

Haute Ecole  
Groupe ICHEC – ECAM – ISFSC



Enseignement supérieur de type long de niveau universitaire

**Etablissement d'un service de veille législative  
européenne des systèmes d'information (EU  
Monitoring) au sein de l'agence Consulting de Sopra  
Steria Benelux.**



Mémoire présenté par :

**Charles de BRIEY**

Pour l'obtention du diplôme de :

**Master - Ingénieur commercial**

Année académique 2019-2020

Promoteur :

**Monsieur Donatien GROLAUX**

Boulevard Brand Whitlock 6 - 1150 Bruxelles



Haute Ecole  
Groupe ICHEC – ECAM – ISFSC



Enseignement supérieur de type long de niveau universitaire

**Etablissement d'un service de veille législative  
européenne des systèmes d'information (EU  
Monitoring) au sein de l'agence Consulting de Sopra  
Steria Benelux.**



Mémoire présenté par :

**Charles de BRIEY**

Pour l'obtention du diplôme de :

**Master - Ingénieur commercial**

Année académique 2019-2020

Promoteur :

**Monsieur Donatien GROLAUX**

Boulevard Brand Whitlock 6 - 1150 Bruxelles



## Remerciements

---

Je tiens à adresser mes plus sincères remerciements à toute une série de personnes.

Tout d'abord, je remercie grandement mon maître de stage, monsieur Cédric Genin ainsi que les collaborateurs de Sopra Steria avec qui j'ai eu la chance de travailler au quotidien et qui m'ont accompagné dans la gestion de projet. Je pense notamment à Julien Heintz, Florian Kariger, Gian Michele Mosca et toutes les autres personnes avec qui j'ai eu la chance d'échanger lors de ce qui a été ma première expérience professionnelle. Tous ont fait preuve de bienveillance et de pédagogie à mon égard.

Je voudrais aussi exprimer ma reconnaissance à mon promoteur, monsieur Donatien Grolaux, qui a relu à plusieurs reprises les versions successives de ce mémoire et dont les remarques étaient toujours pertinentes.

Finalement, merci aux membres de ma famille qui m'ont soutenu et accompagné au cours de cette période si particulière de confinement.



« Je soussigné, de BRIEY, Charles, Master 2, déclare par la présente que le Mémoire ci-joint est exempt de tout plagiat et respecte en tous points le règlement des études en matière d'emprunts, de citations et d'exploitation de sources diverses, signé lors de mon inscription à l'ICHEC, ainsi que les instructions et consignes concernant le référencement dans le texte respectant la norme APA, la bibliographie respectant la norme APA, etc. mises à ma disposition sur Moodle.

Sur l'honneur, je certifie avoir pris connaissance des documents précités et je confirme que le Mémoire présenté est original et exempt de tout emprunt à un tiers non-cité correctement.  
»

Date

Bruxelles, le 17 aout 2020

Signature

*Charles de Briey*

« Nous sommes bombardés d'informations que nous sommes censés digérer plus rapidement que nous ne le pouvons. » Atom Egoyan

« La connaissance s'acquiert par l'expérience, tout le reste n'est que de l'information. »  
Albert Einstein



# Table des Matières

---

<b>Listes des figures et tableaux</b>	<b>11</b>
Liste des figures	11
Liste des tableaux	12
<b>Introduction</b>	<b>13</b>
<b>Chapitre 1 : Contextualisation du Projet</b>	<b>15</b>
<b>1. Environnement</b>	<b>15</b>
1.1. L'Entreprise	15
1.1. Branche Sopra Steria Benelux	20
1.2. L'Agence Consulting	22
<b>2. Approche théorique de la veille législative européenne</b>	<b>23</b>
2.1. Veille	23
2.2. Spécificités de la veille législative	24
2.3. Critères d'efficacité	25
2.4. Institutions et Agences de l'Union Européenne	28
2.4. Fiabilité des sources	33
2.5. Synthèse de la veille	34
2.6. Technologies potentiellement utiles	37
<b>Chapitre 2 : Description du Projet</b>	<b>40</b>
<b>1. Constats de départ</b>	<b>40</b>
1.1. Retard dans la prise de connaissance	40
1.2. Besoin de reconnaissance sur le marché ESN	40
1.3. Mise à profit des compétences existantes	40
<b>2. Objectif principal</b>	<b>41</b>
<b>3. Nature de l'existant</b>	<b>42</b>
<b>4. Acteurs du projet</b>	<b>42</b>
<b>5. Analyse ADEPT</b>	<b>43</b>
5.1. Activités	43
5.2. Données	44
5.3. Environnement	46
5.4. Personnes	49
5.5. Technologie	50
<b>6. Objectifs intermédiaires</b>	<b>52</b>
<b>7. Méthodologie de gestion de projet</b>	<b>54</b>
<b>Chapitre 3 : Mise en Œuvre du Projet</b>	<b>57</b>
<b>1. Récolte d'information</b>	<b>57</b>
1.1. Boîte e-mail dédiée	57
1.2. Agrégateur de flux RSS	57
1.3. Pipe de lectures	59

<b>2. Évaluation de la pertinence</b>	<b>60</b>
<b>3. Résumer l'information</b>	<b>61</b>
3.1. Comment une application de résumés fonctionne-t-elle ?	63
<b>4. Installation de la plateforme</b>	<b>66</b>
4.1. Choix de la solution Wiki	66
4.2. Utilisation de <i>WampServer</i>	68
4.3. Accès à la machine hôte à distance	69
<b>5. Définition du champ d'application</b>	<b>70</b>
<b>6. Stockage de l'information</b>	<b>70</b>
6.1. Contenu permanent vs Actualité	70
6.2. Gestion des utilisateurs	72
6.3. Gestion de la base de données	73
<b>7. Partage de l'information</b>	<b>73</b>
7.1. <i>Newsletter</i>	73
<b>8. Harmonisation du <i>Business Process</i></b>	<b>74</b>
<b>9. Identification des besoins à court terme</b>	<b>76</b>
9.1. Parvenir à partager l'information	76
9.2. Migration du système	77
9.3. Calibration des moyens humains	77
9.4. Claire définition de la portée d' <i>EU Monitoring</i>	78
<b>Chapitre 4 : Bilan</b>	<b>79</b>
<b>1. Accomplissement des objectifs</b>	<b>79</b>
<b>2. Appréciations Générales</b>	<b>81</b>
2.1. Fonctionnalités	81
2.2. Efficacité	83
<b>Chapitre 5 : Potentielles évolutions</b>	<b>86</b>
<b>1. <i>Web Mining</i></b>	<b>86</b>
1.1. <i>Web Content Mining</i>	87
<b>2. Suivi de sujets</b>	<b>88</b>
<b>3. Analyse de sentiment/objectivité</b>	<b>91</b>
<b>4. Voice-to-Text</b>	<b>93</b>
<b>Conclusion</b>	<b>95</b>
<b>Bibliographie</b>	<b>97</b>

# Listes des figures et tableaux

---

## Liste des figures

Figure 1 : Géographie de Sopra Steria et CA relatifs .....	16
Figure 2 : Branches principales de Sopra Steria et leurs nombres de collaborateurs .....	16
Figure 3 : Poids relatifs des secteurs d'activité de Sopra Steria .....	17
Figure 4 : Poids relatifs des services de Sopra Steria .....	18
Figure 5 : Comparaison CA des "Big 4s" et de Sopra Steria .....	19
Figure 6 : Comparaison CA ESN et de Sopra Steria .....	20
Figure 7 : Géographie de Sopra Steria Benelux .....	21
Figure 8 : Axes verticaux et centres de compétences transversaux .....	22
Figure 9 : Les axes de la BSC .....	27
Figure 10 : Temps du trilogue .....	30
Figure 11 : Schéma de l'interopérabilité des systèmes d'information de l'espace de liberté, de sécurité et de justice (eu-LISA) .....	32
Figure 12 : Entraînement de modèles de <i>Machine Learning</i> .....	36
Figure 13 : Aperçu du fonctionnement de <i>Text Mining</i> basé sur l'extraction d'information .....	38
Figure 14 : Business process d'EU Monitoring .....	44
Figure 15 : Data Model d'EU Monitoring .....	46
Figure 16 : Use case d'EU Monitoring .....	50
Figure 17 : Architecture d'EU Monitoring .....	52
Figure 18 : Roue de Demming .....	56
Figure 19 : Rubrique "Today" de Feedly .....	58
Figure 20 : Rubrique "All" de Feedly .....	59
Figure 21 : Pipe de lecture .....	60
Figure 22 : Fonction "Automatique" de <i>Resoomer</i> .....	62
Figure 23 : Fonction "Manuel" de <i>Resoomer</i> .....	63
Figure 24 : Processus de résumé de texte .....	65
Figure 25 : Architecture d'EU Monitoring mise à jour .....	69
Figure 26 : Partie Wiki d'EU Monitoring .....	71
Figure 27 : Partie Articles d'EU Monitoring .....	71
Figure 28 : Privilèges maximum et minimum d'une page .....	73
Figure 29 : Business Process mis à jour avec les outils utilisés .....	75
Figure 30 : Différents types de Web Mining .....	86
Figure 31 : Architecture du système d'extraction des mots-clés .....	89
Figure 32 : Module d'extraction des mots-clés .....	91
Figure 33 : Analyse de sentiment d'un texte .....	93

## Liste des tableaux

Tableau 1 : Rôles et descriptions des acteurs du projet .....	42
Tableau 2 : Caractéristiques des veilles législatives équivalentes .....	48
Tableau 3 : Listes objectifs intermédiaires avec leurs mesures de réalisation.....	54
Tableau 4 : Degré d'importance des sujets à traiter.....	61
Tableau 5 : Sélection du Wiki.....	67
Tableau 6 : Réalisation du projet.....	81
Tableau 7 : Balanced Scorecard d' <i>EU Monitoring</i> .....	84

## Introduction

---

Chaque organisation doit faire face, chaque jour, à une quantité astronomique d'informations. Pour la plupart, il est impossible de faire fi de cette information qui peut s'avérer précieuse. Celle-ci prend différentes formes (vidéo, texte, image, etc), provient de différentes sources (informatives ou législatives) et peut surtout avoir différents niveaux de qualité.

Le problème étant que lorsque les emplois du temps sont déjà surchargés, peu d'organisations ou d'entreprises prennent le temps de considérer suffisamment ce flux incessant. Il faut dès lors les aider à faire le tri, à uniformiser et à simplifier l'information.

Par ailleurs, nul ne peut échapper au fait que les affaires et la gestion quotidienne de n'importe quel secteur d'activités est régi par des législations, tous niveaux de pouvoir confondus. Le RGPD (Règlement Général sur la Protection des Données) en est sûrement l'exemple le plus célèbre en matière informatique.

De nombreuses organisations perçoivent très bien ces défis et font donc appel à des entreprises pour les accompagner à gérer ce flux d'informations et ainsi à prendre de meilleures décisions. L'agence Consulting de Sopra Steria fait partie de celles à qui ces organisations peuvent s'adresser pour les aider. Toutefois, Sopra Steria n'échappe pas aux problématiques mentionnées ci-dessus et son temps n'est pas plus extensible que les autres. Afin qu'elle puisse, malgré tout, gérer efficacement cette information et assurer son rôle de support, l'agence Consulting de Sopra Steria Benelux vise à développer un système de veille législative.

Ce système de veille législative s'appelle « *EU Monitoring* ». Sa portée est bien précise puisqu'elle concerne les systèmes d'information de gestion des frontières et de sécurité intérieure au niveau européen. Le contexte spécifique de l'Union Européenne fait que toute régulation prise à ce niveau de pouvoir a des conséquences sur tous les Etats Membres, lesquels doivent s'y conformer.

Le rôle d'*EU Monitoring* pour les clients de Sopra Steria est double. D'une part, il doit permettre d'anticiper la mise en place de législations au niveau européen. D'autre part, il doit aider à décrypter ces législations afin d'informer les états membres et les organisations actives dans la gestion des frontières et de la sécurité intérieure.

De plus, *EU Monitoring* sera aussi un outil interne à l'entreprise. Il permettra d'identifier des opportunités dans les textes législatifs européens et les articles de presse spécialisés. Une opportunité serait par exemple l'entrée en vigueur d'un nouveau système d'information pour lequel Sopra Steria pourrait décider d'être impliquée.

Dans les deux cas d'utilisation, cela deviendrait alors un outil d'aide à la décision.

Ce document aura dès lors comme question de fond : « Comment concevoir un système de veille législative européenne efficace ? »

Afin de répondre à cette question, nous analyserons le déroulement du projet de développement d'*EU Monitoring* au sein de l'agence Consulting de Sopra Steria Benelux.

