induit un contact avec la société sous-traitante. Notons également que lors de la rupture du contrat entre le cabinet et son prestataire, l'outil devient alors inopérant du fait de l'incapacité technique à assurer sa gestion en interne.

De plus, soulignons un pan spécifique de l'exploitation interne de la *Business Intelligence* : Le pilotage des processus d'enquêtes. En effet, nous avons déjà signalé que la *BI* pouvait se positionner de façon complémentaire aux sondages et cela concerne également leur gestion. Il est possible d'extraire des études réalisées un certain nombre de « métadonnées » d'enquête. C'est-à-dire, des données caractérisant l'enquête, comme sa date de début, sa date de fin, son taux de réponse, le taux de satisfaction client à l'issue de la mission, ou encore le moment de la

journée privilégié pour répondre (dans le cas d'une enquête en ligne). L'ensemble de ces métadonnées peuvent alors être stockées au sein d'une base et, par conséquent, faire l'objet d'une analyse approfondie à l'aide de la *BI*. Les divers éléments à disposition peuvent aider à appréhender un phénomène, tel qu'un faible taux de réponse aux enquêtes par exemple, et à déterminer les variables impactantes sur celui-ci. Le *Groupement d'Intérêt Scientifique (GIS) Marsouin* a par exemple mis en place ce système de pilotage *BI* en 2019 pour un sondage qu'il réalisait auprès des étudiants de Bretagne, afin de parvenir à expliquer un taux d'engagement particulièrement faible<sup>54</sup>.

Maintenant que nous avons observé ces cas d'application théorique, nous allons nous pencher dans une troisième partie de notre étude, sur un projet de déploiement d'outils *BI* concret au sein d'un institut d'étude et sondage spécialisé.

---

<sup>54</sup> IMT Atlantique – *Business Intelligence*, « Pilotage des processus d'enquêtes ». URL : [http://formations.imt-atlantique.fr/BI/BI\\_projet\\_sondage.html](http://formations.imt-atlantique.fr/BI/BI_projet_sondage.html)

### **III. Cas pratique : Mise en place et exploitation des outils *BI* au sein d'un cabinet d'étude et de conseils en communication**

#### **A. Présentation du projet**

Dans cette partie, nous allons mettre en pratique les grands concepts de l'informatique décisionnelle observés précédemment dans un cas concret de déploiement d'outils de *Business Intelligence* en entreprise. L'étude théorique de la naissance de cette discipline et de ses grandes tendances actuelles nous permettent proposer des solutions en adéquation avec les enjeux du secteur. Ainsi, commençons par présenter notre projet, mené sur près de 18 mois en collaboration avec les équipes du cabinet d'étude *Occurrence*.

##### **a. Contexte de mise en place**

###### **1. Besoins concrets**

Tout d'abord, rappelons que ce projet est à l'initiative d'*Occurrence*, qui a opté pour le recrutement d'un apprenti *Data analyst* dans le but d'explorer les pistes d'utilisation potentielles des outils *BI* pour l'entreprise. Ainsi, la première étape de tout projet *data* est celle de la compréhension *data* et métier. C'est-à-dire la compréhension des données dont dispose l'entreprise, mais également les réalités de terrain et le fonctionnement de l'activité. Ayant été recruté en premier lieu pour mes compétences relatives à l'exploitation des données, je ne suis pas un spécialiste de la communication, ni des études. Il m'a donc fallu comprendre les missions, le déroulement des études et des prestations de conseils, afin d'en identifier les points de blocages et leur résolution potentielle à l'aide de la *BI*.

Aussi, nous ne reviendront pas sur le fonctionnement des missions, ni sur le contexte *data* d'*Occurrence*, que nous avons déjà largement développé en première partie de cette étude. Mais à l'issue de cette étape, j'ai pu lister un certain nombre de besoins concrets pour le cabinet que voici :

- **Mise en valeur des résultats** d'analyses : Il s'agit d'un besoin propre à l'activité d'*Occurrence*, à la fois en interne et pour ses clients. Etant un cabinet d'études et de conseils en communication, nous livrons à nos clients une analyse des résultats d'études. Cette analyse repose ainsi sur des données chiffrées pour lesquelles nous devons valoriser les informations les plus impactantes. Aujourd'hui, la restitution est principalement réalisée sous la forme d'un document PowerPoint composé d'une série de graphiques et d'analyses. En ce sens la *Business Intelligence* a un rôle majeur à jouer, proposant des supports d'interprétation plus dynamiques et interactifs.
  
- **Automatiser certains process** interne et externe : Les process de traitement et de nettoyage de la donnée collectée nécessitent aujourd'hui une grande part d'intervention humaine. Plus les données sont conséquentes, plus leur exploitation est longue et fastidieuse. Or, de nombreux outils de *BI* sont aujourd'hui spécialisés dans la fluidification et l'automatisation de ces étapes, générant un gain de temps non négligeable sur l'ensemble de l'activité.
  
- **Moderniser l'offre tableau de Bord** : Outre les études qualitatives et quantitatives, *Occurrence* propose également des missions de conseil, par lesquelles le suivi des performances de la communication. Pour cela, l'entreprise développe des tableaux de bords sur *Excel* pour ses clients. Bien que ce dernier soit extrêmement puissant, il pose également des contraintes, telle que la nécessité pour les clients de saisir leurs données dans le tableur, bien qu'elles soient déjà présentes sur d'autres sources. Les solutions *BI* offrent la possibilité d'acheminer des données directement depuis plusieurs bases afin de réaliser des tableaux de bord multi-sources. Par ailleurs, les options graphiques de ces outils permettent bien souvent de produire un contenu plus esthétique et attractif qu'*Excel*.
  
- **Toucher de nouveaux publics** : Nous l'avons dit, la *Business Intelligence* est à la mode. De fait, l'exploitation de ces nouveaux outils par l'entreprise lui permettrait d'élargir son offre et de gagner de nouveaux contrats.
  
- **Favoriser l'interopérabilité** des outils en interne : Aujourd'hui, l'architecture du système d'information d'*Occurrence* n'a pas été conçu pour créer des passerelles entre les outils et ne favorise pas l'acheminement des données. Aussi, les supports de

l'informatique décisionnelle apportent des solutions de centralisation et d'harmonisation de la donnée à l'aide de passerelle entre les outils et de flux de travail (*workflow*) étendus.

Ainsi, *Occurrence* dispose depuis longtemps d'outils de suivi et d'une offre tableau de bord. A l'aide d'*Excel* notamment, l'entreprise touche donc du doigt les prémices de la *Business Intelligence*, mais elle manque de logiciels adaptés à son activité. Or nous l'avons vu, les solutions d'informatique décisionnelle ont davantage à apporter qu'un simple suivi de l'activité. Elles ont pour but de centraliser l'information et de l'organiser à des fins opérationnelles.

## 2. Equipe

Ce projet de déploiement d'outils ne respecte pas une méthodologie conventionnelle. En effet, nous n'avons pas mis en place de planning prévisionnel, estimé un budget ni adopté de schéma agile respectant un certain nombre d'étapes. Nous avons fonctionné de façon empirique, en réalisant de nouveaux tests à mesure que s'accroissait l'expérience acquise. Cependant, ce projet a été chapeauté par une équipe opérationnelle de petite envergure, composée de collaborateurs ayant une sensibilité avec les thématiques abordées.

Pour ma part, j'ai officié en tant que *Data analyst*, en charge de l'exploration et du déploiement des nouvelles solutions *BI* chez *Occurrence*. J'ai été, durant l'intégralité de la mission, le membre opérationnel de l'équipe, acquérant la compétence et réalisant les expérimentations successives. J'ai ainsi pu mettre à profit mon expertise *data* à destination des collaborateurs afin de développer les bonnes pratiques de la gestion des données en entreprise.

A mes côtés, Johann Damen, travaillant en tant que Tech lead, responsable SI et directeur de la qualité chez *Occurrence*. Notamment présent pour superviser le projet, nous concertions afin de déterminer les différentes étapes que nous avons suivi à mesure que nous identifions les possibilités techniques des outils. Également spécialiste des missions d'étude et salarié de longue date pour le cabinet, il a ainsi permis de recentrer le projet autour des besoins et des réalités du métier.

Enfin, d'autres collaborateurs ont été sollicités lors de besoins spécifiques du projet : un graphiste pour proposer des idées de visuels innovants, les chargés de missions en charge de la réalisation des tableaux de bords *Excel*, etc. Bien qu'étant personnellement rattaché à un pôle,

la *BI* est un besoin global chez *Occurrence* et tous les services sont concernées. Par conséquent, j'ai été amené à collaborer avec de nombreuses personnes, tant pour identifier leurs besoins métier que pour les aider dans leur démarche *BI*.

### 3. Objectifs projet

Avant de passer concrètement au déploiement des outils, nous avons donc établis une correspondance entre les besoins identifiés précédemment et des objectifs que nous souhaiterions idéalement atteindre à l'issue du projet. Bien qu'étant dans une démarche exploratoire, cela nous a ainsi permis de nous fixer un cadre de départ dont nous pourrions nous servir de point de repère lors des différentes étapes et nous recentrer si besoin.

D'une part, il s'agit d'accompagner *Occurrence* dans la transition progressive de ses outils de suivi et d'analyse pour se diriger vers des solutions plus modernes et flexibles, accompagnant les prises de décision, tant en interne qu'en externe :

- En interne :
  - **Réaliser des tableaux de bord de suivi de l'activité** venant en complémentarité des actuels tableaux *Excel* utilisés.
  - **Développer une véritable culture *data*** d'entreprise et sensibiliser les collaborateurs sur les bonnes pratiques en matière de *BI*.
  - À terme, **former d'autres employés** aux nouveaux outils déployés afin d'assurer un transfert des compétences.
  
- En externe :
  - **Proposer de nouveaux services** à forte valeur ajoutée pour les clients d'*Occurrence*.
  - **Déployer des outils plus esthétiques** et surtout plus dynamiques afin de **renouveler l'expertise du pôle *Corporate*** en matière de tableaux de bord.

D'autre part, un autre objectif serait de parvenir à identifier plus précisément les limites des outils actuels et tenter de trouver des réponses à l'aide de solutions *BI*. Ainsi, nous pensons qu'une compréhension des faiblesses actuelle et leur résolution, couplé avec une hausse

générale des compétences *data* interne, sont la clé d'amélioration du niveau de maturité *data* d'*Occurrence*.

Aussi, il est vrai que les objectifs affichés sont ambitieux, mais n'oublions pas que nous sommes dans une démarche de test. L'idée est, avant toute chose, d'explorer les outils et d'identifier les éléments susceptibles de répondre aux besoins de production. Nous ne savons pas si tous ces objectifs seront atteignables, ni la durée nécessaire pour les réaliser, mais tout le travail de recherche qui sera effectué permettra à *Occurrence* de progresser sur le sujet de la *BI*.

## **b. Approche méthodologique**

D'un point de vue méthodologique, rappelons que personne n'est spécialiste du sujet en interne. De ce fait, nous nous trouvons dans le cadre d'un projet de type « exploration », comprenant de nombreuses incertitudes quant à sa réalisation et sa finalité. Cela implique, sur les divers outils explorés, de fonctionner de façon itérative, composé de phase de veille, de tests et d'observation des résultats successifs. Une fois un résultat satisfaisant obtenu, répondant aux problématiques métiers et à nos objectifs, nous procéderons à sa validation et à un déploiement à destination d'usages concrets. Cette méthode projet s'appuie sur le principe de « *Proof of Concept* ».

Concernant les outils explorés, nous commencerons par nous pencher sur le logiciel *Microsoft Power BI*. Nous en détaillerons le fonctionnement et les avantages plus tard, mais nous avons effectué ce choix dans un contexte de continuité des tests. En effet, *Occurrence* avait déjà commencé à s'intéresser à cet outil avant mon arrivé et un précédent *Data analyst* (plus en poste aujourd'hui) avait effectué de premières recherches sur le sujet. Le cabinet a donc souhaité continuer à explorer cette piste avant d'envisager d'autres solutions.

Ainsi, le tableau ci-dessous présente une estimation en amont des différentes grands étapes su projet sur les sujets principaux, ainsi que la répartition du temps investi :

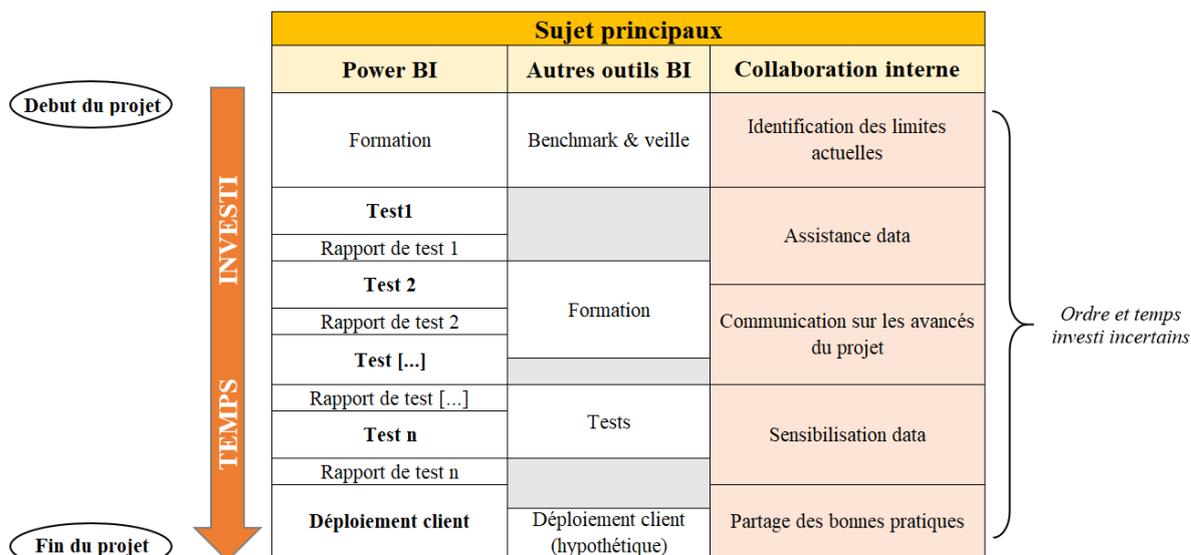


Tableau prévisionnel des étapes majeurs du projet

Ainsi, nous pouvons constater que la priorité est donnée au logiciel *Power BI*, mais nous prévoyons en parallèle de réaliser une veille sur d'autres outils *BI* afin d'élargir le champ des possibilités d'*Occurrence* en matière de déploiement. Nous avons par ailleurs que nous n'oublierons pas les objectifs de collaboration interne fixés précédemment. D'une part, en communiquant avec l'ensemble des collaborateurs sur le projet et les bonnes pratiques afin de les sensibiliser aux enjeux de la donnée. D'autre part, en partageant mon expertise *data* sur les différentes problématiques métiers du quotidien : aide au nettoyage et aux traitements, assistance *Excel*, estimation de la faisabilité de certains projets, etc.