Tout d'abord, nous donnerons un contexte au projet. Cette contextualisation comprendra deux points principaux : La description de l'entreprise et une approche théorique de la veille législative. Cette approche théorique servira à apporter des éléments de réponse à notre questionnement.

Ensuite, nous décrirons le projet. C'est dans cette section que seront exposés l'objectif et les exigences du projet. Ce chapitre servira à aborder notre question de manière plus pratique avec notamment l'analyse ADEPT, qui y sera définie.

Après cela, il conviendra de décrire la mise en œuvre du projet. Tous les outils et procédés mis en place dans le cadre du projet seront expliqués. Cela sera mis en perspective dans le chapitre qui suivra.

En effet, à la suite de la mise en œuvre, nous tirerons un bilan de ce qui a été fait. Cela permettra de dégager des pistes d'amélioration afin de rendre *EU Monitoring* encore plus efficace.

Finalement, les pistes d'amélioration émises lors du bilan seront explorées de manière théorique lors du dernier chapitre.

# Chapitre 1 : Contextualisation du Projet

---

## 1. Environnement

### 1.1. L'Entreprise

#### 1.1.1. Historique

Avant de devenir une seule entité, SOPRA (SOciété de PROgrammation et d'Analyse), créée en 1968, et STERIA (Société d'Etude et de Réalisation en Informatique et Automatismes), créée en 1969, étaient deux entreprises françaises séparées. Elles avaient comme but commun d'aider leurs clients à se moderniser en leur proposant des produits et services digitaux innovants.

Sopra se concentre sur le développement de logiciels. Elle développa donc une plateforme bancaire après son premier grand contrat d'infogérance bancaire en 1971 et Pacha fut sa première solution développée pour les ressources humaines (1972).

Steria de son côté, gagne de plus en plus de contrats dans le secteur public comme notamment en informatisant l'AFP (Agence France-Presse) en 1973 ou encore en prenant part au développement du Minitel (1981). Steria devient ensuite également expert sur le marché des transports, à la suite de l'automatisation de la ligne A du RER en Île-de France (1987).

Sopra et Steria font leurs entrées en bourse respectivement en 1990 et 1999. Le début des années 2000 est témoin d'une demande croissante pour des acteurs globaux capables d'accompagner les entreprises dans leur transformation digitale. Steria répond à ces défis en acquérant plusieurs entreprises. Sopra pour sa part, consolide son expansion en Europe en combinant ses croissances internes et externes. Pour cela, elle développe des axes majeurs que sont le conseil, l'intégration de systèmes et l'édition de solutions digitales (Sopra Steria, 2020).

En 2014 a lieu la fusion des deux entités, ce qui donne naissance à un leader européen qui se nomme désormais Sopra Steria Group.

Aujourd'hui, le groupe fait partie du top 5 des acteurs de la transformation digitale en Europe. Un peu plus de 46.000 personnes, réparties dans 25 pays, travaillent pour l'entreprise qui a généré un chiffre d'affaire de 4,4 milliards d'Euros en 2019 (Sopra Steria, 2020).

#### 1.1.2. Géographie

D'un point de vue géographique, Sopra Steria est avant tout une entreprise à portée européenne étant donné que le reste du monde ne représente que 2 % de son activité, majoritairement en Inde (Figure 1). La France est bien sûr son premier marché, représentant à lui seul 50 %, devant le Royaume-Uni avec 19 %. Le reste de l'Europe représente 20 % (Sopra Steria, 2020).



Figure 1 : Géographie de Sopra Steria et CA relatifs

Source : Sopra Steria. (2020, juin 1). *Sopra Steria Benelux* [Diapositives]. Teams

Les branches comptant les plus grands contingents de collaborateurs sont la France, le Royaume-Uni et l'Inde employant respectivement 19.502, 6.305 et 5.726 personnes. L'Espagne est la troisième plus grande branche en Europe en termes de nombres d'employés (Figure 2).

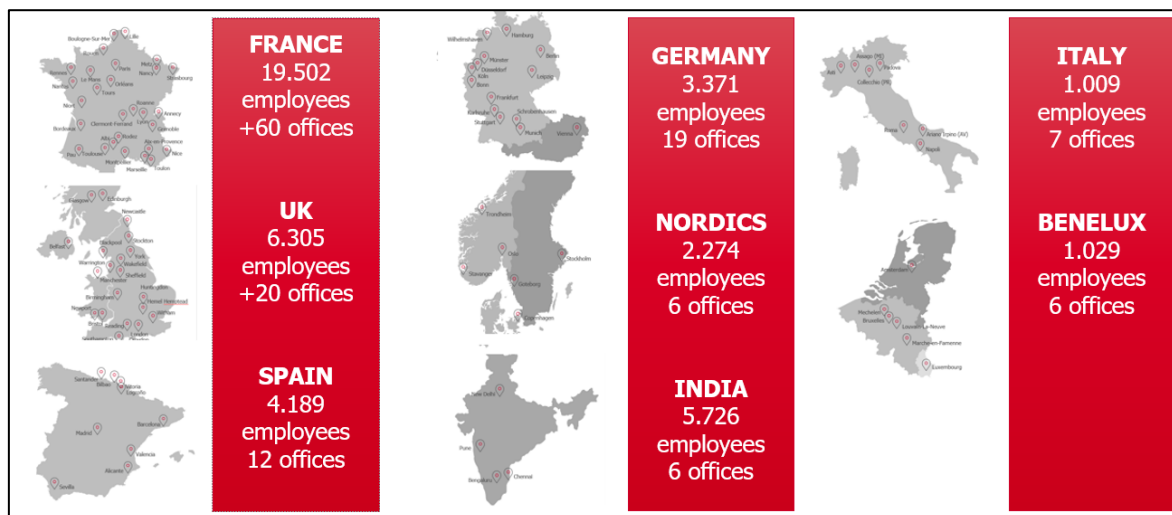


Figure 2 : Branches principales de Sopra Steria et leurs nombres de collaborateurs

Source : Sopra Steria. (2020, juin 1). *Sopra Steria Benelux* [Diapositives]. Teams

### 1.1.3. Secteurs d'activité

Les secteurs dans lesquels Sopra Steria opère à travers ses différentes branches sont nombreux et variés. Comme présenté dans l'historique ci-avant, le secteur bancaire et le secteur public sont des secteurs qui ont toujours été à l'avant-plan pour le groupe et c'est



pourquoi ils représentent aujourd'hui chacun presque un quart de l'activité. D'autres secteurs ont toujours été des pans importants de l'activité de l'entreprise, tels que la télécommunication/médias, les transports, l'assurance ou encore l'énergie représentant chacun entre 4 et 7 % du chiffre d'affaire (Figure 3).

Toutefois, les secteurs qui vont nous intéresser prioritairement dans le cadre de ce projet sont la sécurité, la défense et l'aéronautique. Ces trois secteurs sont regroupés au sein d'une seule agence s'appelant *HLS, Defence & Aeroline*, HLS étant l'abréviation de *HomeLand Security*. *HLS Defence & Aeroline* pèse 17 % dans le portfolio de Sopra Steria, ce qui en fait la troisième source de revenus (Sopra Steria, 2020).

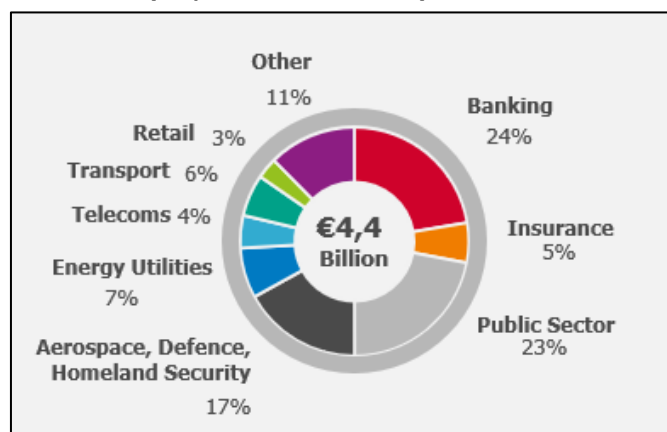


Figure 3 : Poids relatifs des secteurs d'activité de Sopra Steria

Source : Sopra Steria. (2020, juin 1). *Sopra Steria Benelux* [Diapositives]. Teams

Cette agence est particulièrement importante pour la branche Sopra Steria Benelux qui se voit régulièrement décerner des contrats majeurs dans ces domaines. En effet, au cours de ces dernières années, Sopra Steria Benelux a remporté plusieurs appels d'offres rapportant chacun plusieurs dizaines voire, dans certains cas, plusieurs centaines de millions d'euros. Désormais, elle fait partie des leaders sur le marché des systèmes d'information (SI) européens.

#### 1.1.4. Services

Le développement et l'intégration de systèmes représentent un peu plus de la moitié de l'activité totale de l'entreprise. Quatre autres types de services complètent leur offre, à savoir la vente de solutions, la gestion d'infrastructures informatiques, les *Business Process Services* et le *Consulting*. La répartition des activités de ces secteurs est reprise à la Figure 4 (Sopra Steria, 2020).

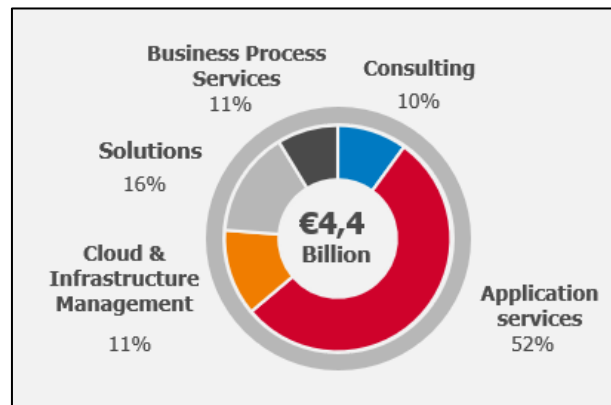


Figure 4 : Poids relatifs des services de Sopra Steria

Source : Sopra Steria. (2020, juin 1). *Sopra Steria Benelux* [Diapositives]. Teams

### 1.1.5. Mission

La mission de Sopra Steria est de mettre le digital au service de l'humain pour construire un avenir positif, en poussant leurs collaborateurs et clients à oser des choix audacieux. Ils comptent remplir leur mission en utilisant les technologies et en allant au-delà, en faisant confiance à l'intelligence collective. Leur slogan : « Sopra Steria. The world is how we shape it. » (Le monde est comme nous le façonnons).

### 1.1.6. Valeurs

Sopra Steria met en avant plusieurs valeurs qu'elle veut collectivement partager (Sopra Steria, 2020) :

#### 1. L'ouverture et la curiosité

Ils encouragent les initiatives afin d'explorer de nouvelles voies et technologies qui puissent bénéficier à tous.

#### 2. Le sens du collectif

Ils comptent beaucoup sur l'intelligence collective et l'esprit d'équipe pour que tous puissent contribuer à faire avancer le monde.

#### 3. La primauté du service au client

Sopra Steria s'engage avec ses clients sur le long terme afin de développer leurs performances avec des technologies toujours innovantes.

#### 4. L'excellence professionnelle

L'approche offerte par Sopra Steria se revendique visionnaire et globale, elle permet à ses partenaires, clients et collaborateurs, de faire des choix audacieux et ainsi de transformer les opportunités en résultats tangibles et durables.

## 5. Le respect de l'autre

Afin de garantir ce sens du collectif, ils adoptent une posture d'écoute et de proximité avec leurs clients, partenaires et collaborateurs.

### 1.1.7. Concurrence

On peut tout d'abord compter les quatre organisations que l'on appelle communément les « Big 4s » que sont Deloitte, Price Waterhouse Coopers, Ernst & Young et KPMG, qui dépassent très largement Sopra Steria non seulement dans leurs domaines d'activités, car la transformation digitale n'est qu'un des nombreux secteurs qu'ils couvrent, mais aussi par leurs chiffres d'affaires, comme le montre la Figure 5 (chiffres de Wikipedia convertis USD vers EUR au 15 juin 2020). Cependant, ces entreprises offrent principalement des services de consultance là où Sopra Steria peut offrir des expertises plus techniques et accompagner ses clients tout au long de l'implémentation des solutions informatiques.

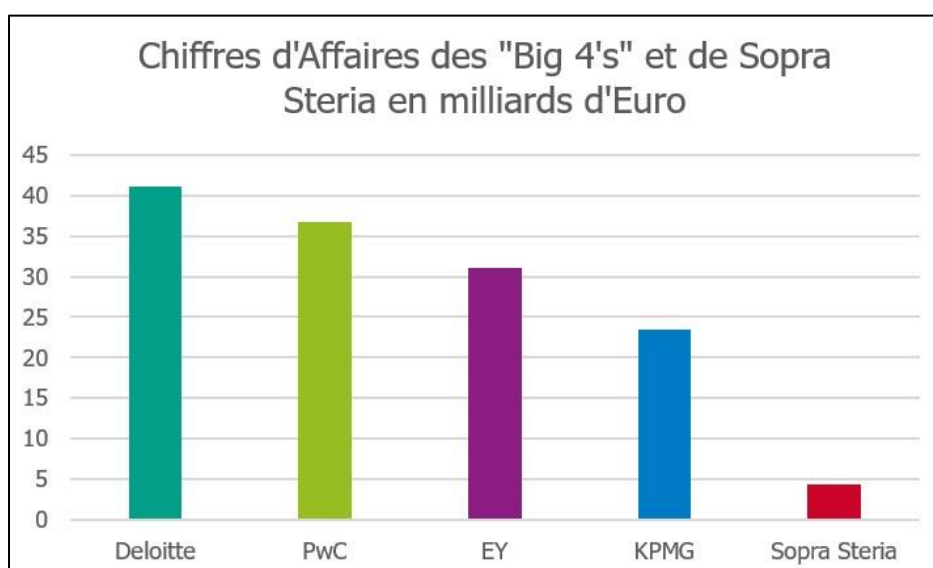


Figure 5 : Comparaison CA des "Big 4s" et de Sopra Steria

Par ailleurs, il existe des entreprises qui se rapprochent plus du business model de Sopra Steria comme Accenture, Cap Gemini, Atos ou encore Infosys pour ne citer que celles-là. Toutes ces organisations sont des Entreprises de Services Numériques (ESN). Bien qu'Accenture fasse figure de géant, les autres organisations, quoique plus importantes, sont comparables à Sopra Steria en termes de champ d'action et de revenus (Figure 6).

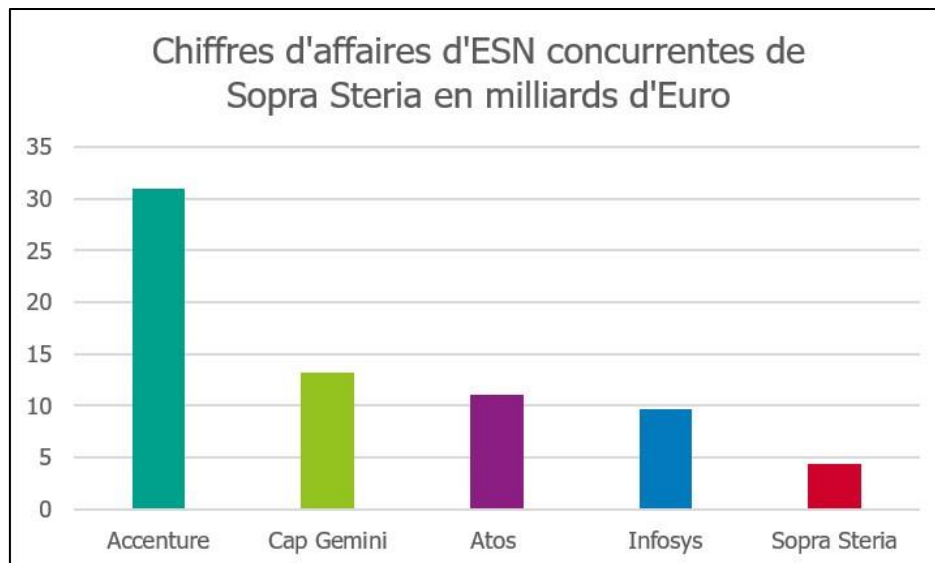


Figure 6 : Comparaison CA ESN et de Sopra Steria

Toutefois, il est intéressant de mentionner que bien qu'étant toutes concurrentes, ces entreprises collaborent régulièrement pour former des consortiums et rassemblent leurs expertises dans le but de couvrir au maximum les exigences de certains appels d'offres. En effet il n'est pas rare, par exemple, qu'un des *Big 4s* fasse appel à Sopra Steria comme gage de haute compétence en développement de systèmes d'information.

### 1.1. Branche Sopra Steria Benelux

Comme le montre la Figure 7, Sopra Steria est présente à six endroits dans le Benelux, dont quatre en Belgique (Malines, Bruxelles, Louvain-la-Neuve et Marche-en-Famenne). Au total, ce sont 1029 personnes qui sont employées par la branche (chiffres datant de juin 2019) ce qui en fait le septième groupe en termes de nombre de collaborateurs, juste après la Scandinavie et juste devant l'Italie. Les revenus en 2019 s'élevaient à 80 millions d'euros (Sopra Steria Benelux, 2020).

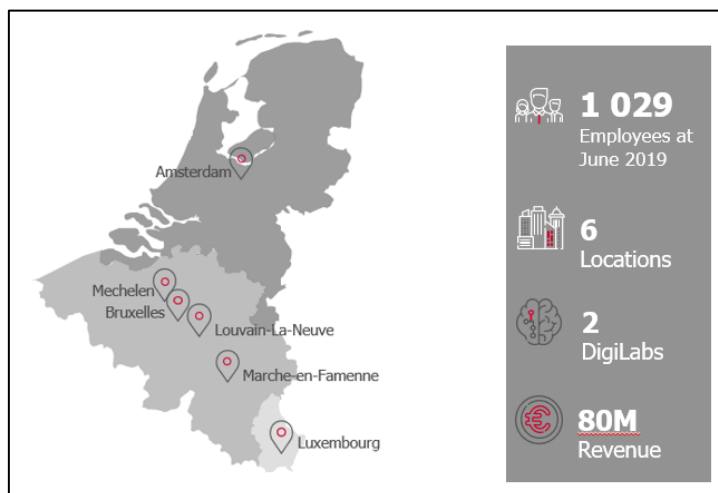


Figure 7 : Géographie de Sopra Steria Benelux

Source : Sopra Steria. (2020, juin 1). *Sopra Steria Benelux* [Diapositives]. Teams

Les secteurs d'activités cités ci-avant s'organisent autour de quatre axes au sein de la branche Benelux (Figure 8). Il y a tout d'abord le secteur public belge et européen. En tout, ce sont plus de 50 ministères et 60 agences publics à travers l'Union Européenne qui collaborent avec Sopra Steria concernant tout type de service public. Cela s'inscrit dans la volonté des divers gouvernements d'accélérer la digitalisation des services publics et l'« e-Governance ». Ces collaborations consistent en grande partie au développement de systèmes. Les contrats comprennent aussi, mais dans une moindre mesure, des services de *Data Centers*, du conseil et de la cybersécurité.

Le second axe traite de la sécurité intérieure. Cela fait maintenant presque vingt ans que Sopra Steria développe des systèmes et des bases de données centralisées servant au contrôle des frontières et à la migration au sein de l'Union Européenne. C'est l'axe sur lequel nous nous concentrerons en priorité dans le cadre de ce projet.

Le troisième axe concerne le secteur privé. On y retrouve les activités bancaires, le transport et l'énergie.

Le quatrième et dernier axe a trait à l'aérospatiale. La présence de Sopra Steria dans ce secteur est encore marginale. Néanmoins, le groupe dans son ensemble, ainsi que la branche du Benelux, ont envie de s'implanter durablement dans ce marché (Sopra Steria Benelux, 2020).

En plus de ces quatre axes, Sopra Steria Benelux est capable d'offrir quatre types de services transversaux, appelés centres de compétences.

Voici une illustration de transversalité des centres de compétences.

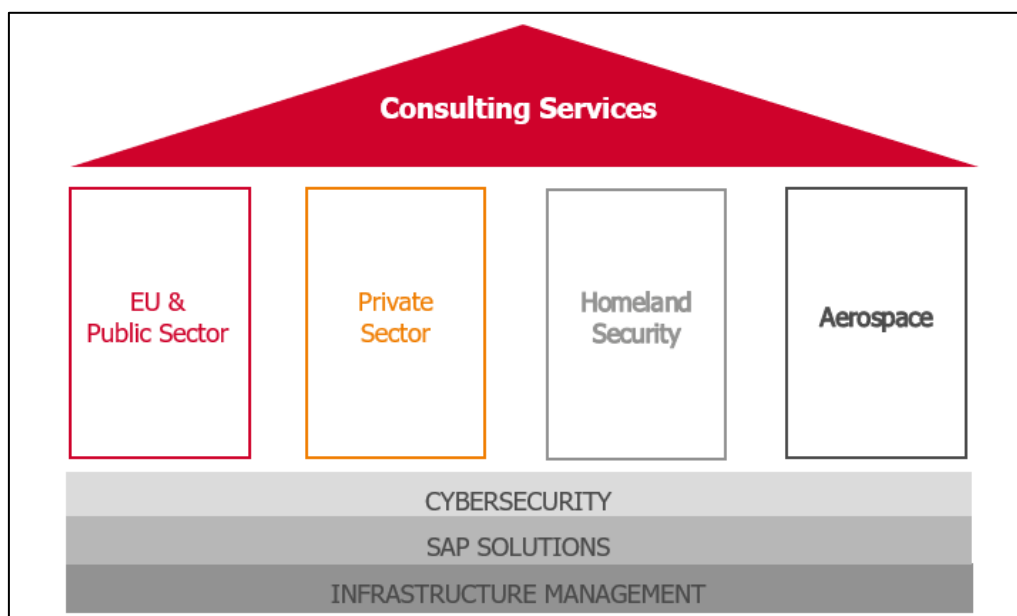


Figure 8 : Axes verticaux et centres de compétences transversaux

Source : Sopra Steria. (2020, juin 1). *Sopra Steria Benelux* [Diapositives]. Teams

Le premier de ces centres de compétences sur lesquels les axes verticaux peuvent reposer est la cybersécurité. L'entreprise se targue de mettre la cybersécurité au cœur du business et son approche « *Security by Design* » se concentre sur les risques inhérents aux affaires.

Le second centre de compétences est l'assistance à la configuration d'*Entreprise Resource Planning* (ERP, Progiciel de Gestion Intégré en français) et plus particulièrement à la configuration et/ou à la migration vers SAP.

Le troisième centre est la gestion des infrastructures, avec comme objectif de garantir des infrastructures agiles et efficaces.

Enfin, le quatrième centre de compétences, dans lequel s'inscrit *EU Monitoring* est le Consulting et est le sujet du point suivant.

## 1.2. L'Agence Consulting

Si la branche du Benelux remporte régulièrement des contrats importants de développement et d'intégration, l'agence de conseil de la branche reste mineure en comparaison à ses équivalents à travers le groupe. Cependant, il y a une volonté claire de devenir un acteur majeur du conseil dans le marché du Benelux, et c'est à cette fin qu'intervient *EU Monitoring*.

Le conseil que propose Sopra Steria est bien entendu un conseil qui a comme objectif la transformation digitale. Dans certains pays, Sopra Steria est un acteur majeur du marché du conseil, par exemple en France ou au Royaume-Uni. En Europe, Sopra Steria emploie en tout 3500 consultants (Sopra Steria, 2020).

Si cela n'est pas (encore) le cas dans le Benelux, ils veulent se positionner comme étant un acteur pertinent du conseil et ainsi pouvoir compléter leur offre. En effet, Sopra Steria se veut être une entreprise qui est capable de présenter une offre « end-to-end ». Cela signifie que l'accompagnement d'un client débute par la définition de ses besoins et la stratégie pour les combler (le conseil), ensuite viennent le développement de solutions, l'intégration, la maintenance et le support. Sopra Steria est donc capable d'être responsable de tous ces aspects, mais dans le Benelux le conseil est l'aspect le moins mature.

Il existe deux pans au conseil qui sont le secteur privé et le secteur public. Les contrats actuels de type privé concernent les banques, le transport ou encore le secteur pharmaceutique. Les contrats de type public concernent quant à eux divers ministères de différents pays, des forces de l'ordre nationales, mais aussi des institutions et agences de l'Union Européenne (UE). Ce sont sur ces dernières qu'*EU Monitoring* va se concentrer.

## **2. Approche théorique de la veille législative européenne**

Dans cette section, nous nous attèlerons à définir ce qu'est une veille législative. Dans la littérature disponible sur internet, il est plus souvent question de veille juridique ou encore de veille réglementaire. Néanmoins, les principes restent les mêmes et nous considérerons ces termes comme des synonymes dans le cadre de cette recherche.

Pour rappel, notre question de fond est « Comment concevoir un système de veille législative européenne efficace ? ». Nous distinguons 4 aspects principaux. Tout d'abord, il faut aborder le principe de veille pour ensuite s'intéresser plus précisément à l'aspect législatif de notre veille. Après cela, il conviendra d'aborder les critères d'efficacité d'une veille. Finalement, nous contextualiserons en apportant le cadre européen de la veille.