## **B. Mise en œuvre**

### **a. Déploiement de Microsoft Power BI**

Maintenant que nous avons posé les bases du projet, nous allons présenter sa mise en œuvre concrète. Les tests réalisés et les phases de déploiement ont eu lieu entre janvier 2021 et le mois de juin 2022, soit une période de 18 mois, largement dédiée à la veille et aux expérimentations en matière de *Business Intelligence*. Aussi, comme nous l'avons dit, notre étude a débuté par la prise en main du logiciel *Power BI*. Il s'agit d'un outil d'informatique décisionnelle complet, leader du marché de la *BI* avec ses concurrents directs tels que *Tableau*

et *Google Data Studio*. Cette partie est ainsi dédiée à la présentation de l'outil et surtout sa mise en place afin de répondre aux besoins concrets d'*Occurrence*.

### **1. Avantages de Power BI**

En premier lieu, notons que le choix de cet outil n'est pas anodin. En effet, il existe aujourd'hui des centaines de logiciels *BI* et il est parfois difficile de se situer face à la taille de l'offre proposée. Mais en réalisant des comparatifs, on constate que toutes les solutions disposent de fonctionnalités différentes et qu'il est essentiel de sélectionner un outil en adéquation avec les besoins de l'entreprise. De fait, *Power BI* dispose de plusieurs avantages susceptibles de faire défaut à la concurrence :

- **Rapport qualité/prix excellent** : La licence individuelle d'utilisation est proposée au tarif de 8,40€ mensuel et est même inclus dans la suite *Microsoft 365 E5*, dont certaines entreprises disposent déjà. Cette solution est donc adaptée au portefeuille des petites entreprises alors qu'il s'agit d'un leader du marché. *Occurrence* disposant de peu de moyens, elle ne pouvait se permettre d'investir massivement dans un outil sans garanti de rentabilité. De fait, *Power BI* est apparu comme le compromis idéal.
- **Outil hybride** : *Power BI* n'est pas uniquement un outil de visualisation des données, il est possible de réaliser en amont de la manipulation et du traitement sur les bases importées afin de les nettoyer et de les harmoniser. Comme nous l'avons identifié, le cabinet dispose de sources de données internes diverses qu'il est nécessaire de consolider avant de procéder aux traitements.
- **Interactivité** : Le logiciel permet la création de tableaux de bords dynamiques, pensé pour être consultés en ligne et de façon interactive. Les utilisateurs peuvent manipuler de nombreux éléments de filtres pour faire varier les visuels ou naviguer dans le rapport. De plus, il s'agit d'un outil doté de fonctionnalités *responsive*, de telle sorte que les *Dashboards* créés puissent être adaptés à un portage sur mobile ou tablettes. Enfin, *Power BI* propose de nombreuses fonctionnalités de partage *Online* et les rapports peuvent facilement être intégrés à des sites web.

- **Automatisable** : Dans une certaine limite, il est possible de planifier l'actualisation automatique des données chargées dans l'outil. De ce fait, les nouvelles données stockées dans la source sont automatiquement ajoutées aux rapports, sans intervention humaine. Cet aspect est déterminant pour *Occurrence* : dans le cadre d'un déploiement client, ce dernier ne doit pas avoir besoin d'être formé pour suivre son activité à travers l'outil.
- **Simple pour une utilisation sommaire** : *Power BI* est adapté aux équipes métiers, élément important si l'on envisage de former des collaborateurs à l'avenir<sup>55</sup>. Il s'agit en effet d'un outil *no-code*, comme nous l'avons déjà évoqué, donc il n'est pas nécessaire de posséder des compétences en programmation pour l'utiliser. Il est cependant possible d'incorporer du script Python ou R sur le logiciel pour effectuer des traitements statistiques de pointe.
- **Esthétique** : Le rendu visuel des tableaux de bords est plus graphique et moderne que sur *Excel* ou *Power Point*. Or, cet enjeu de modernisation est l'un des objectifs que nous avons soulignés dans le cadre de notre projet.
- **Outil Microsoft** : *Power BI* dispose de plus de 150 sources disponibles pour se connecter et charger des données. De plus, étant un outil *Microsoft*, il dispose d'interactions privilégiées avec les autres outils du même éditeur, que nous détaillerons plus tard. Comme *Occurrence* est titulaire de la suite *Microsoft* qu'elle utilise au quotidien, le logiciel *BI* du même éditeur nous a donc semblé particulièrement adéquat.

## 2. Phase 1 : Formation à l'outil sur les standards en interne

Lors de mon arrivée chez *Occurrence*, le premier cas d'utilisation concrète de *Power BI* a été de réaliser un tableau de bord des Standards du pôle « *Corporate* ». Nous les avons déjà présentés (cf *I.A.e.2. – Travail sur les Standards*), les standards représentent des données de référence pour de nombreux indicateurs d'évaluation de la communication. Actuellement sous

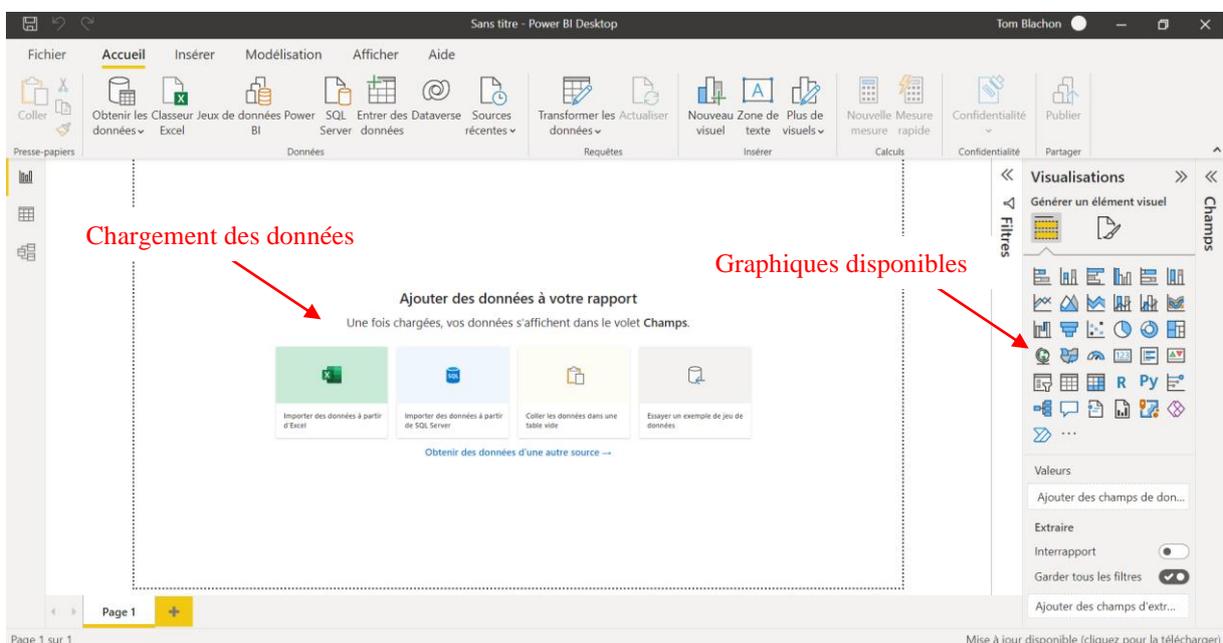
---

<sup>55</sup> BELHAMRI Nicolas, « OUTILS DE BUSINESS INTELLIGENCE : LE TOP 10 EN 2022 (COMPARATIF) », Boryl, URL : <https://www.boryl.fr/blog/top-10-des-outils-de-business-intelligence/#8>

forme de base de données *Excel*, le pôle souhaiterait réaliser une série de graphique afin d'illustrer ces données.

Il ne s'agit toutefois que d'un test. L'objectif est de partir de données concrètes pour se former sur l'outil, en comprendre le fonctionnement global, les limites et les usages potentiels. L'idée n'était pas d'obtenir un résultat exploitable directement, mais plutôt d'observer les résultats du test et d'en faire le bilan. Celui-ci s'est ainsi déroulé en plusieurs étapes :

- **Etape 1 :** Formation théorique à l'aide de la documentation *Microsoft* (en français et de bonne qualité), utilisation de tutoriels pour bien débuter.
- **Etape 2 :** Compréhension du tableau déjà réalisé dans le cadre de la passation. En effet, l'ancien *Data analyst* d'*Occurrence* ayant déjà travaillé sur *Power BI* avant réalisé un travail similaire sur les standards d'un autre pôle. La compréhension de son travail m'a été d'une grande aide pour progresser sur l'outil.
- **Etape 3 :** Premiers tests de l'outil, premier import de données et premiers graphiques. Il m'a fallu pour cela prendre en main de l'interface, dans lequel nous retrouvons bien l'apparence propre aux outils *Microsoft* :



Interface de création de Power BI Desktop

- **Etape 4 :** Initiation à la manipulation de bases à l'aide de *Power Query*. Nous le présenterons plus tard, mais il s'agit d'un outil intégré de transformation des données après leur chargement. Cela permet notamment d'unifier la structure de bases originellement différentes.
- **Etape 5 :** Construction du rapport de façon empirique. Pour cela, je procédais à l'import d'une première base de standards, créais le tableau de bord de celle-ci, puis importait une nouvelle base en profitant de l'expérience acquise progressivement.
- **Etape 6 :** Export de la version finale et publication *Online*, distinction entre *Power BI Desktop* et *Power BI Service* (nous y reviendrons).
- **Etape 7 :** Présentation des résultats du test à l'équipe. Il s'agit d'un tableau de bord de 11 onglets permettant une vue d'ensemble des indicateurs pour chaque catégorie de standards. Comme il s'agissait d'un premier test, j'ai souhaité mettre en valeur l'interactivité de l'outil, à l'aide de boutons de navigation entre les onglets et de filtres sur chaque page du tableau de bord. Lors de la présentation, les retours furent globalement positifs sur l'aspect visuel et le travail a été jugé encourageant pour de potentiels usages à venir.

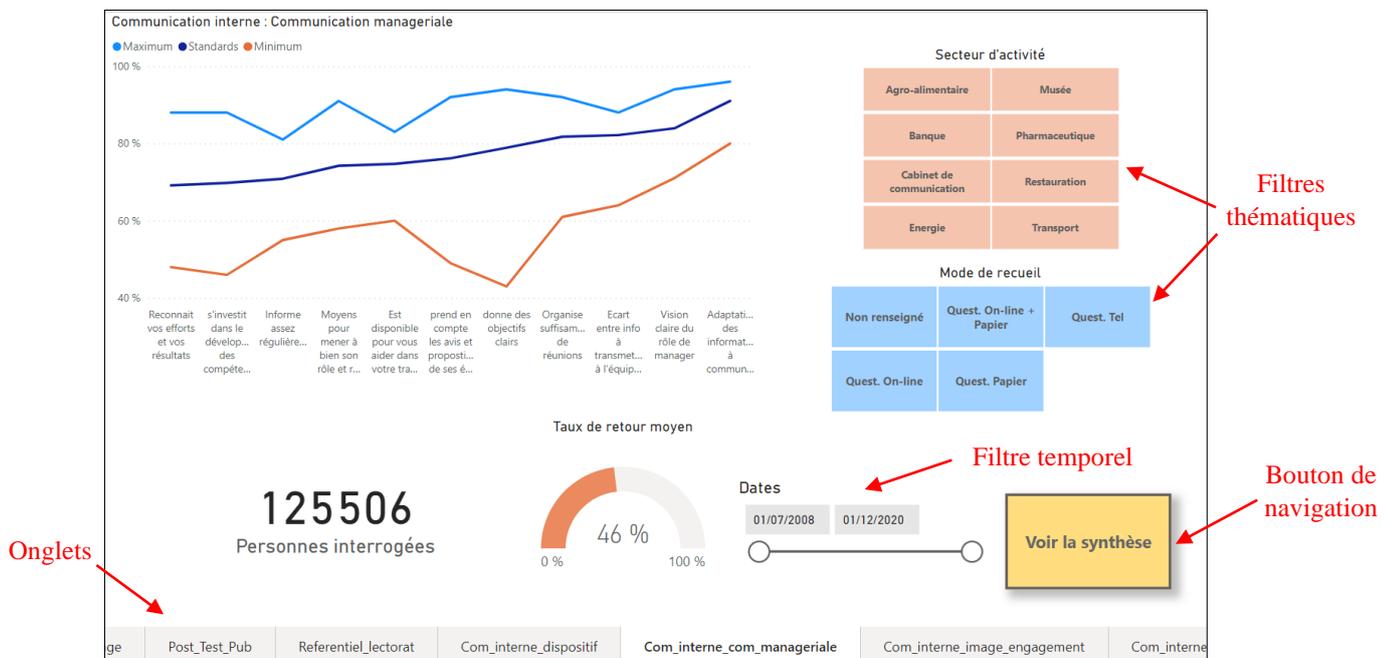


Tableau de bord des standards de communication interne

En conclusion, ce premier essai de l'outil en autodidacte s'est étendu sur une période assez longue, puisque le développement a débuté en janvier et s'est achevé fin mars 2021, à raison d'environ un jour et demi de travail hebdomadaire sur l'outil. Bien qu'à posteriori, toutes les bonnes pratiques en matière de conception de tableau de bord n'ont pas été respectées, ce test fut relativement convainquant et formateur. Il m'a permis de comprendre le processus de création de tableaux de bord dans sa globalité. De plus, la possibilité de centraliser plusieurs sources de données externes au sein d'un même rapport à l'aide de l'outil a également pu être expérimenté.