### **2.1. Veille**

Peu importe l'implication, le terme de veille dans le sens dans lequel on l'aborde ici est relativement récent. En français, on peut observer ses premières apparitions au début des années 1980 (Leitzelman, 2009). Daniel Douach, en 1996, en donnait la définition suivante : « l'art de repérer, collecter, traiter, stocker des informations et des signaux pertinents (faibles ou forts) qui vont irriguer l'entreprise à tous les niveaux de rentabilité, permettre d'orienter le futur (technologique, commercial...) et également de protéger le présent et l'avenir face aux attaques de la concurrence (...). La veille se pratique dans la légalité et le respect des règles de déontologie (...). En d'autres termes, la veille conduit à transférer des connaissances de l'extérieur vers l'intérieur de l'entreprise, en suivant des règles de fonctionnement strictes. »

Selon G. Salzano, M. Ankoud et O. Desfriches-Doria (2013), la veille s'organise selon un processus itératif, chaque itération comprenant plusieurs phases. Ils s'appuient sur les

travaux de Mylène Leitzelman qui, en 2009, présente quatre différentes phases de veille supportées par les outils du web 2.0 :

- 1) Définition des besoins de recherche : Social bookmarking
- 2) Collecte de l'information : Blogs, CMS, flux RSS, social bookmarking, réseaux sociaux
- 3) Analyse et traitement de l'information : Mashups (applications web combinant des contenus en provenance de différents sites)
- 4) Diffusion de l'information : Blogs, Wiki, CMS, flux RSS, Mashups, Widgets

Notons cependant que, davantage que les outils en tant que tels, ce sont l'observation et l'analyse des données produites par ceux-ci qui peuvent contribuer à enrichir un système de veille (Salzano et al, 2013).

#### Définition de Web 2.0 de Technopedia (Traduction de l'anglais)

Le Web 2.0 est un terme qui désigne l'Internet interactif d'aujourd'hui. Il est souvent comparé au Web 1.0, l'Internet des années 1990, et au futur Web 3.0 théorique qui implique des technologies avancées supplémentaires pour améliorer la façon dont nous utiliserons probablement le web dans des décennies.

Nous comprenons que la veille est un procédé qui peut s'adapter à quasi tous les secteurs. On évoque d'ailleurs souvent la veille technologique qui consiste à se tenir informé des innovations dans un secteur déterminé. La veille concurrentielle est également un type de veille répandu qu'on peut définir comme un dispositif de surveillance continue des actions et produits des concurrents actuels ou potentiels (Bathelo, 2016).

Toutefois, le sujet qui nous intéresse est celui de la veille législative. C'est une activité qui consiste à identifier parmi la production législative (nationale, européenne, internationale) les nouveaux textes qui sont applicables à l'entreprise, et à définir les actions nécessaires à mettre en place pour se conformer à ces nouvelles exigences (AQSE, 2016).

## 2.2. Spécificités de la veille législative

Le site internet Manager GO nous donne un bel aperçu des points qu'il faut prendre en compte lorsque nous abordons toute veille législative, juridique ou réglementaire. Une problématique y est tout d'abord identifiée : Comment reconnaître les sources d'information juridique fiables et leur pertinence ? La réponse se compose de deux éléments principaux.

Premièrement, il faut connaître la hiérarchie des normes. En effet, dans le droit national, une loi n'est pas équivalente à un décret ou à une ordonnance. Au niveau de pouvoir européen, il est important de pouvoir discerner (dans l'ordre chronologique) les propositions, les amendements et les réglementations.



Deuxièmement, il est important de pouvoir reconnaître les sources sûres. Nous évoquerons la problématique des fake news un peu plus loin dans le document.

Lorsque les sources fiables sont identifiées, un cycle de quatre étapes doit être respecté (Manager GO, 2020). On se rend compte qu'il y a de fortes similitudes avec les quatre étapes de Leitzelman exposées ci-avant. Nous l'appellerons le cycle de Manager GO pour le différencier du cycle de Leitzelman.

- 1) Il faut tout d'abord rassembler les informations pertinentes et cela, plus largement que l'activité précise de l'organisation. Il est important de ne pas trop restreindre la veille.  
Prenons un exemple hypothétique de notre champ d'application. Lorsqu'un nouveau système de contrôle des frontières entrera en vigueur partout dans l'espace Schengen, le ministère de l'intérieur Belge devra s'y conformer. Il faudra cependant aussi qu'il se conforme au RGPD. Dès lors, le support nécessaire n'est pas restreint à un seul sujet.
- 2) Ensuite, il est nécessaire de suivre l'évolution des projets de loi liés au domaine en question d'un projet de loi. Il faut sans arrêt être informé des avancées d'un texte dans le processus législatif. Pour cela il est pertinent de connaître la procédure législative de l'UE.
- 3) Une fois l'actualité législative analysée, il faut pouvoir la diffuser aux personnes et organisations concernées pour qu'elles puissent s'y conformer le plus rapidement possible.
- 4) L'information ainsi partagée doit finalement être stockée pour pouvoir être accessible à tout moment.

### 2.3. Critères d'efficacité

Dans la question introductive, nous nous demandons comment créer un système de veille législative efficace. Le principe de veille a été défini. Toutefois, il nous faut toujours savoir comment en évaluer l'efficacité.

Pour rappel de l'introduction, *EU Monitoring* vise à être un service pour les clients de Sopra Steria autant qu'un outil interne pour l'aide à la décision. Les décisions entraînées par *EU Monitoring* pourraient concerner des aspects différents selon l'utilisateur du service. En effet, pour les clients, il s'agit de se conformer aux législations et mettre en place des processus adéquats, tandis que pour Sopra Steria, cela devra surtout servir à l'exploration d'opportunités. Cependant, nous pouvons considérer l'évaluation de l'efficacité de la veille de manière similaire.

Depuis que le principe de veille existe, les professionnels et les académiques ont toujours voulu pouvoir en évaluer l'efficacité. C'est évidemment une question primordiale, mais dont la réponse ne coule pas de source pour autant.

La veille sert à extraire des informations et il revient aux preneurs de décisions de considérer ou non ces informations (Roulet et al, 2015). Même dans le cas où les résultats de veille sont pris en compte, comment évaluer leur contribution dans la prise de décision et finalement dans les conséquences de cette décision ? L'évaluation est d'autant plus complexe que la veille n'est qu'un facteur parmi d'autres, tous interconnectés.

En effet, comme d'autres actifs intangibles (gestion de la connaissance, gestion des ressources humaines, etc), la veille ne crée pas de valeur en soi, mais uniquement lorsqu'elle est combinée avec ces différents actifs. Nous ne pouvons donc dès lors pas estimer facilement son impact sur les performances. C'est d'autant plus vrai lorsque nous considérons la temporalité d'une décision et ses retombées différées dans le temps pouvant paraître bonnes au début et se révéler mauvaises par après (Roulet et al, 2015).

Ces aspects sont clairement des freins à l'évaluation, mais il est malgré tout possible de définir l'efficacité de la veille comme le degré d'atteinte des objectifs, la comparaison entre les résultats attendus avec les résultats effectifs (Roulet et al, 2015). L'évaluation dépend aussi de la mission que nous attribuons à la veille. Les praticiens affirment en général que l'objectif principal de l'activité de veille est l'aide à la décision. C'est le cas d'*EU Monitoring*.

On attend donc que la veille contribue à améliorer la communication entre les niveaux hiérarchiques de l'entreprise, l'implication des collaborateurs dans le processus de décision, de management et de planification. Toutes ces attentes sont intangibles et dès lors soulignent à nouveau la difficulté d'établir un outil de mesure. Aurélie Roulet et ses collègues relèvent néanmoins trois catégories d'outils plus ou moins sophistiqués : les approches « *Measures of effectiveness* », les approches « *Return on Investment (ROI)* », et les approches de « *Balanced Scorecard (BSC)* ».

### 2.3.1. Measures of Effectiveness

J. Herring (1999) propose une approche qui mesure l'efficacité d'une veille selon quatre mesures. Cette approche s'inspirant des mesures développées pour les technologies de l'informatique, prend en compte les économies de temps, les coûts évités, les économies d'argent et l'augmentation des revenus. Selon Roulet et al., cette approche pertinente ne résout cependant pas la problématique de l'identification de l'apport réel de la veille parmi les autres actifs intangibles.

### 2.3.2. Return on Investment (ROI)

Le retour sur investissement est une des mesures principales de la performance d'une entreprise. Normalement, il s'agit de comparer les retombées financières d'une activité avec les investissements effectués pour faire fonctionner cette activité.

Cependant, L. Davidson (2001), propose de se servir du principe de *Return on Competitive Intelligence Investment (ROCII)*. Ce principe issu de la publicité

catégorise les bénéfices en résultats tactiques et stratégiques. Cela regroupe de nombreux indicateurs comme la mesure des facteurs de risques, d'atteintes des objectifs, d'augmentation de valeurs et de satisfaction. Ces indicateurs sont mesurables mais la transposition en valeur monétaire est malgré tout difficile.

On peut malgré tout calculer les économies réalisées au cours d'un projet pendant lequel la veille a été utilisée. Si on considère ces économies comme imputables à la veille, nous pouvons avoir une mesure monétaire de son apport. Une fois encore, nous ne mesurerions pas la veille en tant que telle, mais bien en tant que composante d'un tout.

### 2.3.3. Balanced Scorecard (BSC)

Ce système de mesure alternatif créé en 1992 par R. Kaplan et D. Norton décline la stratégie de l'entreprise en quatre perspectives comme le montre la Figure 9: Client, Processus interne, Innovation & Apprentissage et Finance. Ce tableau de bord prospectif prend malgré tout en compte l'aspect financier, car cela fait évidemment partie d'une entreprise. Il y a peu d'intérêt d'améliorer le management si cela ne se traduit pas en une augmentation des bénéfices.

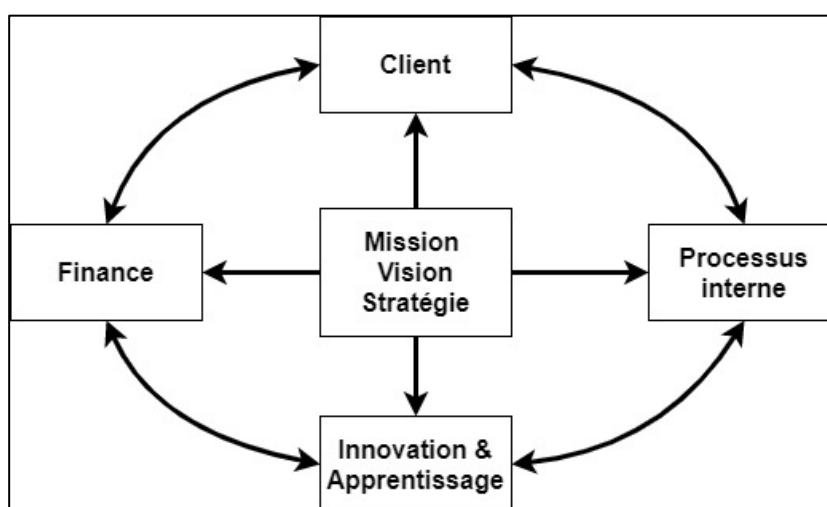


Figure 9 : Les axes de la BSC

Source : Roulet, A., Bezençon, C., & Madinier, H. (2015). Évaluation de la performance et de l'impact de la veille. *I2D - Information, données & documents*, 52(3), 70. <https://doi.org/10.3917/i2d.153.0070>

Pour chacun de ces quatre axes, il faut identifier un objectif formulé sans valeur claire (ex. : Augmenter le nombre de clients satisfaits), un indicateur (ex. : Taux de clients satisfaits), une valeur cible (ex. : 95% de clients satisfaits) et une initiative stratégique, c'est-à-dire ce qui va être fait pour atteindre l'objectif.

Ce modèle souple s'aligne sur la stratégie de l'entreprise et c'est pour cela qu'il est possible de l'utiliser pour évaluer une veille (Roulet et al, 2015).

## 2.4. Institutions et Agences de l'Union Européenne

À présent, intéressons-nous aux spécificités de l'environnement législatif européen. Cela permettra ensuite de cerner facilement la portée précise du projet. En effet, comme nous l'avons identifié en tant que prérequis d'une veille législative, il faut connaître le fonctionnement législatif européen ainsi que les acteurs pertinents.

### 2.4.1. Procédure Législative Ordinaire de l'UE

Avant d'utiliser le terme de procédure législative ordinaire, on faisait référence à la procédure de codécision. Elle fut introduite pour la première fois en 1992, avant que son utilisation soit étendue en 1999. Cela représente maintenant la principale procédure de prise de décision au sein de l'UE s'appliquant à environ 85 champs d'action. Tout commence avec le droit d'initiative de la Commission.

#### i. La Commission Européenne

La Commission européenne, organe exécutif, est l'institution qui se rapproche le plus d'un gouvernement d'un pays. Elle est dirigée par 27 commissaires (un par pays membre). Ensemble, ils s'appellent le « collège ». Ce sont eux qui orientent la politique et la stratégie de la Commission. Le collège est renouvelé tous les cinq ans, à la suite des élections européennes et reflète les résultats de celles-ci, sans toutefois que leurs membres soient directement élus (Commission Européenne, 2016).

L'institution est organisée en plusieurs « directions générales » (DG). Comme la comparaison faite plus en avant dans le texte, ces DG sont l'équivalent des ministères. C'est à ces services qu'il revient de mettre en œuvre et de gérer les politiques, la législation et les budgets de l'UE. Les DG appliquent la politique de la commission à travers des agences, comme eu-LISA par exemple. eu-LISA est une agence européenne, client de Sopra Steria qui sera détaillée au point 2.4.2.

Le rôle de la commission dans la procédure législative ordinaire est donc de proposer de nouvelles initiatives, rôle dont elle a le monopole. Avant cela, elle se doit d'évaluer les implications économiques, sociales et environnementales de la législation à mettre en place. C'est dans cette optique que la commission procède à des études d'impact.

Pour l'accompagner dans ce processus, elle a recours à des groupes d'experts pour la conseiller sur des questions techniques. De plus, les parties prenantes comme les ONG, les autorités ou encore les entreprises, sont, elles aussi, consultées.

Les différents parlements nationaux sont également invités à faire part officiellement des potentielles réserves qu'ils auraient à formuler quant à la pertinence de légiférer à propos d'une problématique au niveau européen plutôt qu'au niveau national.

Une fois ces considérations traitées, la commission rédige un texte qui commence en général par :

Proposition de  
**RÈGLEMENT DU PARLEMENT EUROPÉEN ET DU CONSEIL**

La proposition est donc transmise au Parlement et au Conseil pour qu'ils révisent et adoptent parallèlement le texte de loi.

**ii. Le Parlement Européen**

À l'instar de la Commission, le Parlement est renouvelé tous les cinq ans, mais est cependant le seul organe de l'Union à être directement élu. De manière similaire à la Commission, le Parlement est divisé en groupes selon les différentes compétences. Ces subdivisions s'appellent les commissions parlementaires.

Il est intéressant de noter que les compétences de ces commissions parlementaires ne correspondent pas exactement aux directions générales. Autrement dit, une commission parlementaire peut avoir un cadre de compétences reprenant potentiellement les compétences de plusieurs directions générales de la Commission européenne. Par exemple, la commission parlementaire des libertés civiles, de la justice et des affaires intérieures (LIBE) légifère sur des textes pour lesquels la direction générale de la migration et des affaires intérieures (DG HOME) est compétente, mais aussi sur d'autres textes relevant du cadre de la direction générale de la justice (DG JUST).

Les traités européens (textes fondateurs et formateurs de l'Union Européenne) ont confié au Parlement un large éventail de compétences et de pouvoirs, dont le pouvoir législatif (Parlement Européen, 2020).

Le pouvoir législatif est ce qui lui confère le devoir d'adopter la législation de concert avec le Conseil. Ces deux institutions opèrent sur un pied d'égalité.

Lors de ce travail législatif, un rapport sur la proposition reçue de la Commission est rédigé. Ce rapport est alors voté, et amendé au besoin, dans le cadre de la commission parlementaire compétente. Le Parlement est alors prêt pour les discussions avec le Conseil et la Commission.

**iii. Le Conseil Européen**

La proposition en provenance de la Commission est assignée à un des 150 groupes de travail du Conseil Européen, formé d'experts représentant tous les pays membres. Ils analysent en détail la proposition afin d'arriver à un accord quant à la position à prendre dans les

négociations avec le Parlement, ce qui comprend notamment des amendements à la proposition.

Une fois cet accord atteint, il est envoyé au COREPER, où siègent des représentants permanents qui sont des ambassadeurs des pays membres. Le COREPER, se charge de préparer les travaux du Conseil Européen. Pour cela, le COREPER doit se mettre d'accord. Ce n'est qu'à ce moment-là que le Conseil est prêt à participer au trilogue avec le Parlement et la Commission (Conseil Européen, 2020).

Le Conseil sera représenté par la Présidence tournante de l'Europe. La Présidence est prise en charge par un pays membre pendant six mois avant qu'un autre pays membre ne soit désigné, suivant un ordre bien établi. À l'heure d'écrire ce texte, la Présidence est assurée par l'Allemagne.

#### iv. Le Temps du Trilogue

Quand le Parlement et le Conseil ont chacun pu amender la proposition de la Commission, vient le temps du Trilogue. Nous appelons trilogue la phase des discussions entre les trois institutions (Figure 10). Le Parlement et le Conseil négocient sur les amendements qu'ils ont fait chacun de leur côté. La Commission, quant à elle, défend son projet de loi initial et joue également un rôle de médiateur entre les deux autres institutions. Lorsqu'ils sont tous d'accord sur une première version (draft), le texte est présenté en séance plénière du Parlement afin d'être voté. Si ce draft est approuvé par le Parlement, le Conseil doit également donner son accord. S'il donne effectivement son accord, on dit alors que la législation est approuvée en première lecture. Sinon, les discussions reprennent dans le cadre d'une seconde lecture. Si après cette seconde lecture, le projet de loi n'est toujours pas sujet à un accord de toutes les parties, un comité de conciliation est mis en place afin de finalement faire accepter le texte par toutes les parties.

Lorsque la législation est acceptée par tous, le texte est alors publié au Journal Officiel et devient une loi de l'Union européenne (Conseil Européen, 2020).

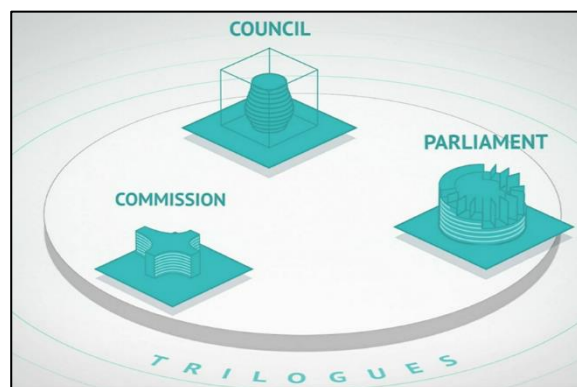


Figure 10 : Temps du trilogue

Source : Conseil Européen. (2020, 31 janvier). Le processus de prise de décision du Conseil. European Council. <https://www.consilium.europa.eu/fr/council-eu/decision-making/>

#### 2.4.2. Les Agences Européennes

Certaines agences représentent des clients stratégiques pour Sopra Steria et sont souvent à l'origine de contrats de très grande envergure. Par exemple, eu-LISA, présentée ci-après, a récemment émis un ensemble d'appels d'offres dont la valeur totale avoisine le milliard

d'euros. Nous nous concentrerons principalement sur les agences qui dépendent de la direction générale de la migration et des affaires intérieures (DG HOME).

### i. eu-LISA

L'Agence européenne pour la gestion opérationnelle des systèmes d'information à grande échelle au sein de l'espace de liberté, de sécurité et de justice (eu-LISA) représente l'un des plus gros clients de Sopra Steria Benelux. Pour cette raison, *EU Monitoring* se concentrera en grande partie sur les législations concernant cette agence. C'est pourquoi il est pertinent de décrire cette agence et ses activités (eu-LISA, 2020).

La mission d'eu-LISA est d'ajouter continuellement de la valeur aux états membres en supportant leurs efforts pour une Europe plus sûre grâce à la technologie. Son rôle est donc de fournir des solutions et des services efficaces et de grande qualité et d'adapter les technologies aux besoins évolutifs des états membres.

eu-LISA fût créée en suivant la procédure législative ordinaire (décrite ci-avant) sur la base d'une proposition de la Commission et sur la base du fait que le Conseil puisse adopter des mesures pour assurer la coopération des départements pertinents des états membres dans le cadre de la liberté, la sécurité et la justice. Cette agence est régulée et possède une personnalité légale. De plus, dans chaque état membre, eu-LISA profite des capacités les plus larges accordées à toute personnalité juridique dans la cadre de la loi nationale.

Cette agence offre l'expertise requise pour développer et gérer les six grands systèmes d'information et leurs infrastructures respectives, tous les jours et à toute heure de la journée et de la nuit. En plus d'accompagner les états membres dans l'échange d'informations, elle applique les standards de protection de données pour chacun des six systèmes. Trois de ces systèmes sont déjà opérationnels depuis plusieurs années :

1. *European Asylum Dactyloscopy Database (Eurodac)* : Ce système, en enregistrant les empreintes digitales des demandeurs d'asile, permet d'appliquer les principes de la convention de Dublin, qui prévoit quel pays est responsable d'une demande d'asile.
2. *Schengen Information System (SIS)* : supporte l'échange d'informations à propos de personnes et d'objets entre les services de police nationaux, de gardes-frontières, de douanes, de visas et des autorités judiciaires.
3. *Visa Information System (VIS)* : permet aux autorités nationales désignées de consulter les données, dont la biométrie, pour les visas de courte durée dans l'espace Schengen.

Trois autres systèmes sont en cours de développement ou le seront prochainement :

4. *Entry/Exit System (EES)* : enregistre électroniquement le temps et l'endroit des entrées et des sorties des ressortissants des pays tiers à l'Union.



5. *European Travel Information Authorisation System (ETIAS)* : à l'instar de l'ESTA américain, ETIAS sera un système d'autorisation d'entrée sur le territoire pour les ressortissants de pays qui n'ont pas besoin de visa.
6. *European Criminal Records Information System – Third Country National (ECRIS-TCN)* : permettra aux autorités des états membres d'identifier les dossiers criminels détenus par les autres états membres, concernant des individus venant de pays hors de l'Union.

Hormis l'évolution et le développement individuel de chacun de ces systèmes, eu-LISA a été chargée d'en assurer l'interopérabilité. Cela se composera de plusieurs éléments, dont l'ensemble permettra d'offrir des données plus sûres et plus rapides d'accès (Figure 11).

Non seulement cela servira aux services de gestion des frontières, mais aussi aux autres forces de l'ordre, lorsqu'ils travailleront sur le terrain ou dans des bureaux. Ces travaux d'interopérabilité sont extrêmement complexes, variés et intéressants, mais dépassent le cadre de ce document.

Notons néanmoins que le contrat pour le composant appelé *Shared Biometric Matching System – sBMS* (Système de Correspondance Biométrique Partagé) a très récemment été remporté par Sopra Steria Benelux. Il s'agit d'un contrat majeur qui occupera plusieurs dizaines de personnes pendant plusieurs années. Cela montre l'importance d'eu-LISA et plus largement des agences européennes pour l'entreprise.

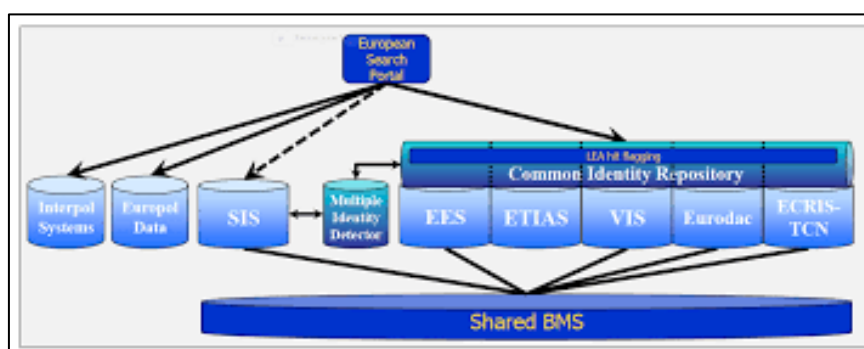


Figure 11 : Schéma de l'interopérabilité des systèmes d'information de l'espace de liberté, de sécurité et de justice (eu-LISA)

Source : eu-LISA. (s. d.). Large-Scale IT Systems. Consulté le 16 février 2020, à l'adresse <https://www.eulisa.europa.eu/Activities/Large-Scale-It-Systems>



## ii. Frontex

L'*European Border and Coast Guard Agency (EBCGA)*, plus communément appelée Frontex, joue un rôle clé dans l'effort de l'Union Européenne pour la protection des frontières terrestres et maritimes de l'espace Schengen. Créée en 2004, l'agence Frontex est à un tournant de son existence avec des prévisions d'augmentation significatives de ses ressources financières, techniques et humaines. Afin d'exécuter ses prérogatives de protections des frontières, Frontex a besoin d'avoir accès aux données que les systèmes gérés par eu-LISA mettent à sa disposition.