### 3. Phase 2 : Tests de substitution des outils actuels (Excel)

Une fois la première phase de formation terminée, nous avons voulu aller plus loin en effectuant un nouveau test : la reproductibilité d'un tableau de bord existant. En effet, afin de s'assurer que *Power BI* soit en mesure de substituer *Excel* en matière de tableau de bord, nous avons souhaité créer à l'identique un *Dashboard* déjà réalisé sans avoir à en modifier la source de données.

Celui-ci avait été créé sur *Excel Online* à partir de données issues d'autres onglets au sein du même classeur. Cela constituait donc également une occasion d'essayer de récupérer des données à partir d'une source *Cloud* et d'établir des relations entre les différentes bases importées (chaque onglet du tableur représentant une base distincte dans *Power BI*).

Ce second test a aussi été une opportunité de m'initier aux mesures en langage *DAX*. Ces « mesures » sont des calculs que l'on peut effectuer sur des champs afin de retourner une ou plusieurs valeurs qui s'adaptent automatiquement en fonction des filtres appliqués sur le rapport. Elles sont notamment indispensables lorsque l'on souhaite calculer des *KPI* sur nos champs dans *Power BI*. Le langage *DAX* peut quant à lui être comparé aux formules *Excel*. Bien que les fonctions soient différentes, il s'agit du langage de rédaction de toutes les formules et expressions dans *Power BI*.

Réalisé en deux semaines, ce test de reproductibilité a été concluant puisque le tableau réalisé sur le logiciel est semblable au *Dashboard Excel Online* d'origine, tant dans son aspect graphique que par ses interactions avec le filtre temporel mis en place. Notons toutefois que la

construction de ce tableau a nécessité quatre fois plus de temps de développement que son homologue sur *Excel*. Bien qu'un expert puisse aller plus vite, il s'agit tout de même d'un élément très important à prendre en compte lors de l'utilisation de l'outil *BI* auprès de clients, notamment en termes de facturation sur le temps passé.

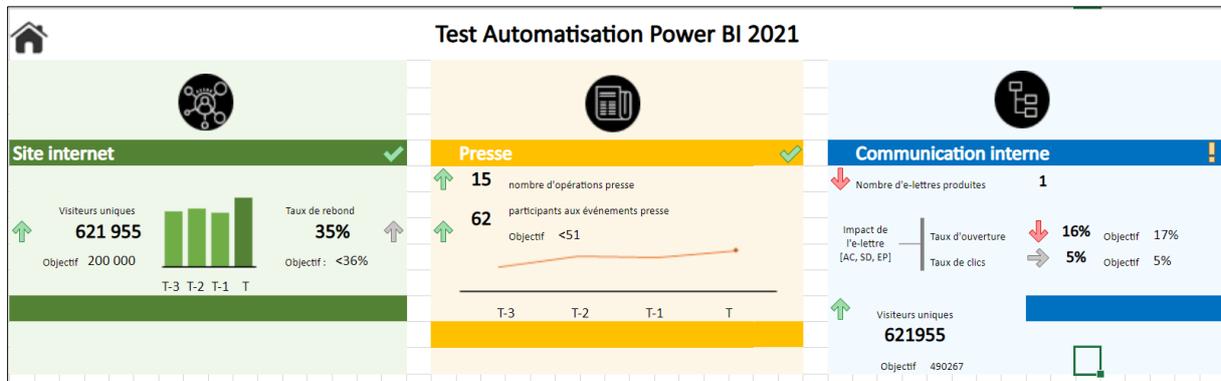


Tableau de bord test réalisé sur Excel Online (données fictives)

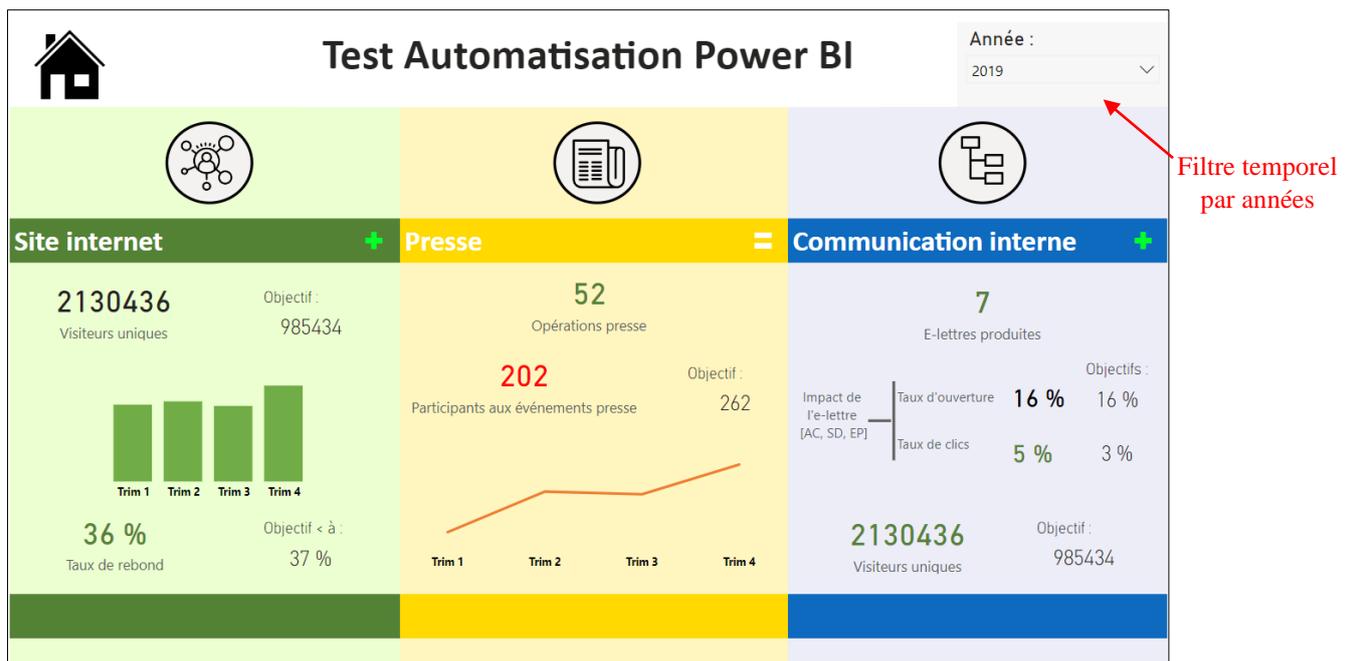


Tableau de bord test réalisé sur Power BI (données fictives)

#### 4. Phase 2.5 : Formation à l'automatisation

En parallèle du test de reproductibilité, nous avons souhaités explorer la piste de l'automatisation des rapports. Nous l'avons dit plut tôt, l'un des avantages de *Power BI*, de la même façon que ses concurrents, est la possibilité de planifier l'actualisation automatique des données du tableau de bord. Ces derniers sont avant tout des outils de pilotage et de suivi, il semble donc primordial que les indicateurs et les graphiques présentés soient maintenus à jour sans que cela ne requiert une intervention humaine quotidienne. Chez *Occurrence*, les données saisies sur *Excel* s'actualisent automatiquement directement sur le *Dashboard*. Il est donc nécessaire de parvenir à obtenir un résultat similaire à l'aide de *Power BI*.

Pour comprendre comment fonctionne cette notion d'actualisation, nous devons nous pencher sur l'architecture globale de *Power BI*. En effet, le logiciel est en réalisé divisé en deux outils principaux :

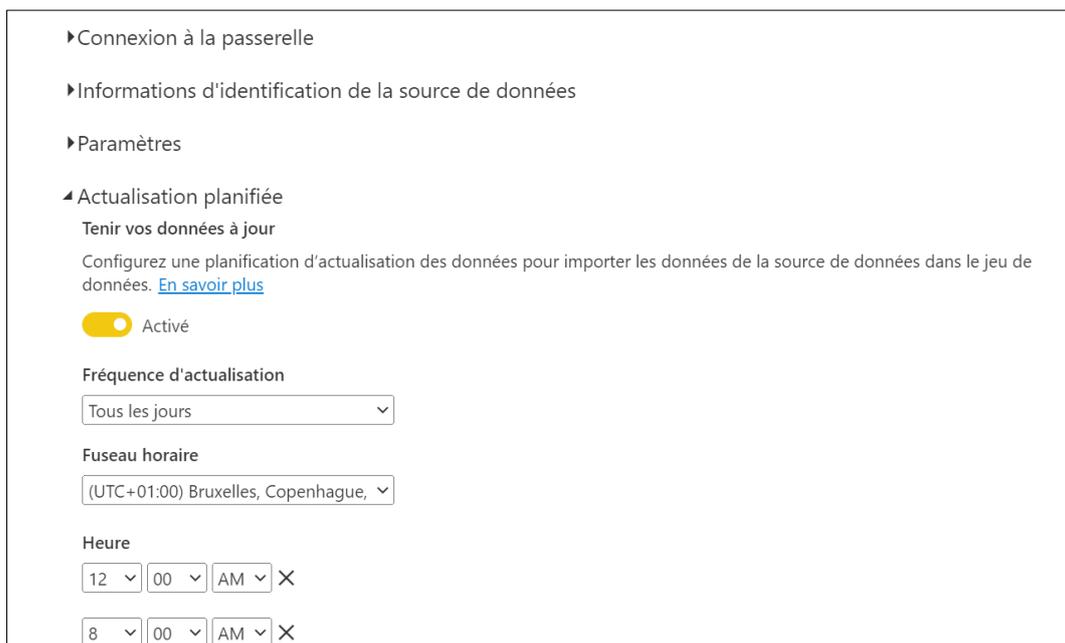
- **Power BI Desktop** : Logiciel *On premise* (installé sur une machine et stocké localement), il s'agit de l'outil de création des rapports et permet de charger des données depuis une source de donnée externe locale ou en ligne. Ce logiciel permet notamment la transformation des données à l'aide de *Power Query* et la modélisation des tableaux de bord.
- **Power BI Service** : Logiciel Saas (stockée dans un *Cloud*), il permet la gestion des tableaux de bords existants, leur partage et leur publication.

Ainsi, le cheminement classique (appelé *workflow*) consiste à charger les données depuis une ou plusieurs sources vers *PBI Desktop*, de construire son tableau de bord sur ce logiciel, puis de le publier sur *Power BI Service* afin de le partager.

Une fois le rapport publié, son actualisation se fait directement via le Service *Power BI*. Or, pour des données stockées en local, telles que des bases *Excel* ou *SQL*, le service *Cloud* ne peut pas accéder aux données. Pour y parvenir, une « passerelle de données » ou *Data gateway* doit être installée localement. Il s'agit d'un logiciel développé par *Microsoft* permettant de faire le lien entre ses outils *Cloud*, tels que *Power Apps* ou encore *Power BI Service* et des données locales.

Nous l'avons vu, la grande majorité des données d'*Occurrence* sont stockées localement. De ce fait, il nous a été nécessaire d'installer et de paramétrer un logiciel de passerelle afin de permettre l'actualisation des données.

Une fois ces étapes réalisées, la procédure de planification de l'actualisation automatique des données est relativement simple. Elle se configure directement depuis *Power BI Service* après activation de la passerelle. Avec une licence *Power BI Pro*, l'actualisation se limite à 8 fois par jours maximum, ce qui est déjà bien suffisant pour les données de pilotage traitées par *Occurrence*.



The screenshot shows the 'Actualisation planifiée' (Scheduled Refresh) configuration interface in Power BI Service. It includes the following elements:

- Navigation menu: 'Connexion à la passerelle', 'Informations d'identification de la source de données', 'Paramètres', and 'Actualisation planifiée' (expanded).
- Section: 'Tenir vos données à jour' (Keep your data up to date).
- Description: 'Configurez une planification d'actualisation des données pour importer les données de la source de données dans le jeu de données. [En savoir plus](#)'.
- Toggle: 'Activé' (Active) with a yellow switch.
- Frequency: 'Fréquence d'actualisation' dropdown set to 'Tous les jours' (Every day).
- Timezone: 'Fuseau horaire' dropdown set to '(UTC+01:00) Bruxelles, Copenhague'.
- Time: 'Heure' section with two rows of dropdowns for hour, minute, and AM/PM, each with a clear (X) button. The first row is set to 12:00 AM, and the second row is set to 8:00 AM.

*Interface de gestion de l'actualisation planifiée sur Power BI Service*

Notons pour finir que les outils de la suite *Microsoft Online*, tels que *One Drive* et *SharePoint* ne requièrent pas de passerelle pour être actualisés, car ils se trouvent dans le même environnement que *Power BI Service*. De ce fait, le tableau de bord réalisé précédemment (cf. Phase 2) a pu être automatisé sans passerelle de données car la source des données est *Excel Online*.

## 5. Phase 3 : Déploiement en interne dans un cas d'usage concret

Au mois de décembre 2021, nous avons procédé à une nouvelle phase de test sur *Power BI*. Il s'agissait du premier test de déploiement réel d'un tableau de bord en interne à des fins de pilotage.

L'idée était en effet de construire un tableau de bord doté de seulement 3 indicateurs et ayant pour source le tableau *Excel* du développement présenté plus haut (cf. *I.B.a.2 - Outils et pilotage*). Celui-ci serait présenté en réunion manager de façon hebdomadaire pour surveiller notamment l'espérance en chiffre d'affaires de l'entreprise.

Ce test a de nouveau permis de progresser sur plusieurs sujets : le paramétrage de l'actualisation planifié, le partage sécurisé des données (entre détenteurs de licence *Power BI Pro*), ainsi que sur le rendu graphique. En effet, le tableau ne contenant que peu d'indicateurs, le visuel se devait d'être impactant. En voici donc le résultat final :



*Tableau de bord du Développement sur Power BI (données de test)*

La complexité de l'exercice a été de reproduire en langage *DAX* le calcul de ces trois *KPI* qui avaient déjà été réalisées sur *Excel*. Il s'agissait de formules à multiples conditions qu'il a été nécessaire de décrypter pour parvenir à les transposer.

Toutefois, si la réalisation du tableau de bord a pu être effectué, ce test n'est pas véritablement un succès. Lors de sa mise en pratique, comme *Power BI* n'est pas un espace de saisie des données, les managers devaient tout de même ouvrir le fichier *Excel* pour modifier la base. De ce fait, le *Dashboard* réalisé n'a pas su convaincre les décideurs qui ne sont pas encore parvenu à l'utiliser en tant qu'outil de pilotage à part entière.

## 6. Phase 4 : Déploiement lors de missions clients

Après plusieurs phases de tests concluants et d'expérimentations ayant permis de monter en compétences sur différents aspects, nous avons décidé de commencer à déployer ce service auprès des clients du cabinet.

Ainsi, depuis le mois de mars 2022, nous avons déjà réalisé trois propositions de conception de tableaux de bords auprès d'entreprises partenaires. Malheureusement, aucune d'entre elles n'a abouti, mais elles ont tout de même permis de tirer des enseignements de ces expériences afin d'anticiper au mieux les propositions à venir. Pour des raisons de confidentialité, nous avons anonymisé ces entreprises que nous appellerons « Société 1 », « Société 2 » et « Société 3 » :

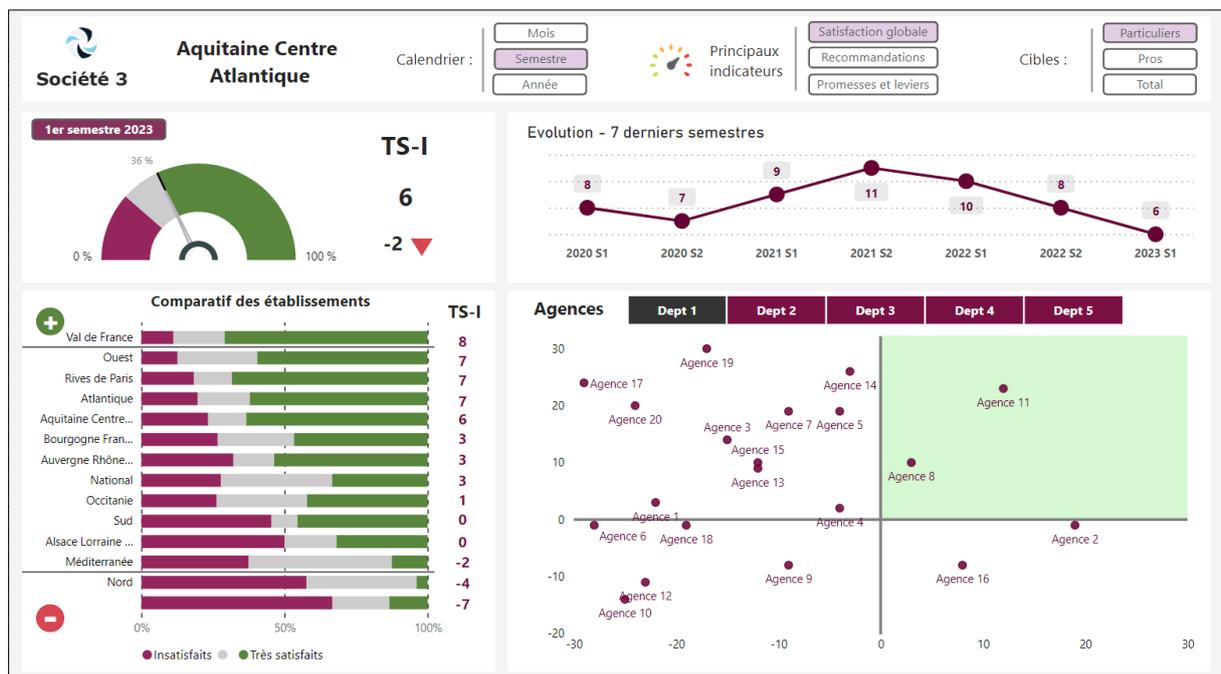
- **Société 1** : Cette société nous a missionné pour la création d'un tableau de bord de pilotage de sa communication. Nous lui avons alors proposé de le concevoir sur *Power BI* et avons réalisé une maquette (Annexe 2)<sup>56</sup>. Le client a été séduit par celle-ci mais a finalement souhaité une solution plus facilement gérable en termes de maintenance, notamment lorsque leur contrat avec *Occurrence* prendra fin. De plus, nous avons rencontré un problème de licence. En effet, le client ne disposait pas de la licence *Power BI* et était retissant à l'idée d'y souscrire.
- **Société 2** : Cette entreprise souhaitait un tableau de bord et a immédiatement exprimé la volonté d'un déploiement sur *Power BI*. Mais il s'agissait d'un tableau incluant plusieurs centaines d'indicateurs. Or nous l'avons dit, le temps de développement sur *Power BI* est plus important que sur *Excel*. Aussi, l'ampleur du travail à réaliser et du temps consacré nous a semblé trop important. Nous avons toutefois réalisé une maquette

---

<sup>56</sup> Annexe 2 : Maquette *Power BI* réalisée pour la Société 1

sur deux onglets afin d'estimer le temps de développement requis en vue de futures missions (Annexe 3)<sup>57</sup>. Nous avons également interrogé un spécialiste de *Power BI* en freelance afin qu'il estime le temps à consacrer. Conformément à nos doutes, l'emploi de *Power BI* pour cette mission aurait été bien trop chronophage pour la facturation convenue. Ainsi, nous avons préféré décliner la demande du client et produire un *Dashboard Excel* standard.

- **Société 3** : il s'agissait de la réponse à un appel d'offre (AO) émis par un grand groupe. Les prérequis spécifiaient, entre autres, la production d'un outil de pilotage de différentes agences réparties partout en France, développé sur *Power BI*. Dans le cadre du récent rachat du cabinet *Occurrence* par l'*IFOP* (Institut Français d'Opinion Publique), nous avons réalisé une candidature commune à cet AO. J'ai pour cela développé une nouvelle maquette de tableau de bord, en collaboration avec les équipes de l'*IFOP*. Nous avons malheureusement perdu cet AO, mais cela a permis de monter en compétences sur *Power BI*, en proposant des graphiques et visuels innovants.



Maquette de tableau de bord *Power BI* réalisée pour la Société 3

<sup>57</sup> Annexe 3 : Maquette *Power BI* réalisée pour la Société 2

## 7. Limites de Power BI et obstacles au déploiement

Le projet de déploiement de *Power BI* au sein d'*Occurrence* aura nécessité beaucoup de temps et d'investissement. Au total, il aura fallu compter plus d'un an entre la découverte de l'outil et les premières propositions de valorisation concrètes. Mais le temps investi est à la hauteur des possibilités techniques du logiciel. En effet, les différents tests réalisés nous ont confirmés qu'il était en mesure de répondre aux besoins spécifiques de l'entreprise.

Cependant, ces expériences ont également été révélatrices des limites techniques de l'outil dans un contexte d'exploitation tel que celui d'*Occurrence* :

- **Difficultés de maintenance** : Dans un contexte client, *Power BI* peut manquer de flexibilité. En effet, assurer la maintenance du tableau de bord en cas de dysfonctionnement requiert une bonne connaissance de l'outil et des caractéristiques des données. Aussi, certains clients privilégient *Excel* pour son aspect d'apparence plus abordable. Lorsque la mission s'achève, *Occurrence* peut proposer un service de maintenance, mais dans d'autres cas, le tableau de bord est laissé à la gestion du client. Cette dernière option est presque inenvisageable avec *Power BI*, à moins qu'il ne dispose d'utilisateurs du logiciel en interne.
- **Pas d'espace de saisie** : Sur *Power BI*, les données sont chargées par défaut en « mode importation », ce qui signifie qu'elles sont copiées et mises en cache dans *Power BI*. A chaque fois que les données sont actualisées, les anciennes valeurs sont écrasées et remplacées par de nouvelles provenant de la source. Par conséquent, si cette copie peut être manipulée et mise en forme sur le logiciel, aucune valeur ne peut être modifiée directement dans *Power BI*. A l'inverse d'*Excel*, il est donc nécessaire de dissocier l'espace de saisie et le tableau de bord, ce qui peut être déroutant pour des non-initiés. Par exemple, c'est ce qui a été reproché au tableau de bord réalisé pour les managers d'*Occurrence* car ces derniers souhaitent pouvoir modifier leurs données directement sur le même outil.
- **Les licences** : Nous l'avons déjà brièvement évoqué, mais l'utilisation du logiciel requiert une licence individuelle appelée « *Power BI Pro* » ou sa version avancée intitulée « *Power BI Premium* ». Sans cela, il est impossible de diffuser et de partager les tableaux de bord réalisés, ni de consulter les *Dashboards* que l'on nous partage de

façon sécurisée. Il est donc nécessaire que le client, comme le concepteur, dispose d'une licence. Si *Occurrence* propose un développement sur *Power BI* à ses clients, il est nécessaire de les alerter en amont sur cette souscription, sous peine de générer de la frustration. Comme dans le cas de la Société 1, le travail réalisé peut même finir par être inutilisable pour cette raison.

- **Temps de développement** : Pour un tableau de la même envergure, le temps de création sera plus long et plus complexe que sur *Excel*. De ce fait, il est nécessaire d'anticiper ces délais dès la facturation de la mission pour ne pas produire à perte.

Pour résumer, *Power BI* ouvre à de nouvelles possibilités et propose des résultats d'une qualité difficilement atteignable à l'aide d'*Excel*. Cependant, l'outil doit parvenir trouver sa cible : il y a certes une demande liée aux tendances du marché, mais il ne convient pas à tous les besoins.

## **b. Visualiser les données spatiales : SIG et cartes de réseaux**

La *BI* inclut la restitution graphique des données, ainsi, les outils de datavisualisation spatialisées peuvent être compris dans ce domaine. La restitution de données spatiales notamment, implique un certain nombre de particularités et des données spécifiques. Disposant d'une compétence en amont sur certains de ces outils, j'ai souhaité proposer une mise en application concrète de ces solutions *BI* spécifiques auprès d'*Occurrence*.

### **1. Échantillonnage**

L'échantillonnage est une étape indispensable lors de la construction d'une étude quantitative. Une fois que la cible de l'enquête est déterminée, il est rare d'être en mesure d'interroger l'intégralité de la population visée. Ainsi, l'échantillon des répondants à l'étude doit être représentatif de cette population pour être reconnue d'un point de vue statistique. Pour cela, le diffuseur de l'enquête se doit de respecter la méthode des quotas, c'est-à-dire, les

proportions de la population d'étude en matière de genre, d'âge, de catégories socioprofessionnelles (CSP), etc.

Outre la représentativité, la significativité de l'échantillon doit également être respecté. C'est-à-dire s'assurer d'un nombre de répondants suffisamment important pour que les résultats obtenus soient assimilables à l'ensemble la population cible d'un point de vue statistique. Plus cette dernière est importante, plus la taille de l'échantillon doit être conséquent. Chez *Occurrence*, on les études quantitatives sont ainsi généralement réalisées auprès de 1000 répondants.

Notons que l'échantillonnage est souvent réalisé par le prestataire en charge de la diffusion du questionnaire, surtout lors d'enquêtes à l'échelle nationale. Les quotas sont bien définis et les organismes spécialisés disposent de panels de répondants garantissant leur respect. Mais dans certains cas spécifiques, où la population n'est pas ciblée au sein d'une entité administrative clairement délimitée, il est plus difficile de fixer les proportions à suivre.

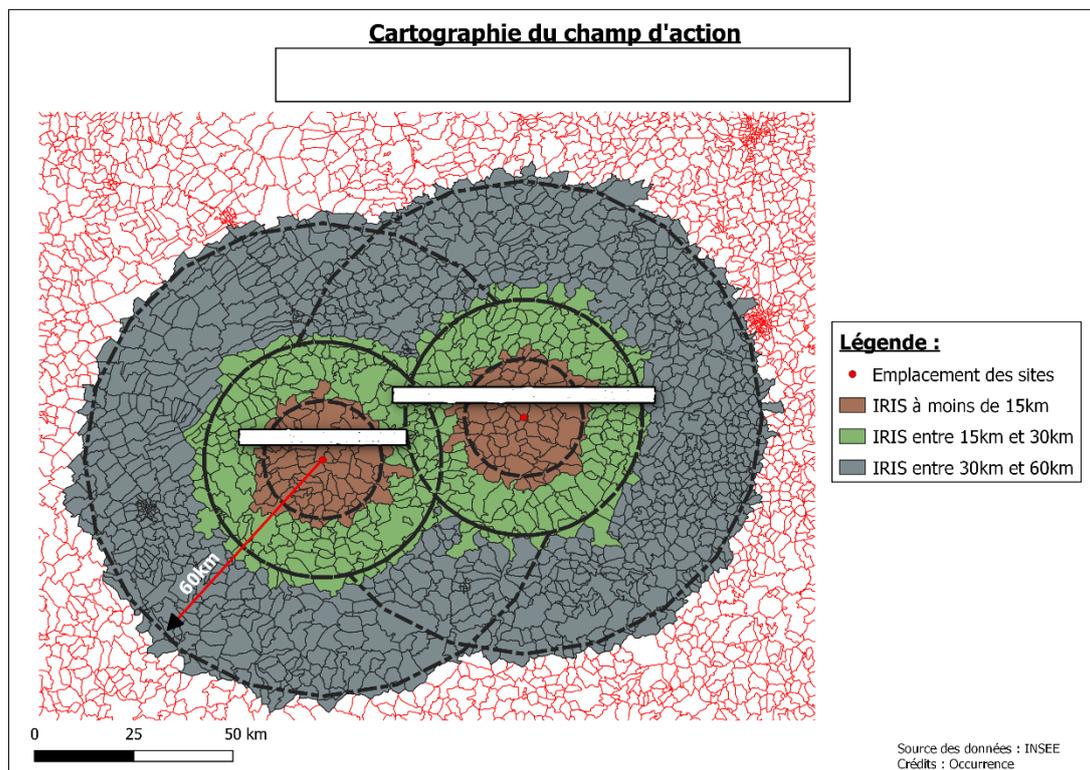
Ainsi, à l'aide d'un logiciel de traitement de données géospatiales et des données de recensement de l'INSEE, il est possible de réaliser manuellement un échantillonnage plus spécifique, comme nous avons eu l'occasion de l'expérimenter chez *Occurrence* (cf. partie suivante)

## 2. Mieux comprendre la répartition des données à l'aide des cartes

Nous venons d'évoquer les logiciels de traitement de données géospatiales. Également appelés SIG (Système d'Information Géographique), il s'agit d'outils assimilables à de la *Business Intelligence*, spécialisé dans le traitement de bases dotées de coordonnées géographiques. A la manière de *Power BI*, il est possible de d'y charger des sources de données externes que nous pouvons manipuler pour produire des visualisations sous forme de carte.