Par ailleurs, Frontex développe ses propres systèmes d'information qui auront à leur disposition des moyens importants. Il n'est donc pas inimaginable qu'à un moment donné, Frontex devienne un client de Sopra Steria.

À titre d'exemple, en début d'année, Frontex a émis un appel d'offres afin de désigner l'entreprise qui les aidera à se connecter au Système d'Information Schengen (SIS). Sopra Steria y a bien sûr répondu, car il s'agirait d'une porte d'entrée chez Frontex. De plus, Sopra Steria a développé un système « clé sur porte », la SIB (Sopra Steria Interconnexion Box) qui permet de gérer les communications des systèmes et bases de données SIS nationaux avec le système SIS central. Ce système, déjà installé dans neuf états membres de l'UE, peut parfaitement convenir aux besoins de Frontex (Frontex, 2020).

## iii. Europol

Dans leur lutte contre la grande criminalité internationale et le terrorisme, les 27 états membres de l'Union Européenne peuvent compter sur le soutien d'Europol. Son action n'est cependant pas limitée aux frontières de l'Union, car elle est amenée à collaborer avec des pays partenaires étrangers et des organisations extérieures à l'UE, comme Interpol par exemple.

Ses domaines d'actions sont multiples et variés comme le terrorisme, le trafic international de stupéfiants, la fraude organisée ou encore le trafic des êtres humains. Comme mentionné pour Frontex, Europol s'est vu attribuer la gestion de systèmes informatiques lui permettant de rendre son action plus efficace (Europol, 2020).

Europol et Frontex auront également des responsabilités très importantes dans la mise en place de l'interopérabilité des systèmes d'informations gérés par eu-LISA.

## 2.4. Fiabilité des sources

À l'heure actuelle, il n'est pas rare d'être témoins de la création et de la propagation de *fake news* (fausses nouvelles). Il est impensable qu'*EU monitoring* relaie ce genre d'information. Les *fake news* sont pourtant toujours plus présentes et profitent de relais importants. En France, en 2018, un sondage affirmait que 24% des sondés avaient déjà relayé de l'information à caractère politique qui s'est révélée fausse. 8% déclarent même l'avoir fait

plus d'une fois (Statista, 2018). Aux Etats-Unis, dans le cadre de l'élection présidentielle de 2016, une étude a révélé qu'un quart des tweets contenant des liens externes amenaient vers des sites réputés pour publier des *fake news* (LaLibre, 2019).

Sachant cela, il est important de pouvoir faire confiance aux sources que nous utiliserons dans nos recherches d'information. Une première astuce dans notre cas est de se fier au site dont l'URL se termine avec l'extension « europa.eu ». Cependant, ces sources, bien que primordiales, ne sont pas suffisantes pour couvrir toute l'actualité et les « bruits de couloirs » européens, il faut donc savoir reconnaître les sources fiables pour lesquelles des outils existent.

Le DigiLab, le service de recherche et d'innovation au sein de Sopra Steria a développé un système s'appuyant sur des technologies sous-jacentes à l'intelligence artificielle (IA) comme le *Web Crawling* ou le *Machine Learning* et ils peuvent déterminer la fiabilité d'une information en analysant sa sémantique et sa syntaxe.

Le projet développé par le DigiLab fait partie d'une initiative de la Radio-Télévision Belge Francophone (RTBF) dans un but de *fact-checking* (vérification des faits). Ce service s'appelle Faky et rassemble, en plus, des services tels que « les Décodeurs » du Journal Le Monde référençant des sources non-fiables.

## 2.5. Synthèse de la veille

Comparons à présent les deux cycles de veille. La troisième et la quatrième étape du cycle de Manager GO sont ensemble équivalentes à la seule quatrième étape du cycle de Leitzelman. Par ailleurs, chacun des cycles reprend des étapes non comprises dans l'autre.

C'est pourquoi nous allons à présent réaliser notre propre cycle théorique rassemblant toutes les étapes des deux cycles.

Prérequis :

- Connaître les différents acteurs et documents impliqués dans la législation européenne;
  - S'assurer de la fiabilité des sources;
1. Définir les besoins de la veille;
  2. Rassembler les informations pertinentes plus largement que le but premier de la veille;
  3. Analyser et traiter l'information;
  4. Suivre l'évolution d'un projet de loi;
  5. Diffuser l'information;
  6. Stocker l'information.

Nous avons désormais un cycle de six étapes avec deux prérequis. Cela nous permet d'appréhender ce à quoi le *business process d'EU Monitoring* devrait ressembler. Ces étapes

permettent également d'identifier une série de points importants qu'un système de veille doit comprendre.

- Un accès à suffisamment d'information pertinente ;
- Un système de suivi ;
- Un système de diffusion d'information ;
- Un système de partage d'information

## Machine Learning

Dans la section qui va suivre ainsi que plus loin dans ce texte, il va être question à plusieurs reprises d'entraînement de modèles et de *Machine Learning* (apprentissage automatique). Il est donc intéressant de faire une courte digression à ce sujet.

Dans sa forme la plus basique, le *Machine Learning* se produit de la façon suivante :

Tout d'abord, des données brutes et connues sont récoltées pour ensuite être nettoyées en enlevant les éléments faux, incomplets, redondants, etc. Nous avons alors des données prétraitées.

Ces données prétraitées vont être séparées en deux grands ensembles, comme illustré à la Figure 12. Le premier contient les données d'entraînement et le second les données de test. Avec le premier ensemble, nous allons entraîner la machine. La proportion de chaque ensemble peut varier mais en général, cela tourne autour de deux tiers d'entraînement pour un tiers de test.

L'entraînement vise à créer un modèle. Ce modèle peut servir à différentes choses, comme reconnaître un chat d'un chien sur une photo ou encore à reconnaître un profil médical plus enclin à développer des cellules cancéreuses.

Une fois ce modèle développé, il sera mis à l'épreuve par l'ensemble des données de test. Notons que ces données sont historiques et connues. C'est-à-dire que l'on sait déjà quelles photos montrent un chat ou quels profils ont développé des cellules cancéreuses. Ainsi, on peut évaluer la précision des prédictions du modèle lorsqu'il est appliqué aux données de test.

On classe les résultats en quatre catégories : Vrai positif (VP), faux positif (FP), faux négatif (FN) et vrai négatif (VN). On peut alors calculer la précision du modèle avec cette formule :  $Précision = \frac{VP+VN}{VP+VN+FP+FN}$ . On obtient alors un pourcentage d'exactitude et selon le désir de précision, le modèle sera gardé ou non (Davis, 2019).

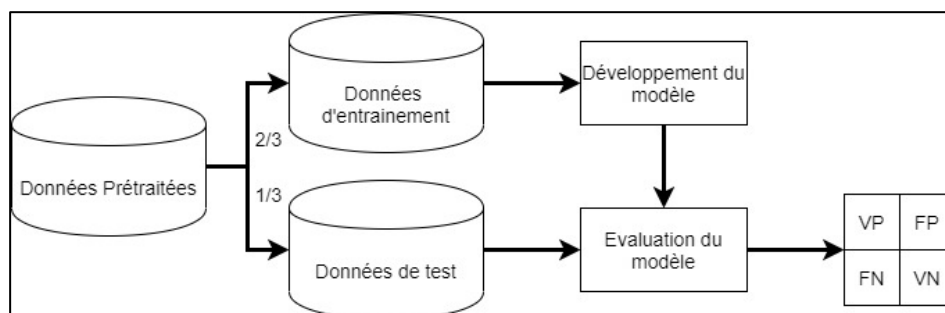


Figure 12 : Entraînement de modèles de *Machine Learning*

Si le modèle est acceptable, on peut alors y entrer des données inconnues pour les traiter et en déduire l'exactitude de l'information produite.

Ici, nous sommes dans une approche binaire. Pour plus de précision et de nuances, il faut associer plusieurs modèles.

## 2.6. Technologies potentiellement utiles

Nous le comprenons à présent, un système de veille traite de l'information. La quantité d'information peut même être très importante. C'est pourquoi nous nous intéressons à certaines technologies qui pourraient nous aider à traiter cette information. Les technologies abordées ici sont principalement des technologies du Web sémantique défini comme « une toile de données qui peuvent être traitées directement et indirectement par des machines pour aider leurs utilisateurs à créer de nouvelles connaissances » par Tim Berners-Lee, directeur du World Wide Web Consortium (W3C).

Lors de la mise en œuvre du projet et dans le bilan qui suivra, nous analyserons lesquelles de ces technologies sont réellement applicables à *EU Monitoring*.

### 2.6.1. Extraction de données non-structurées – *Text Mining*

Aujourd'hui, entre 85 et 90 % de l'ensemble des données d'entreprises ont une forme non-structurée. Il est estimé que ces données doublent de volume environ tous les dix-huit mois (Davis, 2019). Afin de rester compétitif, il est primordial de puiser dans ces données. Le type de données non-structurées principal sont les données textuelles. Ce sont également les textes qui nous intéressent principalement dans le cadre *EU Monitoring*.

Il est possible d'extraire ces données. Toutefois, les données non structurées ainsi extraites sont toujours des informations inconnues. Le *Text Mining* (fouille de texte) est l'extraction d'informations précédemment inconnues en extrayant des informations de différentes sources textuelles (Figure 13). L'extraction de contenu nécessite alors l'application de techniques de *Data Mining*.

Le *Basic Content Mining* est un type de fouille de textes. Les techniques utilisées dans la fouille de textes sont nombreuses mais celle qui est à la base de plusieurs autres est l'extraction d'informations (Johnson et Gupta, 2012). Cette technique s'efforce de suivre une logique qui est de structurer les données non-structurées afin de pouvoir ensuite y appliquer le *Data Mining* classique (Davis, 2019).

Pour extraire des informations de données non structurées présentes sur le web, on utilise le *pattern matching*. Cette méthode permet de retracer les mots et les phrases clés, puis de découvrir la connexion des mots clés dans le texte. Lorsque le volume de texte est important, cette technique est très utile. L'extraction d'informations transforme un texte non structuré en une forme plus structurée. Ainsi, au lieu de disposer d'un texte difficile à traiter pour un ordinateur, nous disposons de listes de données plus digestes pour une machine. L'information est d'abord extraite des données ainsi recueillies. Après avoir extrait les connaissances à partir de ces données, le système peut prédire les informations manquées par l'extraction précédente en utilisant des règles découvertes. Les règles d'extraction d'informations qui font des prédictions incorrectes sont rejetées (Gupta et Lehal, 2009).

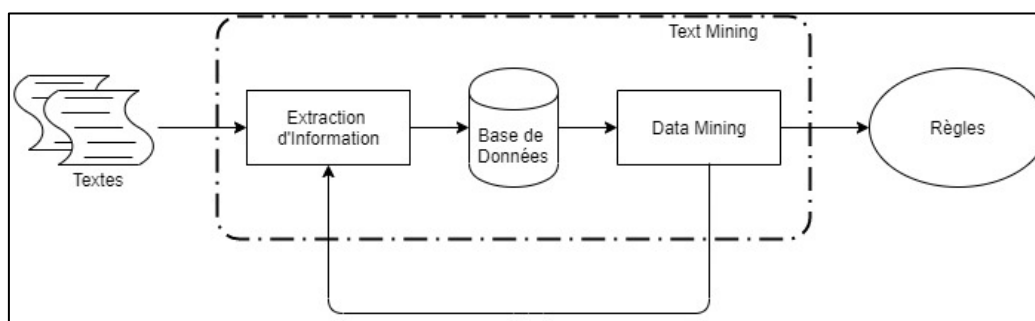


Figure 13 : Aperçu du fonctionnement de *Text Mining* basé sur l'extraction d'information

Source : Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60-76. <https://doi.org/10.4304/jetwi.1.1.60-76>

### 2.6.2. Traitement du langage naturel - *Natural Language Processing*

L'information législative est essentiellement sous forme de textes et de fichiers audio. Elle est donc exprimée en langages humains. Cependant, si l'on voulait s'aider d'ordinateurs pour traiter cette information, on aurait besoin que ceux-ci comprennent ces langages humains. C'est à cela que sert le *Natural Language Processing* (NLP).

Le *Natural Language Processing* est une branche de l'intelligence artificielle et consiste à appliquer des algorithmes pour identifier et extraire les règles du langage naturel de telle sorte que les données linguistiques non structurées soient converties sous une forme que les ordinateurs peuvent comprendre. Lorsque le texte a été fourni, l'ordinateur utilisera des algorithmes pour extraire le sens associé à chaque phrase et en recueillir les données essentielles (Garbade, 2018).