Lorsque l'on possède des données géographiques, les cartes de répartition sont des formes de datavisualisation très impactante. C'est pourquoi j'ai souhaité déployer cette compétence chez *Occurrence* dans le cadre de missions d'études. Nous avons alors opté pour le logiciel QGIS, ayant l'avantage d'être gratuit et libre de droit. Voici donc trois exemples de missions pour lesquelles nous avons pu déployer cet outil en situation réelle :

- **Construction d'un échantillonnage** spécifique : Une société souhaitait mener une étude dont la cible serait les personnes habitant dans un rayon de 15km, 30km et 60km autour de trois usines en France. J'ai alors utilisé QGIS pour identifier les communes françaises situées dans les rayons d'étude. J'ai ensuite croisé ces données avec celles du recensement national réalisé par l'INSEE en 2017. Nous disposons ainsi d'une base des données de recensement, commune par commune, de notre population cible. Une fois cette base exportée, nous avons pu déterminer les quotas à respecter pour chaque rayon et pour chaque usine.



*Cartographie d'échantillonnage spécifique réalisée sur QGIS*

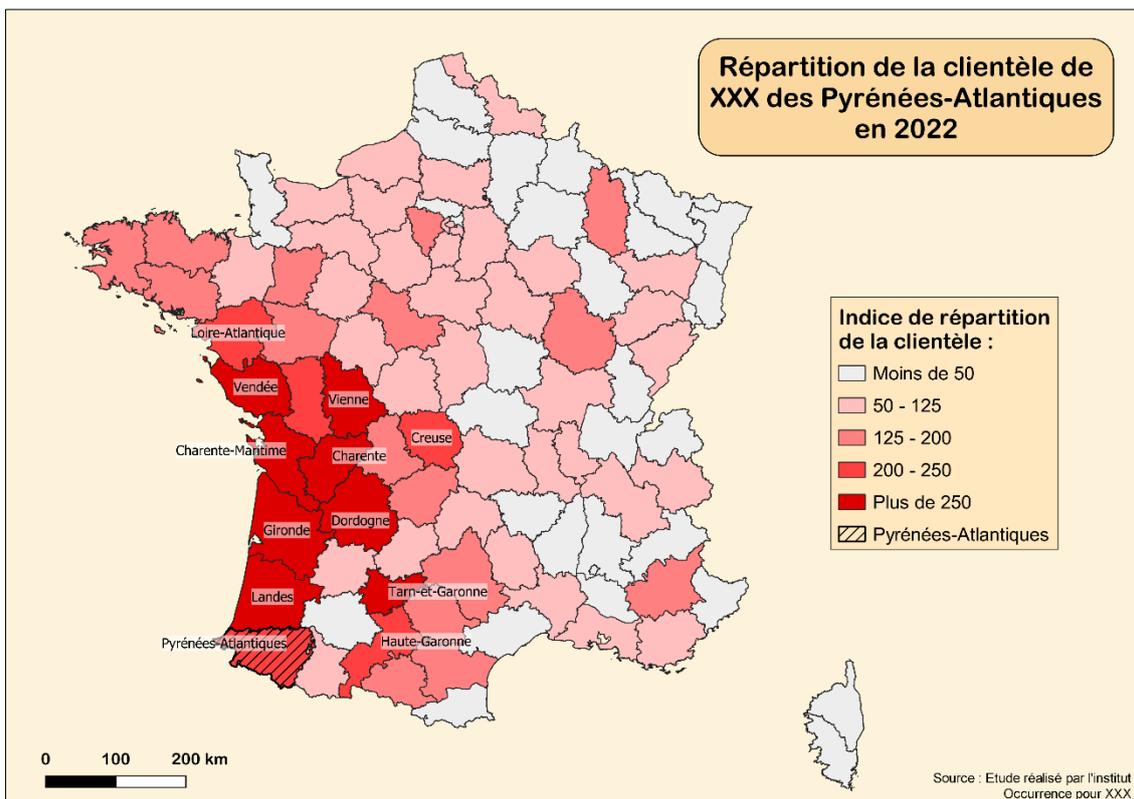
- **Carte de répartition des répondants** : Il s'agit d'une enquête sur les travailleurs « invisibles » réalisée pour la *Fondation Travailler Autrement*<sup>58</sup> et parue dans le *Journal du dimanche* du 13 mars 2022<sup>59</sup>. J'ai alors réalisé des cartes sur QGIS afin de représenter le taux de répartition des « Invisibles » par régions et dans les principales métropoles

<sup>58</sup> DALLE-MOLLE Maris, « Enquête sur « Les Invisibles » pour la Fondation Travailler Autrement », 13 mars 2022. URL : <https://www.fondation-travailler-autrement.org/2022/03/13/invisibles-plongee-france-back-office-etude/>

<sup>59</sup> SOUFFI Emmanuelle, « Enquête sur « Les Invisibles », dans le *Journal du dimanche*, 13 mars 2022. URL : <https://www.lejdd.fr/Politique/enquete-sur-la-france-des-oublies-4099073>

françaises (pour illustrer le fait que ceux-ci étaient moins présents dans les milieux urbains qu'à la campagne). Ces cartes ont ainsi été présentées par le *JDD* dans leur article sur l'étude, bien qu'ils en aient modifié l'aspect général. Cette mission a donc été un succès et génératrice d'une véritable valeur ajoutée de la part d'*Occurrence*.

- **Cartes de connaissance de la clientèle** : Dans le cadre d'une vaste mission d'étude de la clientèle d'un organisme disposant d'agences partout en France à l'échelle départementale. Pour chacune de ces agences, le groupe souhaitait connaître la provenance de ses clients. Pour illustrer cela, j'ai donc réalisé 21 cartes sur QGIS (une par agence), mettant en avant les principaux départements de provenance. Le nombre de cartes que nous sommes parvenus à proposer lors de cette mission confirme l'intérêt de cet outil pour le cabinet d'étude, apportant une expertise dont il ne disposait pas jusqu'alors.



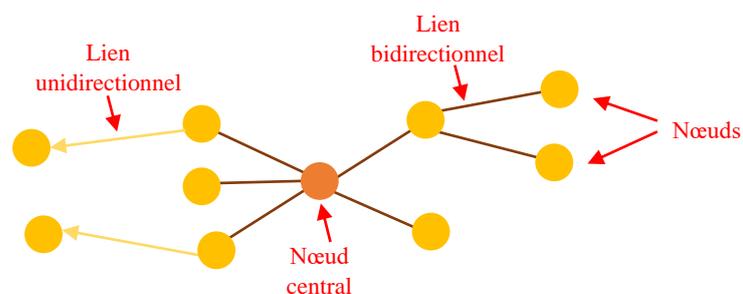
*Carte de répartition de la clientèle de l'agence des Pyrénées-Atlantiques (QGIS)*

### 3. Cas spécifique de la spatialisation des réseaux (Géphi)

Pour en finir avec les outils *BI* de traitement des données spatiales, notons qu'il existe également des données spatiales non géographiques : les cartes de réseaux. Ces dernières permettent en somme de spatialiser en deux ou trois dimensions les relations entre plusieurs individus. L'objectif peut ainsi être de déterminer qui sont les individus au centre du réseau, l'identification de groupes distinctifs au centre du réseau ou la force des relations entre certains individus par exemple. Ici, la notion d'individus ne désigne pas nécessairement des personnes, il peut s'agir de sites web, de posts sur des réseaux sociaux, etc.

Pour définir cette relation au sein d'une cartographie, seuls deux éléments sont indispensables :

- **La liste des nœuds** : il s'agit de la liste des individus de notre réseau, souvent composé d'un numéro d'identifiant unique (*id*) et d'un nom (*label*).
- **La liste des liens** : il s'agit de la liste des relations entre les différents nœuds du réseau. Une ligne est notamment composée de l'*id* du nœud de départ et de l'*id* du nœud d'arrivée. Nous pouvons également mentionner l'intensité du lien entre deux nœuds, sous la forme d'une valeur numérique. Il est aussi possible de préciser si le lien est unidirectionnel ou non. C'est-à-dire, si la relation entre les deux nœuds ne va que dans un seul sens.



*Cartographie de réseau fictive*

Je n'ai pas directement eu l'idée de déployer cet outil chez *Occurrence* car je ne parvenais pas à en discerner les usages pour notre activité, malgré ma connaissance, bien que sommaire, d'un outil approprié. Mais un collaborateur m'a approché en souhaitant réaliser un graphique de réseau et m'a demandé si je disposais des compétences nécessaires. Il s'agissait de matérialiser la relation entre environ 3000 sites internet dans le cadre d'une mission

d'observatoire du web. Pour répondre à ce besoin, j'ai utilisé le logiciel *Gephi*, dédié à la création de cartes de réseau et ayant l'avantage d'être libre et gratuit. Je suis ainsi parvenu à obtenir un résultat mettant en avant plusieurs groupes distincts au sein du graph. Cependant, nous n'avons finalement pas retenu cette solution et avons engagé une société prestataire spécialisée dans ce domaine afin d'obtenir un résultat interactif et facilement intégrable sur le web, dépassant mes compétences actuelles dans ce domaine.

Bien qu'il s'agisse d'un outil de *BI* puissant en termes de datavisualisation, il est certain que son usage demeurera relativement mineur, compte-tenu des besoins d'*Occurrence* face à la spécificité de l'outil.

### **c. Exploration de la suite *Microsoft* : Anticiper la *BI* en amont et multiplier les supports d'interprétation**

Nous avons pu observer précédemment les apports de différents outils d'informatique décisionnelle, ainsi que la façon dont ils répondent aux besoins de l'entreprise. Maintenant que leur utilité est avérée et leur déploiement relativement installé, nous avons souhaité aller plus loin dans notre démarche d'implantation. Cette seconde phase, plus exploratoire, vise à envisager *Power BI* au sein d'un processus plus global, allant de la collecte à l'intégration des tableaux de bords dans des supports d'entreprise dynamiques.

Ayant l'avantage de faire partie de la suite des outils *Microsoft*, *Power BI* dispose de connexions privilégiées avec les autres logiciels de la suite. L'éditeur, pour des raisons pratiques et commerciales, a en effet conçu ses solutions comme interopérables afin de constituer un système global de gestion de l'entreprise pouvant se suffire à lui-même. Nous avons donc réalisé une veille sur ces outils, ainsi qu'un certain nombre de tests pour envisager de nouvelles opportunités.

#### **1. Joindre les outils de collecte et les tableaux de bord : *Power Apps***

Nous l'avons dit, l'une des limites de *Power BI* est qu'il ne dispose pas d'un espace de saisie des données. Les datas chargées dans le logiciel ne sont qu'une copie de celles présentes

dans la base source. Ainsi, il n'est pas possible de modifier, ni de saisir de nouvelles données dans la base d'origine à partir de *Power BI*. Or, comme nous l'avons vu plus haut, cette difficulté a entraîné l'échec de certaines tentatives de déploiement de l'outil, notamment le tableau de bord du développement interne d'*Occurrence* en décembre 2021 (cf. III.B.a.5. – *Déploiement en interne*).

Pour pallier ce problème, j'avais déjà proposé une solution sommaire : saisir des données sur une base *Online* ou stockée dans le *Cloud* (type *Excel Online*), puis intégrer un lien hypertexte vers cette base dans le tableau de bord *Power BI*. De même, insérer dans la base de données un lien cliquable renvoyant sur le *Dashboard* une fois de nouvelles valeurs saisies. Mais cette solution ne constitue en réalité qu'un chemin d'accès facilité vers la source. Elle ne permet pas de modifier et d'ajouter de nouvelles données depuis le même outil.

Toujours dans une logique exploratoire, je me suis donc intéressé au logiciel *Power Apps*. Faisant partie de la suite *Microsoft*, il s'agit d'un outil *no-code* de conception d'applications métiers simples. Notamment, les applications conçues sont principalement imaginées pour la gestion de clients, où l'on peut renseigner, modifier et consulter des données sur ceux-ci, à la manière d'un logiciel CRM simplifié. Ainsi, afin d'adapter un tel outil aux études réalisées par *Occurrence*, nous pouvons envisager une application de formulaire, où les champs à remplir représenteraient les réponses à un questionnaire.

Notons que le choix de *Power Apps* n'est pas anodin et présente un réel intérêt pour notre problématique de départ. En somme, ce logiciel possède un module d'intégration de tableau de bord *Power BI*. Par conséquent, si nous parvenons à relier la source où sont stockées les données de notre formulaire avec *Power BI*, nous pouvons obtenir dans un même outil un espace de saisie et un *Dashboard* de présentation et d'analyse des résultats.

En juin 2021, j'ai donc procédé au développement d'une telle application, une fois de plus, en me formant de façon autodidacte au logiciel de création. Après une phase de construction d'un mois à temps partiel, une version finale en 4 onglets navigables (Annexe 4)<sup>60</sup> a été développée :

- Un **onglet d'accueil** : Pour que l'utilisateur choisisse s'il souhaite se rendre sur l'espace de saisie des données ou sur le tableau de bord.

---

<sup>60</sup> Annexe 4 : Application de questionnaire conçue sur *Power Apps*, détail des onglets navigables.