Leand Romaf, un ingénieur en logiciels expérimentaux qui se passionne pour l'enseignement du fonctionnement des systèmes d'intelligence artificielle, déclare que "ces dernières années, il y a eu des percées significatives pour permettre aux ordinateurs de comprendre le langage tout comme nous" (Garbade, 2018).

L'objectif ultime du NLP est de lire, déchiffrer, comprendre et donner un sens aux langues humaines d'une manière qui soit valable. La plupart des techniques de NLP s'appuient sur la *machine learning* pour tirer un sens des langages humains.

Toutefois, le traitement du langage naturel est considéré comme un problème difficile en informatique. C'est la nature du langage humain qui rend la NLP difficile. Les règles qui dictent la transmission d'informations à l'aide des langages naturels ne sont pas faciles à comprendre pour les ordinateurs.

Certaines de ces règles peuvent être de haut niveau et abstraites, par exemple, lorsque quelqu'un utilise une remarque sarcastique pour transmettre une information.

D'autre part, certaines de ces règles peuvent être de bas niveau, par exemple, l'utilisation du caractère "s" pour signifier la pluralité des éléments. Pour comprendre le langage humain

de manière exhaustive, il faut comprendre à la fois les mots et la façon dont les concepts sont liés pour délivrer le message voulu. Alors que les humains peuvent facilement maîtriser une langue, l'ambiguïté et les caractéristiques imprécises des langues naturelles sont ce qui rend la NLP difficile à mettre en œuvre pour les machines (Garbade, 2018).

Cette technologie repose en grande partie sur l'analyse syntaxique et sémantique.

La syntaxe désigne l'agencement des mots dans une phrase de manière qu'ils aient un sens grammatical. En NLP, l'analyse syntaxique est utilisée pour évaluer comment le langage naturel s'aligne sur les règles grammaticales. Les algorithmes de *machine learning* sont utilisés pour appliquer les règles grammaticales à un groupe de mots et en tirer un sens.

La sémantique, quant à elle, se réfère au sens véhiculé par un texte. L'analyse sémantique est l'un des aspects difficiles du traitement du langage naturel qui n'a pas encore été entièrement résolu. Elle consiste à appliquer des algorithmes informatiques pour comprendre le sens et l'interprétation des mots et la façon dont les phrases sont structurées (Garbade, 2018).

### 2.6.3. Automatisation des procédés – *Robot Process Automation*

Nous le constatons, le cycle de la veille législative est un processus reprenant plusieurs étapes qui se répètent. Pour cette raison, attardons-nous sur la *Robot Process Automation* (RPA).

La RPA utilise les intrants et les logiques métiers afin de rendre le *business process* automatique. Concrètement, il s'agit de créer un logiciel, appelé robot, permettant de capturer et interpréter les transactions et les traitements de données afin d'obtenir des réponses et de communiquer avec d'autres systèmes numériques (zdnet, 2020).

La RPA est particulièrement intéressante pour les *business process* qui se composent principalement de tâches répétitives et dont les besoins intellectuels sont faibles. Cela permet alors de libérer beaucoup de temps pour les humains en rendant la machine capable de les remplacer (lebigdata, 2020).

Maintenant que nous avons formulé une approche théorique de la veille législative, que nous sommes informés du fonctionnement du système législatif européen ainsi que plus familiers avec certaines technologies, nous pouvons nous intéresser à la description du projet qui offrira une approche concrète à notre question de fond.

## Chapitre 2 : Description du Projet

---

### 1. Constats de départ

La volonté de Sopra Steria de développer ce système de veille est motivée principalement par trois constats de départ, qui sont selon elle trois aspects critiques à améliorer. Elle compte ainsi sur *EU Monitoring* pour tenter de compenser ces difficultés actuelles.

#### 1.1. Retard dans la prise de connaissance

L'activité de Sopra Steria au Benelux étant un conseil relativement neuf et moins développé que pour certains de ses concurrents, il est doté de moins de ressources. Il n'est pas rare que les consultants doivent agir en réaction aux événements qui se produisent, et dans le cas qui nous concerne, la mise en place de législations européennes.

En effet, dans le processus décisionnel de l'Union européenne, il arrive que ce ne soit qu'une fois la législation adoptée que Sopra Steria prenne connaissance de celle-ci. Alors seulement, ils cherchent à être pleinement informés de la marche à suivre et à être opérationnels aussi bien dans l'approche conseil que dans le développement de systèmes.

C'est un clair désavantage compétitif, car les concurrents ont pour leur part déjà acquis des connaissances, ont eu le temps de mettre en place une stratégie et sont prêts à offrir leurs services au moment de la sortie de la législation. Il est donc nécessaire de trouver un moyen de se mettre au courant le plus vite possible des potentielles législations européennes en devenir.

#### 1.2. Besoin de reconnaissance sur le marché ESN

L'agence *Consulting* de Sopra Steria Benelux veut se positionner comme étant un acteur de conseil pertinent dans le secteur des entreprises de services du numérique (ESN). Sopra Steria étant actuellement prioritairement considérée comme une entreprise de développement, ses relations avec ses clients passent essentiellement par les techniciens. Il lui faut donc une stratégie qui lui permette de se rendre crédible également au niveau des preneurs de décisions.

#### 1.3. Mise à profit des compétences existantes

Au travers des nombreux contrats que l'entreprise a exécutés ou est en train d'exécuter pour des institutions ou agences européennes, elle a acquis de larges compétences à propos des systèmes d'information (SI) gérant les frontières, l'immigration ou encore la sécurité intérieure. Étant donné que les états membres doivent se conformer en vue de pouvoir prendre part à la gestion et à l'alimentation en données de ces SI, ils ont besoin d'être guidés dans ce processus. Dès lors, il est primordial que Sopra Steria mette à profit cette



connaissance dans ces domaines afin de figurer comme compétent dans l'accompagnement des états membres. Ce processus s'appelle le « *cross-selling* ».

Cela rejoint en quelque sorte le point précédent, car si la communication sur ses compétences est suffisante, Sopra Steria apparaîtra encore plus fiable et crédible.

## 2. Objectif principal

Bien que l'introduction et la contextualisation du projet apportent déjà des éléments permettant d'identifier l'objectif du système, il est important de formuler un objectif principal afin de définir clairement le point de départ du projet.

L'objectif principal est donc, comme abordé au chapitre précédent, de créer un nouveau service qui s'appellera « *EU Monitoring* » et qui consistera en de la veille législative.

Pour rappel, la veille législative est un ensemble d'activités qui visent à analyser le développement des législations et de souligner les secteurs où il est possible de s'améliorer. Ce processus permet aux acteurs intéressés d'examiner systématiquement les processus de création et d'implémentation des réglementations.

La veille législative peut être appliquée à toutes sortes de secteurs et celui qui nous intéresse pour ce projet concerne la législation des systèmes d'information européens. En d'autres mots, cela implique de se tenir au courant de tout ce qui se déroule en termes de discussions, de législations, d'études d'impacts et autres, au niveau des pouvoirs décisionnels européens.

L'objectif de l'entreprise est d'être capable de se préparer en amont, dès que les discussions débutent à la Commission Européenne et de suivre ensuite l'évolution jusqu'à la mise en application des réglementations.

Dans un premier temps, le service ne serait accessible qu'au personnel de Sopra Steria pour éprouver le concept et pas encore à des personnes ou organisations tierces.

Dans un second temps, le service sera ouvert à la clientèle de l'entreprise. *EU Monitoring* sera un service d'information qui permettra à ses utilisateurs d'être à jour avec les décisions et discussions européennes et ainsi se conformer plus facilement.

Bien que nous distinguons deux temps de déploiement distincts, il faut prendre en considération toutes les exigences potentielles que nous pourrions rencontrer dans le futur. Ainsi, le système sera évolutif et permettra un déploiement plus large.

### 3. Nature de l'existant

Ce projet ne prend racine que sur une idée et une volonté claire de l'agence Consulting, tout le reste est à faire. La direction à prendre n'est donc pas facile à déterminer, mais la latitude laissée pour les différents choix à faire et qui nous incombent, rend ce projet d'autant plus intéressant.

### 4. Acteurs du projet

Le Tableau 1 ci-dessous reprend les acteurs majeurs opérant sur le projet.

Tableau 1 : Rôles et descriptions des acteurs du projet

Rôle	Description	Entité ou personne(s) concernée(s)
Co-sponsor	<p><u>Besoin</u> : cfr ci-avant, « constats de départ »</p> <p><u>Responsabilités</u> :</p> <ul style="list-style-type: none"> <li>Mettre à disposition les ressources nécessaires pour la réalisation du projet.</li> <li>Apporter un cadre.</li> </ul> <p><u>Autorité</u> : Nicolas Courtin - Directeur du Conseil</p>	Sopra Steria Consulting Benelux
Co-sponsor	<p><u>Besoin</u> : Exploration d'opportunités</p> <p><u>Responsabilités</u> :</p> <ul style="list-style-type: none"> <li>Donner un retour sur le travail effectué.</li> <li>Apporter une expertise.</li> </ul> <p><u>Autorité</u> : Didier Gilbert - Directeur de l'agence <i>HLS, Defense &amp; Aeroline</i></p>	Agence HLS, Defense & Aeroline
Chef du Projet	<p><u>Compétences</u> : Expertise en consultance, très bonnes connaissances des institutions européennes et des systèmes d'information.</p> <p><u>Responsabilités</u> :</p> <ul style="list-style-type: none"> <li>Définitions du but et du cadre du projet</li> <li>Apport d'un feedback régulier</li> <li>Suivi de l'avancement du projet</li> </ul> <p><u>Autorité</u> : Nicolas Courtin - Directeur du Conseil</p>	Cédric Genin (Head of Consulting Public Sector)
Gestionnaire du projet	<p><u>Compétences</u> : Connaissance des systèmes d'information et esprit d'analyse</p> <p><u>Responsabilités</u> :</p> <ul style="list-style-type: none"> <li>Définition des besoins d'<i>EU Monitoring</i></li> <li>Définition d'une méthodologie de projet en collaboration avec le chef de projet</li> <li>Première phase de développement d'<i>EU Monitoring</i></li> </ul>	Charles de Briey (Etudiant stagiaire)

	<ul style="list-style-type: none"> <li>• Complétion du registre d'<i>EU Monitoring</i></li> <li>• Analyse des pistes de valorisation du projet à plus long terme</li> </ul> <p><u>Autorité</u> : Cédric Genin (Head of Consulting Public Sector)</p>	
Collaborateurs du projet	<p><u>Compétences</u> : Expertise en consultance et connaissance des sujets de l'agence <i>HLS, Defense &amp; Aeroline</i></p> <p><u>Responsabilités</u> :</p> <ul style="list-style-type: none"> <li>• Soutien de l'étudiant</li> <li>• Apport d'idées</li> <li>• Aide à la production de livrables</li> </ul> <p><u>Autorité</u> : Cédric Genin (Head of Consulting Public Sector)</p>	<p>Julien Heintz (Business Consultant Senior)</p> <p>Florian Kariger (Business Consultant Junior)</p>

Les acteurs nommés dans ce tableau ne sont que les personnes qui sont amenées à travailler régulièrement pour *EU Monitoring*. Cependant, grâce à la culture d'entreprise, il n'est pas rare de pouvoir demander un support à des collaborateurs étrangers au projet de manière inopinée et, malgré tout, recevoir ce support.

## 5. Analyse ADEPT

Maintenant que nous avons une idée d'où nous nous situons avant le début de ce projet, nous pouvons identifier les besoins du projet. Afin de déterminer ces besoins, nous allons utiliser l'analyse ADEPT. Le cadre d'analyse ADEPT est un cadre simple, de haut en bas, utilisé pour aider les analystes à comprendre le système sous différents angles. Elle comprend cinq aspects que sont Activités, Données, Environnement, Personnes et Technologies. Ces perspectives guident l'analyse en découvrant les exigences fonctionnelles ou non fonctionnelles du système qui apportent de la valeur à l'organisation. À la suite de cette analyse, nous devrions disposer d'un *data model*, d'un diagramme de type « *use case* » et d'une architecture technique. (Sombach, 2019)

### 5.1. Activités

Les activités sont les principales étapes de la mise en œuvre de *business process* qui créent de la valeur pour les parties prenantes internes et externes. (Sombach, 2019)

Il faut désormais définir le *business process* d'*EU Monitoring*.

Tout d'abord, le processus de veille législative comme nous la concevons, et comme l'a montré l'approche théorique du premier chapitre, commence par une récolte d'informations. Il faut pouvoir avoir accès à suffisamment d'informations de qualité et pertinentes.