- Un **onglet de formulaire** : Composé de 9 champs à remplir. Un ID de répondant obligatoire et unique, ainsi que 8 questions fictives. Une fois les données saisies et enregistrées, celles-ci sont stockées en *back-end* dans une base *Excel Online*.
- Un **onglet de moteur de recherche** : Dans celui-ci, il est possible de parcourir la base pour rechercher un répondant spécifique à l'aide de son identifiant et de modifier ses informations.
- Un **onglet de tableau de bord** : Il présente le tableau de bord conçu sur *Power BI* et est relié au fichier *Excel Online* où sont inscrites les données saisies. Celui-ci livre une analyse des résultats de l'enquête et conserve l'ensemble de ses fonctionnalités interactives.

*Application Test Power Apps : Onglet de saisie*

En résumé, il s'agit d'une solution prometteuse ayant été validée en interne, pouvant notamment être utile lors d'enquêtes quantitatives avec peu de questions et nécessitant une saisie des données « administrée ». Ce terme désigne les études où le répondant ne saisit pas lui-même ses réponses, lors d'un entretien ou d'une enquête téléphonique par exemple. Cependant, le développement d'une application spécifique dédié à une seule enquête et étant abandonnée par la suite ne convient pas à toutes les missions d'*Occurrence*. Le temps de

création et les contraintes techniques de l'application ne permettent pas d'envisager une telle solution pour des enquêtes quantitatives trop volumineuses notamment.

## 2. Automatiser la collecte et le nettoyage en amont : Power Automate et Power Query

Une autre contrainte de *Power BI* est l'harmonisation et la qualité des données. En effet, l'une des bases de la *data* analyse est de disposer de données propres et nettoyées en amont des traitements, sans quoi, cela peut engendrer des biais d'interprétation. Or, rappelons-le, les missions de conception de tableaux de bord et de suivi de la performance déjà réalisées par *Occurrence* se font sur des données clients. Ces données peuvent ainsi être issues de sources hétérogènes, prendre des formes différentes, ou parfois même être mal saisies par nos interlocuteurs peu sensibilisés à l'importance de la qualité des données.

Pour pallier ces difficultés, plusieurs solutions sont envisageables, tel qu'un nettoyage manuel des bases avant leur import, ou encore une double saisie des valeurs sur *Excel* (solution aujourd'hui privilégiée) dans des bases propres et déjà harmonisées. Mais si cela constitue une tâche fastidieuse pour le client, cette option n'empêche pas les erreurs de saisies humaines.

Une autre possibilité aujourd'hui encore inexploitée chez *Occurrence* est l'utilisation d'outils d'*ETL*. Acronyme de l'anglais « *Extract, Transform, Load* », il s'agit de processus de transformation des données lors de leur acheminement depuis la source, juste avant qu'elles ne soient chargées dans le logiciel de traitement *BI*. Cela permet de collecter des données brutes et d'appliquer une suite de requête sur celles-ci afin de les nettoyer et d'harmoniser les bases issues de sources différentes de façon automatique.

Une fois de plus, la suite *Microsoft* propose deux outils adaptés : *Power Query* et *Power Automate*. Le premier est le logiciel *no-code* d'*ETL* officiel de *Microsoft* et est directement intégré dans *Power BI*. Lors du chargement de données sur *Power BI Desktop*, il est possible, avant de procéder à l'analyse, de « transformer les données ». L'éditeur *Power Query* s'ouvre alors et nous pouvons indiquer l'ensemble des instructions à effectuer sur la base lors du chargement des données. Les traitements peuvent ensuite être effectués sur ces bases « propres » dans *Power BI*. Ainsi, à chaque actualisation du tableau de bord, ces étapes de transformation seront à nouveau reproduites sur les nouvelles données de façon automatique. Attention cependant, cela ne modifie pas les données dans la source d'origine. J'ai eu

notamment l'occasion d'expérimenter cette solution sur les fichiers de standards d'*Occurrence* afin que toutes les bases possèdent le même aspect final.

Le second outil, *Microsoft Power Automate*, n'est pas à proprement parler un logiciel d'*ETL*, mais il peut le devenir. Cette solution, toujours *no-code* comme le privilégie *Microsoft*, permet en effet de créer des flux de travail ou *workflow* automatisés entre différents outils. En somme, nous pouvons compiler une suite de tâches à effectuer à partir d'un élément déclencheur. Il est alors possible de créer un *workflow* d'*ETL* configuré manuellement, en indiquant de d'extraire des données depuis une source, de les transformer, puis de les charger sur *Power BI*. Il s'agit donc d'un outil très complet et particulièrement puissant, que nous envisageons d'explorer à l'avenir afin notamment d'automatiser certaines tâches de traitement fastidieuses et ainsi améliorer la performance du cabinet.

### 3. Assouplir les supports de tableaux de bord : *Sharepoint Online* et *Microsoft Teams*

Enfin, intéressons-nous à deux derniers outils de la suite *Microsoft* permettant d'intégrer les solutions *BI* déployées à un questionnement plus global de fluidification de l'activité. En effet, nous avons pu constater que le partage des tableaux de bord, aussi bien en interne qu'auprès des clients soulève plusieurs problématiques. La question des licences dont il faut disposer pour accéder au *Dashboard* de façon sécurisée, mais également la question des droits d'accès. Aujourd'hui, nous privilégions le partage via un URL web sécurisé, au détriment d'une intégration douce au sein de supports collaboratifs d'entreprise.

Pourtant, nous n'avons eu de cesse de le répéter, les outils *Microsoft* sont conçus pour interagir entre eux, à l'image de *Power BI* avec *Power Apps* ou *Excel Online*. Ainsi, l'intégration des *Dashboards* au sein d'autres supports sont possibles. L'une des plateformes privilégiées en entreprise est *Sharepoint Online*. Celle-ci permet, pour une société utilisant *Microsoft*, de créer des sites web d'équipe. Il s'agit de sites internet à l'accès restreint et stockées sur le *Cloud Microsoft*. Ceux-ci sont utilisés dans le cadre de projets internes pour favoriser l'échange et la collaboration entre plusieurs acteurs à travers plusieurs onglets. Il est notamment possible d'y créer des bases de données dédiés au pilotage d'un projet sous forme de liste, mais également d'incorporer un tableau de bord *Power BI* au sein d'un onglet.



## **a. Limites et axes de vigilances**

Tout d'abord, procédons à certaines mises en garde et soulignons les points d'attention relatif à notre projet de déploiement. Nous avons déjà eu l'occasion d'évoquer les limites techniques de *Power BI* (cf III.B.a.7. – *Limites de Power BI*), mais cette partie est plus largement dédiée aux limites de l'informatique décisionnelle dans le secteur des études. Par ailleurs, nous souhaitons attirer l'attention sur les points sensibles de ce déploiement, les plus susceptibles de constituer des obstacles s'ils sont ignorés à l'avenir :

- **BI de petite envergure** : Nous n'avons pas effectué ce déploiement dans le but de traiter des données massives. Même les bases des clients, souvent plus conséquentes que celles d'Occurrences, ne constituent pas réellement de la *Big data*. De la même manière, nous n'avons pas mis en place de modèles statistiques avancés pour observer les tendances à venir dans la stratégie de pilotage. Enfin, nous agissions au sein d'une infrastructure *data* extrêmement simple, où les sources de données demeurent majoritairement des fichiers *Excel*. Il est donc important de garder en tête qu'*Occurrence* n'est pas un expert de la *Business Intelligence*, malgré ces solutions nouvellement acquises. De fait, le cabinet ne doit pas sur vendre, ni surestimer ses prestations auprès de ses clients, sous peine de créer de la déception.
- **Temps de développement** : Il s'agit d'un point de vigilance qui ne s'applique pas uniquement à *Power BI*. De manière générale, les outils *BI* offrent plus de possibilité que les solutions proposées actuellement. Cependant, le temps de développement de ces solutions est également plus long. L'entreprise monte en qualité de prestation, mais la durée des missions va augmenter également. Donc il faut repenser la façon de travailler : facturation, temps investi, etc.
- **La qualité des données, plus que jamais indispensable** : Nous ne le répéterons jamais assez, les problèmes de qualité des données sont la principale cause d'échec des projets *data*. Un certain nombre d'actions jusqu'alors réalisées manuellement sont automatisées par les outils *BI*. Ainsi, des bases propres en amont sont la clé pour assurer un bon fonctionnement des logiciels et éviter les biais d'analyse. Dans le monde des études de surcroît, la diffusion de résultats biaisés peut rapidement engendrer une perte de crédibilité de l'institut et fortement impacter son image de marque.

- **La BI requiert des compétences data plus importantes** : L'analyse et la consolidation de données sont un véritable métier dont les outils ne sont pas à mettre entre toutes les mains. Malgré le développement et l'attractivité des solutions *no-code*, il est nécessaire de faire preuve d'une grande vigilance lors des opérations de manipulation des données. Une mauvaise jointure, ou une erreur de format peut par exemple complètement fausser les résultats obtenus.
- **Continuer d'assurer la sécurité des données** : Les nouveaux outils *BI* favorisent l'interopérabilité et les échanges transverses. Ainsi, il est facile de publier et de partager les résultats en ligne et à grande échelle. C'est pourquoi les questions de cyber sécurité et le respect du RGPD doivent être centrales afin de ne pas divulguer de données personnelles ou confidentielles par erreur.

### **b. Bonnes pratiques pour assurer une pérennisation de l'utilisation des outils BI**

Pour clore cette étude, il nous semble bon de réfléchir à l'avenir du projet mené sur ces 18 mois de trail. En effet, certains outils de *Business Intelligence* sont désormais implantés et font aujourd'hui partie du quotidien de la vie d'*Occurrence*. Mais l'impulsion donnée aujourd'hui ne doit pas s'essouffler, elle doit s'inscrire dans la durée. Comme toute nouveauté, il faut un temps d'adaptation à l'adoption de l'informatique décisionnelles pour accompagner le pilotage. Si ce sujet est laissé à l'abandon et n'est pas réactivé régulièrement, il est certain que les anciennes pratiques se réinstalleront progressivement.

Ainsi, observons une série de bonnes pratiques à développer afin d'assurer une pérennisation de l'implantation de la *Business Intelligence* chez *Occurrence* :

- **Communiquer sur les outils** : il faut que les collaborateurs aient conscience des capacités techniques des solutions *BI*. Ils doivent être en mesure de les vendre aux clients et d'identifier par eux-mêmes les situations où ces outils peuvent être déployées et apportent une plus-value. En somme, autonomiser les collaborateurs sur ce sujet afin de développer une maîtrise collective du sujet.

- **Instaurer une culture *data* d'entreprise** : Nous en avons déjà parlé, chez *Occurrence*, les collaborateurs ont conscience des enjeux de la *data* mais n'en perçoivent pas les usages concrets pour augmenter les performances du cabinet. Il faut progressivement faire évoluer ces mentalités pour que tout le monde perçoive la conversion en valeur des données. C'est en effet en percevant l'exploitation de ces dernières comme indispensable pour rester concurrentiel que le niveau de maturité *data* de l'entreprise augmentera.
  
- **Former certains collaborateurs** : Ce projet de déploiement de la *BI* doit être collectif. Aujourd'hui, je suis le seul à travailler sur ces outils, mais une transmission de compétences doit être opérée. D'une part, pour que la compétence ne soit pas perdue si je suis amené à quitter l'entreprise, ou même simplement en cas d'absence. D'autre part, pour en généraliser la pratique : de plus en plus d'AO mentionnent explicitement l'utilisation de ces outils, donc le métier se doit d'évoluer en même temps que le marché pour rester concurrentiel.
  
- **Intégrer le principe d'amélioration continue** :
  - **R&D et veille** : Envisager l'avenir en surveillant les tendances du marché, réaliser une veille sur les nouveaux outils de *BI*, continuer de réaliser des tests et d'innover afin de déployer de nouvelles offres et de prolonger la démarche *BI*.
  - **Monter en compétence sur les outils existants** : *Power BI* notamment est une solution extrêmement vaste qui requiert plusieurs années de pratique pour pouvoir prétendre être expert. De nombreuses fonctionnalités restent à explorer, tel que l'intégration de code *R* et *Python* dans l'outil, ou encore la gestion avancée des droits d'accès par exemple.
  
- **Anticiper et accompagner un changement d'architecture du Système d'Information** : Comme dit plus tôt, le SI d'*Occurrence* est ancien et son architecture ne correspond plus aux standards actuels des entreprises. Pour évoluer dans une démarche *data* globale, le cabinet sera inévitablement amené à réviser en profondeur son SI, pour disposer, à minima, d'un *Datawarehouse* et de bases de données relationnelles stockées sur un serveur *SQL* ou équivalent. Il serait préférable que les outils *BI* déployés aujourd'hui soient compatibles avec le fonctionnement d'*Occurrence* de demain.

Ainsi, nous espérons que la mise en application de ces bonnes pratiques permettra définitivement d'implanter la *Business Intelligence* au sein d'*Occurrence* et de multiplier ses performances à l'avenir. N'oublions pas que cela ne peut être qu'une démarche collective et que les données sont l'affaire de tous, en particulier dans une PME comme celle-ci. C'est pourquoi la communication sur ce sujet, le partage des savoirs et le transfert de compétences sont la clé d'un succès à long terme.

## Conclusion

Pour conclure cette étude, nous avons véritablement pu constater que le déploiement d'outils de *Business Intelligence* constitue une aide précieuse, sinon indispensable, pour faire face à la forte concurrence du secteur des études et de la communication.

En effet, nous avons tout d'abord mené des recherches théoriques sur le sujet et celles-ci nous ont permis de saisir l'impact considérable de la *BI* sur le pilotage des entreprises, ainsi que ses potentielles voies d'utilisation concrète au sein des instituts de sondage.

Puis, dans un second temps, nous avons conduit un projet de déploiement d'outils d'informatique décisionnelle au sein de la société *Occurrence*, s'étendant sur plus de 18 mois. Partant d'une expérience de terrain quasi-nulle dans ce secteur, cette étude m'a ainsi permis de monter en compétences sur de nombreux outils, à commencer par *Microsoft Power BI*. Lors des différentes étapes de test itératifs, j'ai ainsi pris progressivement conscience de la puissance de la *BI* et de la plus-value qu'elle pourrait générer à moyen et long terme. Parallèlement, j'ai pu constater les contraintes et difficultés de déploiement propre à ce secteur, notamment à destination des entreprises clientes. La *Business Intelligence* est originellement prévue pour une démarche interne, raccordée au service d'information de l'entreprise détentrice de l'outil. Paradoxalement, les appels d'offres du secteur exprimant la volonté de disposer d'un tableau de bord sur *Power BI* se sont multipliés au fil des mois. Ce phénomène est donc venu confirmer les enjeux de déploiement de cet outil afin de rester concurrentiel dans ce secteur d'activité.

Enfin, l'exploration et l'intégration d'autres solutions *BI* spécialisées ont permis d'élargir encore les possibilités d'exploitations et les offres proposées par *Occurrence*, voire de se positionner en anticipation de missions futures. Apportons une mention toute particulière aux projets cartographiques réalisés sur QGIS, ayant directement été intégrées aux services proposés par le cabinet.

Cependant, si cette approche exploratoire du projet a été un succès, notre étude présente tout de même certaines limites que nous ne pouvons ignorer. D'une part, rappelons qu'aucun projet auprès d'entreprises clientes n'a encore abouti à un déploiement concret. Nous avons fait face à plusieurs obstacles, en partie liés à notre méconnaissance du logiciel, mais également à l'impossibilité techniques de certaines demandes. Ainsi, des incertitudes persistent et ne pourront être déterminées qu'après une mise en place opérationnelle : la tarification de la

prestation, les processus de déploiement à l'externe, le suivi et la maintenance des outils sur la durée...

D'autre part, sur les cinq objectifs fixés à l'origine, nous n'en avons atteint que trois. Les deux objectifs restants concernent les enjeux communicationnels et collaboratifs associés au projet, à savoir, le développement d'une culture *data* d'entreprise et la formation d'autres collaborateurs aux outils. Cela pose un constat : ce projet a été mené à bien et est arrivé à son terme, mais je ne suis pas parvenu à impliquer les équipes dans celui-ci. A titre d'exemple, certains collaborateurs du pôle ne parvenaient pas à identifier quelles étaient mes missions de R&D, qui sont toujours restées relativement opaques. Ce problème communicationnel doit constituer un axe de vigilance fort. En effet, nous avons souligné le fait que la *BI* était un enjeu collectif au sein de l'entreprise. De ce fait, une non-implication des collaborateurs dans le projet, même involontaire, constitue une menace pour la pérennité des outils.

C'est pour cette raison, et en gardant en tête les limites et points de vigilances de ce projet, que nous avons clos notre étude sur la présentation des bonnes pratiques à adopter afin d'inscrire ce déploiement dans le temps. Le sujet de la *Business Intelligence* est un sujet vivant, qui ne doit pas s'achever au lendemain de la publication de ce travail de recherche. *Occurrence* n'est en réalité qu'aux balbutiements de l'exploitation de la *Business Intelligence* et dispose encore de nombreuses perspectives d'évolutions, dont le tournant majeur sera la réorganisation de son SI à l'avenir.

Pour clore cette étude, dressons un bilan positif et retournons observer l'échelle de maturité mises en place par l'*Observatoire de la Maturité Data des Entreprises*. Nous avons originellement estimé le niveau de maturité d'*Occurrence* entre 1 et 2 sur 6. Aujourd'hui, l'implantation d'outils d'informatique décisionnelle au sein du cabinet réhausse cette note qui tend aujourd'hui vers 3. Pour s'assurer ce score, il ne reste plus qu'à déployer les outils disponibles dans le cadre du pilotage interne de l'entreprise.

# Bibliographie

## Ouvrages :

ARRUABARRENA Béa, « Datavisualisation : principes, enjeux et perspectives pour des utilisateurs non experts », dans BROUDOUX Évelyne, *Big Data - Open Data : Quelles valeurs ? Quels enjeux ? Actes du colloque « Document numérique et société »*, Rabat, 2015, Louvain-la-Neuve, De Boeck Supérieur, « Information et stratégie », 2015, pp. 151-163 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/--9782807300316-page-151.htm>

CARLIER Alphonse, *Business Intelligence et management*, Paris, Afnor, 2013, 277 p.

DARGENT Claude, *Sociologie des opinions*, Paris, Armand Colin, 2011, 240 p.

DESABIE Jacques, *Théorie et pratique des sondages*, Paris, Dunod, 1966, 482 p.

FERNANDEZ Alain, *Les Nouveaux Tableaux de bord des managers, Le Projet Business Intelligence clés en main*, Paris, Eyrolles, 2013, 662 p.

GARNIER Alain, *L'Information non structurée dans l'entreprise - Usages et Outils*, Paris, Eyrolles, 2007, 245 p.

GOUARNE Jean-Marie, *Le Projet décisionnel : Enjeux, Modèles, Architectures du Data Warehouse*, Paris, Eyrolles, 1997.

JOLY Bruno, *La communication*, « Chapitre 3 : La communication en entreprise », Paris, De Boeck Supérieur, « Le point sur... Marketing », 2009, pp. 69-129.

LE ROUX Brigitte et LEBARON Frédéric (dir.), *La méthodologie de Pierre Bourdieu en action : espace culturel, espace social et analyse des données*, Malakoff, Dunod, 2015, 262 p.

LÉVY Mick, *Sortez vos données du frigo. Une entreprise performante avec la Data et l'iA*. Paris, Dunod, 2021, 302 p.

LIEBOWITZ Jay, *Strategic intelligence : Business Intelligence, competitive intelligence, and knowledge management*, Boca Raton, Auerbach Publications, 2006, 242 p.

MAURIN Pierre, *Le tableau de bord*, Paris, Ellipses, 2021, 136 p.

MILLIE-TIMBAL Marie-Hélène, *Construire des tableaux de bord vraiment utiles : associer les indicateurs à la stratégie de l'entreprise et à son management*, Le Mans, Gereso, 2020, 190 p.

OLIVIER Martin, *L'analyse quantitative des données*, Malakoff, Armand Colin, 2020, 128 p.

REGNIER Nicole, *Une expérience de diffusion "intelligente" d'informations de tableau de bord qui généralise le concept de portail, Réseaux sociaux et d'information*, Mémoire, CNAM, 2012, 125 p.

RISSOAN Romain, « Outil 11. Big data vs Business Intelligence au service de la marque employeur », dans *La boîte à outils de la stratégie big data*, Paris, Dunod, 2018, pp. 44-45.

SELMER Caroline, *Concevoir le tableau de bord. Méthodologie, outils et exemples visuels*, Paris, Dunod, « Fonctions de l'entreprise », 2015, 232 p.

TAYLOR James, *Decision Management System : a practical guide to using Business rules and predictive analytics*, Indianapolis, IBM Press, 2011, 308 p.

### **Articles :**

ANDONOVA Yanita. « De l'invisibilité des dispositifs numériques à la légitimation de la communication en entreprise », dans *Sociologies pratiques*. 2015. N°30, p. 43-52.

BOURDIEU, Pierre. « L'opinion publique n'existe pas », dans *Les Temps modernes*. [en ligne]. URL : [https://www.acrimed.org/IMG/article\\_PDF/article\\_a3938.pdf](https://www.acrimed.org/IMG/article_PDF/article_a3938.pdf)

BOYADJIAN Julien, « Big data et traçabilité numérique. Les sciences sociales face à la quantification massive des individus », dans *Politix*, n°122, Paris, Collège de France, 2017, pp. 234-237 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-politix-2018-2-page-234.htm>

BOYADJIAN Julien et VELCIN Julien, « L'analyse quantitative des médias sociaux, une alternative aux enquêtes déclaratives ? La mesure de la popularité des personnalités politiques sur Twitter », *Questions de communication*, n°31, 2017, pp. 111-135 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-questions-de-communication-2017-1-page-111.htm>

MERCANTI-GUÉRIN Maria, « L'amélioration du reciblage par les Big Data : une aide à la décision qui menace l'image des marques ? », dans *Revue internationale d'intelligence économique*, vol. 5, 2013, pp. 153-165 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-internationale-d-intelligence-economique-1-2013-2-page-153.htm>

REUTTER Jérôme, ALLAIN Élodie, LANDAGARAY Patrice, « L'évolution des rôles du contrôleur de gestion à l'ère de la *Business Intelligence* », dans *ACCRA*, n°11, 2021, pp. 85-107 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-accra-2021-2-page-85.htm>

RIKHARDSSON Pall et YIGITBASIOGLU Ogan, « *Business Intelligence & analytics in management accounting research: Status and future focus* », dans *International Journal of Accounting Information Systems*, vol. 29, 2018, pp. 37-58 [en ligne]. URL : [https://econpapers.repec.org/article/eeeijoais/v\\_3a29\\_3ay\\_3a2018\\_3ai\\_3ac\\_3ap\\_3a37-58.htm](https://econpapers.repec.org/article/eeeijoais/v_3a29_3ay_3a2018_3ai_3ac_3ap_3a37-58.htm)

SABER Brahim, « Open source comme système d'informatique décisionnelle », dans *Revue internationale d'intelligence économique*, vol.3, 2011, pp. 93-101 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-internationale-d-intelligence-economique-1-2011-1-page-93.htm>

STANOEVSKA-SLABEVA Katarina, WOZNIAK Thomas et RISTOL Santi, *Grid and Cloud Computing: A Business Perspective on Technology and Applications*. Berlin, Springer, 2010, 224 p [en ligne]. URL : <https://link.springer.com/book/10.1007/978-3-642-05193-7>

SOUFFI Emmanuelle, « Enquête sur « Les Invisibles », dans *le Journal du dimanche*, 13 mars 2022 [en ligne]. URL : <https://www.lejdd.fr/Politique/enquete-sur-la-france-des-oublies-4099073>

SYBORD Christine, « Intelligence économique et système d'aide à la décision : de l'opérationnel « big data » au stratégique « little knowledge » », dans *Revue internationale d'intelligence économique*, vol.7, 2015, pp. 83-100 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-internationale-d-intelligence-economique-1-2015-1-page-83.htm>

TRAVAILLÉ Denis, MARSAL Christine, « Automatisation des tableaux de bord et cohérence du contrôle de gestion : à propos de deux cas », dans *Comptabilité Contrôle Audit*, tome 13, 2007, pp. 75-96 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-comptabilite-contrôle-audit-2007-2-page-75.htm>

« Les sondages d'opinion », dans *Courrier hebdomadaire du CRISP*, n°829, 1979, pp. 1-24 [en ligne]. URL : <https://www-cairn-info.proxybib-pp.cnam.fr/revue-courrier-hebdomadaire-du-crisp-1979-4-page-1.htm>

### **Sites internet :**

AURINE Guillaume, « Une brève histoire de la *Business Intelligence* », 7 juin 2016. URL : <https://www.salesforce.com/fr/blog/2016/06/une-breve-histoire-de-la-business-intelligence.html>

BELHAMRI Nicolas, « OUTILS DE *BUSINESS INTELLIGENCE* : LE TOP 10 EN 2022 (COMPARATIF) », *Boryl*, URL : <https://www.boryl.fr/blog/top-10-des-outils-de-business-intelligence/#8>

Boulevard du web, « Cartographie des plateformes de visioconférence les plus populaires à travers le monde ». URL : <https://boulevardduweb.com/plateformes-visio-2021/#:~:text=Les%20parts%20de%20march%C3%A9%20des%20plateformes%20de%202020%20%C3%A0%202021&text=Zoom%20a%20gagn%C3%A9%2022%2C3,8%25%20de%202020%20%C3%A0%202021.>

DALLE-MOLLE Maris, « Enquête sur « Les Invisibles » pour la Fondation Travailler Autrement », 13 mars 2022. URL : <https://www.fondation-travailler-autrement.org/2022/03/13/invisibles-plongee-france-back-office-etude/>

Digital League Rhône-Alpes, « La maturité *Data*, c'est quoi ? », 26 mai 2021. URL : <https://www.digital-league.org/blog/2021/05/26/la-maturite-data-cest-quoi/#:~:text=La%20maturit%C3%A9%20data%20d'une,m%C3%A0me%20en%20fonction%20du%20secteur>

GARTNER, « « *Dirty data* » is a business problem, not an IT problem, says Gartner », 2007. URL : <https://www.gartner.com/newsroom/id/501733>

IMT Atlantique – *Business Intelligence*, « Pilotage des processus d'enquêtes ». URL : [http://formations.imt-atlantique.fr/BI/BI\\_projet\\_sondage.html](http://formations.imt-atlantique.fr/BI/BI_projet_sondage.html)

Observatoire de la Maturité *Data* des Entreprises. URL : <https://observatoire-data.fr/>

Observatoire des pétitions (par *Occurrence*) – Comprendre l’opinion via les pétitions. URL : <https://observatoirepétitions.fr/>

Occurrence, cabinet d’études et conseil en communication. URL : <http://Occurrence.fr/>

Opinionpublique, « Les instituts de sondage français produisant des études d’opinion ». URL : <https://opinionpublique.wordpress.com/les-instituts-de-sondage-francais-produisant-des-etudes-dopinion/>

Page Wikipédia dédiée à la *Business Intelligence*. URL : [https://fr.wikipedia.org/wiki/Informatique\\_d%C3%A9cisionnelle](https://fr.wikipedia.org/wiki/Informatique_d%C3%A9cisionnelle)

R. Gaetan, « *Data Governance* ou gouvernance des données : Qu’est-ce que c’est ? », *Lebigata.fr*. URL : <https://www.lebigdata.fr/data-governance>

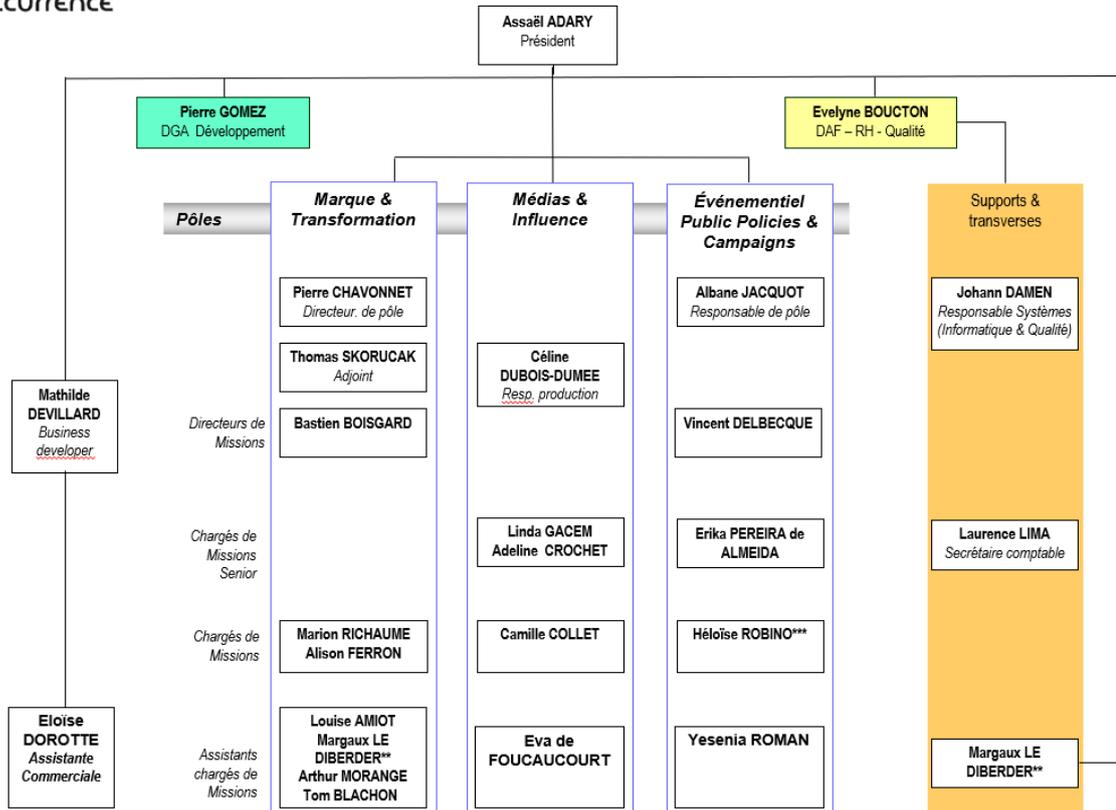
Talend - Un leader de l’intégration de données *Cloud* (*ETL* moderne). URL : <https://www.talend.com/fr/>

Tableau, « Culture des données ». URL : <https://www.tableau.com/fr-fr/why-tableau/data-culture>

TREHOLAN Magali, « Les *Big Data* prémunissent-ils des biais dans les études ? », *Survey Magazine*. URL : <https://www.soft-concept.com/surveymag/big-data-biais-etudes.html>

# Annexes

## Annexe 1 - Organigramme de la société Occurrence :

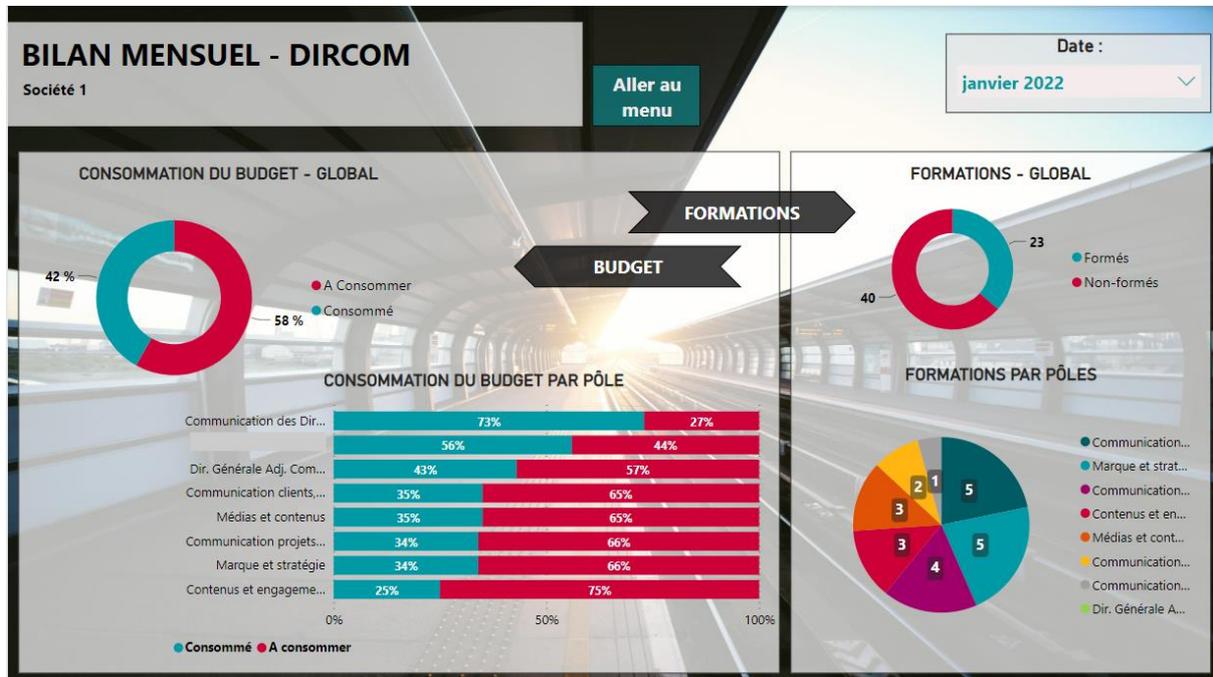


\*\* Ressource transversale rattachée à AA

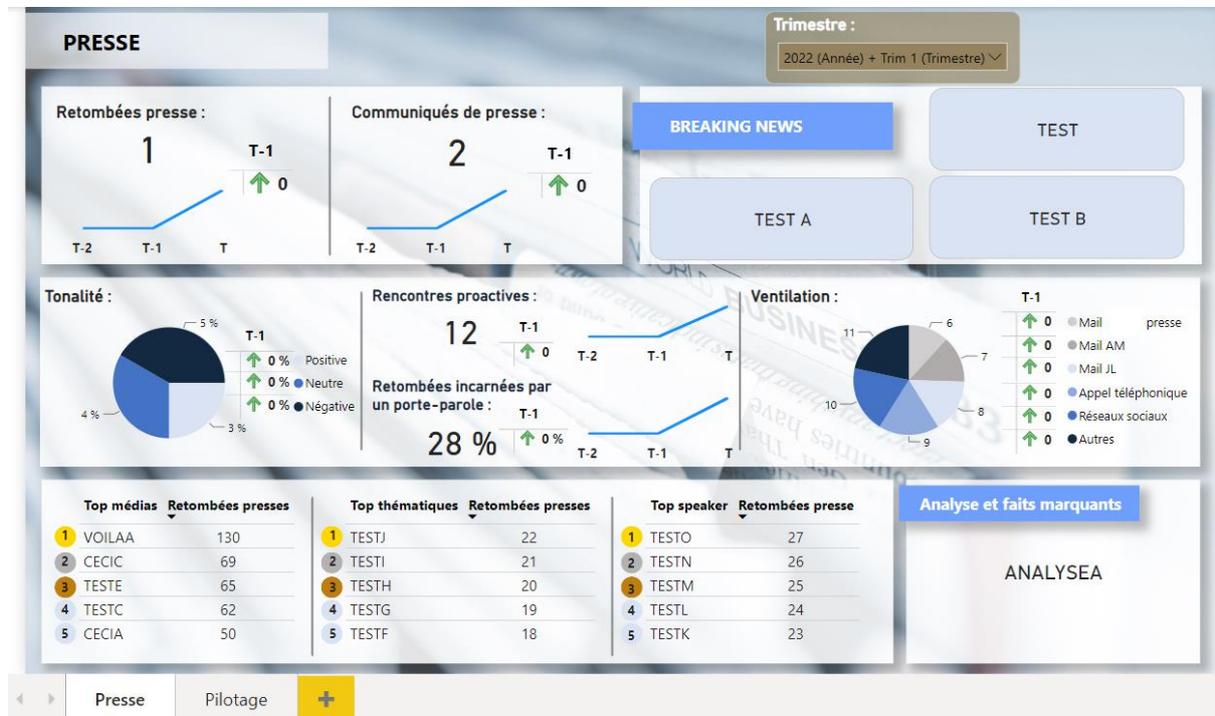
\*\*\* VIE

Organigramme OCCURRENCE – 14/01/21

**Annexe 2 – Maquette de tableau de bord de pilotage réalisée sur Power BI pour la « Société 1 » :**



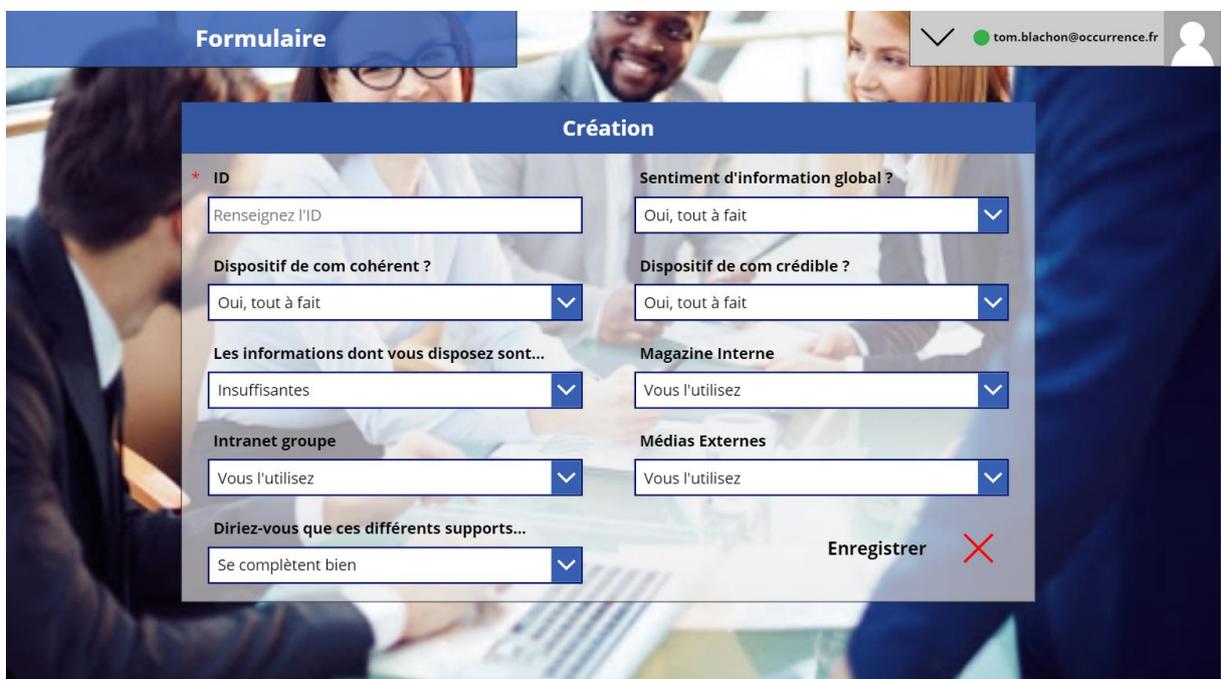
**Annexe 3 – Maquette de tableau de bord de pilotage réalisée sur Power BI pour la « Société 2 » :**



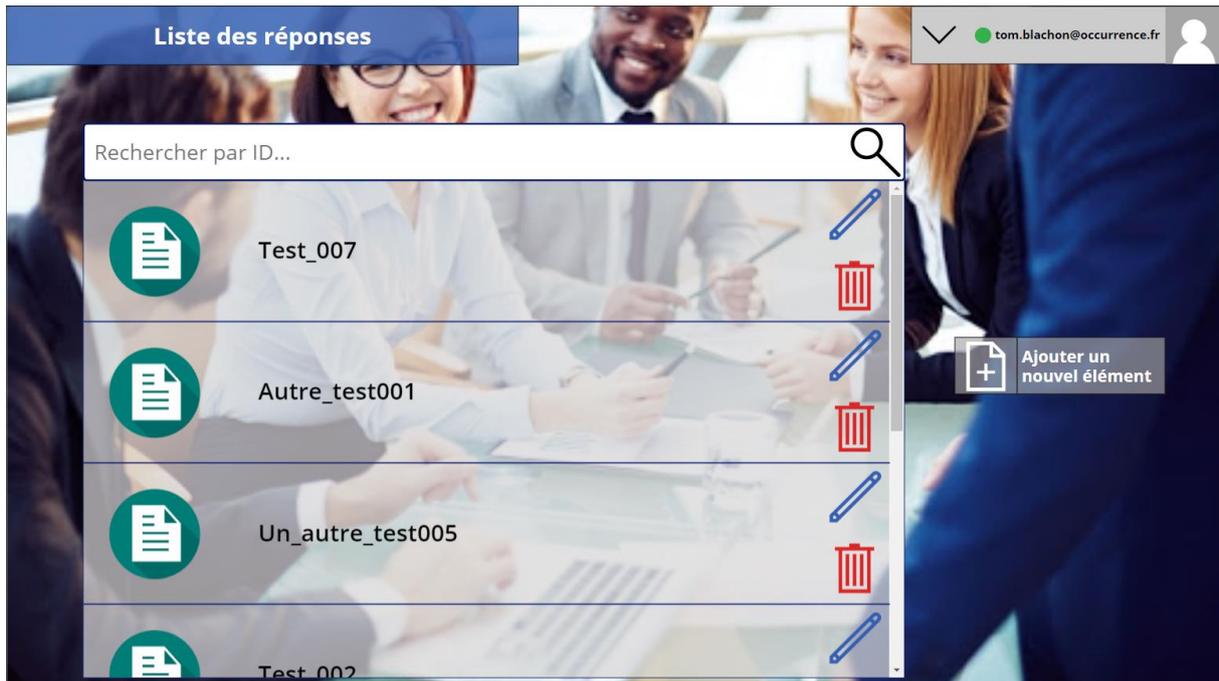
**Annexe 4 – Application de questionnaire conçue sur Power Apps, détail des onglets navigables :**



*Application Test Power Apps : Page d'accueil*



*Application Test Power Apps : Onglet de saisie*



Application Test Power Apps : Onglet de consultation de la base



Application Test Power Apps : Tableau de bord Power BI (données fictives)