Une fois que l'information est en notre possession, il faut savoir si celle-ci est pertinente ou non. Cet aspect du *business process* n'est pas le plus compliqué, mais est tout de même très important car il faut arrêter le processus dans le cas où la pertinence ne serait pas avérée.

Lorsque l'information est jugée pertinente, il est temps de la résumer et c'est l'étape qui prend le plus de temps. En effet, pour résumer un texte, une vidéo ou autre, il faut prendre le temps de ne rater aucun élément important.

Ensuite, le champ d'application de l'information sera clairement défini afin de savoir par la suite comment stocker le résumé. Il va de soi que pour retrouver un résumé, il faudra qu'il soit bien référencé et pour ce faire, le champ d'application devra être correctement défini.

Finalement, deux étapes sont à exécuter en même temps. D'une part le stockage de l'information et d'autre part, le partage de l'information. Tout ce procédé est schématisé à la Figure 14.

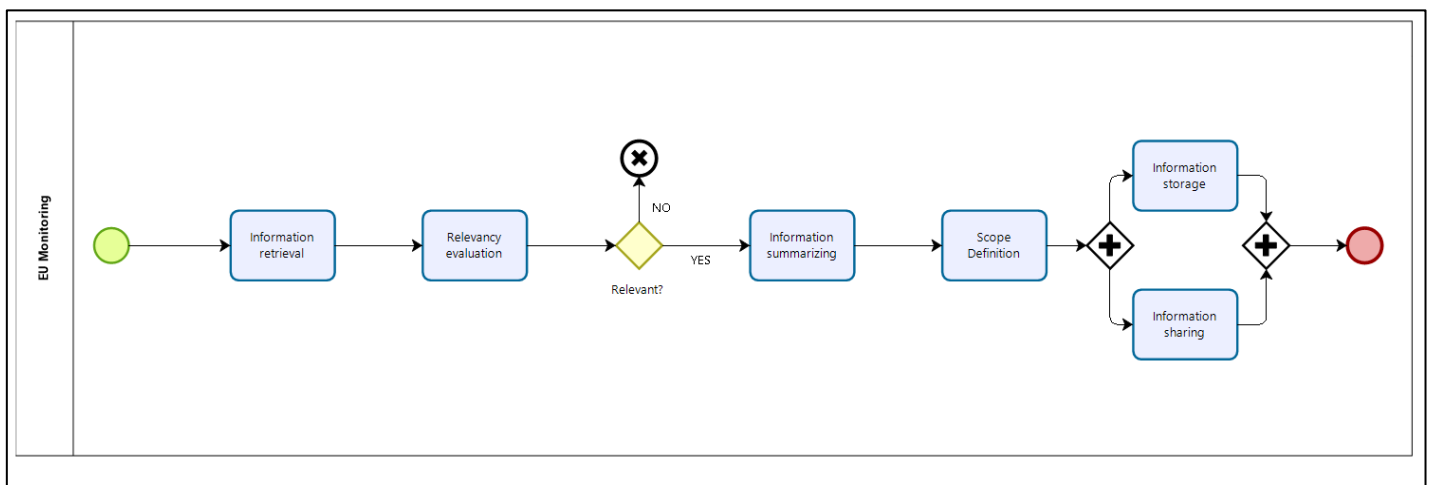


Figure 14 : Business process d'EU Monitoring

Une approche que nous voulons adopter est de viser le moins d'action humaine et d'automatiser au maximum chaque étape.

Dans ce processus de travail, on peut distinguer deux étapes supplémentaires au cycle que nous avons formulé lors de l'approche théorique. En effet, l'évaluation de la pertinence et la définition du champ d'application font également partie du procédé car ces étapes sont utiles à sa réalisation.

Par ailleurs, deux étapes du cycle théorique ne sont pas reprises dans le *business process*, à savoir la définition des besoins et le suivi des projets de loi. Au moment de la réalisation du *business process*, il ne semblait pas utile de définir systématiquement les besoins. En effet, *EU Monitoring* sera, dans un premier temps, un service généraliste et non personnalisé. Quant au suivi des projets de loi, il est réalisé mais pas non plus de manière automatisée. Nous nous y intéresserons lors de l'exploration de pistes d'amélioration.

## 5.2. Données

Les données sont tous les faits, images et sons qui peuvent ou non être pertinents pour un *business process* en particulier. Il existe quatre formes majeures de données (Sombach, 2019).

- Les données concrètes : données clairement définies générées par des systèmes formels (c'est-à-dire des connaissances explicites) ;
- Les données douces : données intuitives ou subjectives obtenues par des moyens informels (c'est-à-dire connaissance tacite, parler à quelqu'un, opinions) ;
- L'information : donnée dont la forme et le contenu sont adaptés à une tâche spécifique (c'est-à-dire des connaissances implicites, non écrites) ;
- La connaissance - combinaison d'instincts, de règles, d'idées, qui déterminent les actions que nous entreprenons et les décisions que nous prenons.

Les données que nous partageons et stockons dans le cadre d'*EU Monitoring* sont des données concrètes. En effet, elles sont clairement définies et générées par un système formel qu'est *EU Monitoring*. On cherche des données concrètes que l'on transmet via ce système sous une forme potentiellement différente, mais toujours de manière explicite.

Cependant, toutes ces données concrètes centralisées par *EU Monitoring* forment ensemble de l'information. Afin de garder l'information centralisée, il a été décidé de créer un recueil qui prendra la forme d'une plateforme.

Cependant, il se peut que les données transmises n'aient pas toutes une source concrète, mais qu'elles soient le résultat de la compilation de l'information et de la connaissance présentes dans l'entreprise. Comme évoqué plus en avant, il serait avantageux que les compétences acquises lors des contrats servent à l'alimentation d'*EU Monitoring*.

Les données se regrouperont en deux classes principales, la hiérarchie étant reprise à la Figure 15.

Il y a d'abord les articles « contenu », où chacun est défini par un numéro unique, un titre, une source, une date de création, un auteur et une indication de limitation (limitée aux employés de Sopra Steria), qui par défaut n'est pas limitée.

Ensuite, il y a les « utilisateurs », tous définis au moins par un numéro unique, une date d'activation du compte, une adresse e-mail, un nom et un prénom. Cette classe se subdivise en trois sous-classes. Premièrement la sous-classe « clients » de Sopra Steria, qui renseigne l'organisation qui emploie l'utilisateur. Deuxièmement, la sous-classe des employés de Sopra Steria (« collaborateurs »), qui doit renseigner l'agence pour laquelle le collaborateur travaille. Troisièmement, la sous-classe des administrateurs de la plateforme d'*EU Monitoring*, qui créent les comptes d'utilisateurs et le contenu. Il n'est pas possible d'avoir un utilisateur qui n'est pas compris dans l'une de ces trois sous-classes.

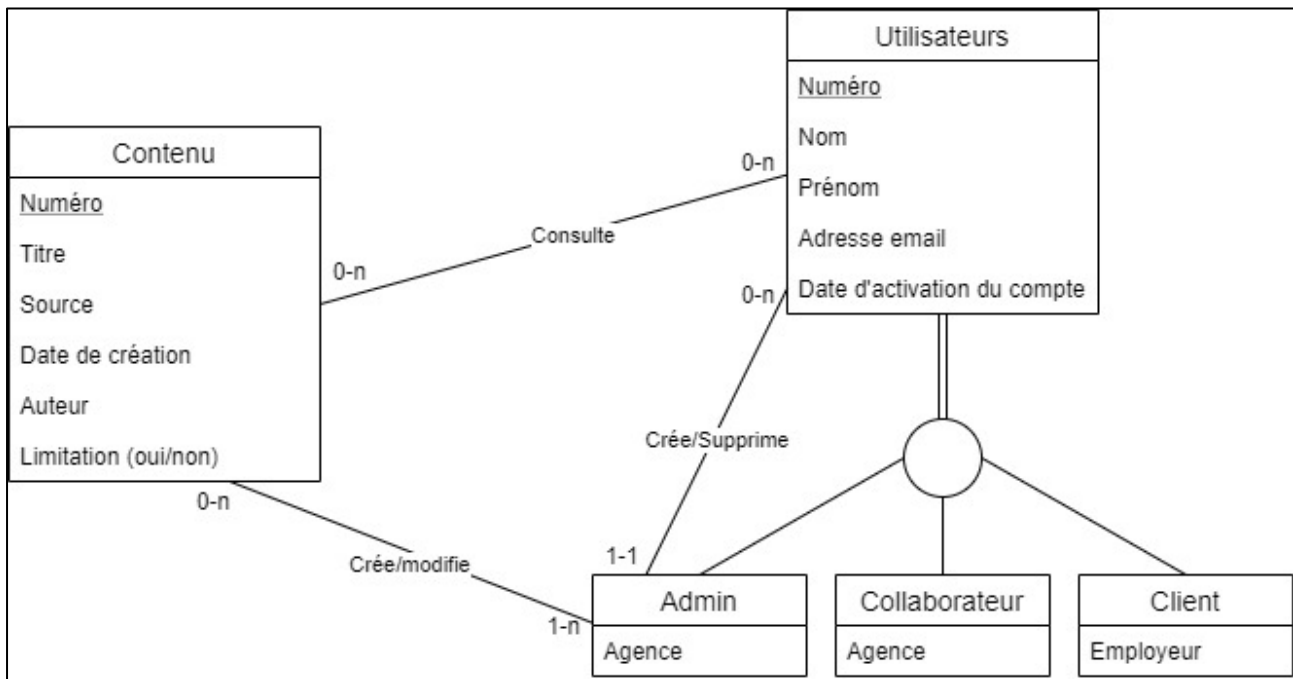


Figure 15 : Data Model d'EU Monitoring

Un élément du contenu peut être créé et modifié par un ou plusieurs administrateurs qui eux peuvent créer de zéro à plusieurs éléments. Un utilisateur est créé et supprimé par un seul administrateur et celui-ci peut en créer de zéro à plusieurs également. Les utilisateurs peuvent consulter de zéro à plusieurs éléments de contenu qui eux peuvent être consultés de zéro à plusieurs fois.

### 5.3. Environnement

La combinaison de facteurs internes et externes qui influencent les opérations d'une organisation forme l'environnement. Cela prend en compte aussi bien les clients que les concurrents (Sombach, 2019).

#### 5.3.1. Environnement interne

L'entreprise déjà décrite au début de ce document compose l'environnement interne au projet. Sopra Steria est divisée en agences dont celles déjà mentionnées, l'agence *HLS, Defense & Aeroline* et l'agence *Consulting*. Alors que la première est propre à un secteur d'activité, la seconde est une agence transversale. Cela signifie qu'elle sert de support à toute autre agence dont les contrats nécessiteraient du conseil.

Comme déjà énoncé, *EU Monitoring* sera un service de l'agence *Consulting*. Toutefois, l'accent sera porté sur les sujets pertinents pour l'agence *HLS, Defense & Aeroline*.

S'informer aussitôt que les discussions précèdent une proposition de réglementation dans la procédure législative ordinaire laisse plus de temps à ces deux agences d'appréhender les tenants et aboutissants de potentielles futures législations. Ainsi, on passe d'une approche réactive à une approche proactive. Cela répond au premier constat de départ. Il s'agit aussi d'explorer les nouvelles opportunités qui s'offrent à Sopra Steria.

### 5.3.2. Environnement externe

Les concurrents principaux de l'entreprise ont été cités dans la présentation de l'entreprise au début de ce document. Lors des recherches menées pour déterminer lesquelles de ces entreprises proposeraient un service comparable, nous n'avons rien trouvé. Cela ne signifie pas pour autant qu'il n'existe pas déjà de solution de veille législative équivalente à *EU Monitoring*, mais aucune de ces solutions déjà sur le marché n'a le même objectif de veille des systèmes d'information européens. Elles servent à des secteurs variés pouvant être le droit du travail ou du lobbying européen. Ci-dessous sont repris les résultats de ces recherches.

#### i. Solutions équivalentes

Il y a d'une part des entreprises de lobbying auprès des institutions européennes, comme **Athenora** ou **AlienorEU**. Leurs sites internet mentionnent qu'elles pratiquent des veilles législatives correspondant aux besoins spécifiques de leurs clients et que cela fait partie de leurs services de lobbyistes. Cependant, ils ne donnent pas plus de détails quant à leurs manières de faire.

**DR2 Consultants** quant à eux, peuvent se comparer aux entreprises Athenora et Alienor, car leur business model s'apparente à du lobbying, mais dans leur cas, la veille réglementaire est un service à part entière. C'est également ceux dont la portée de la veille se rapproche le plus de celle que nous voulons mettre en place au vu de l'impact européen de leur offre. Leur site précise qu'ils fournissent des mises à jour des derniers développements des politiques pertinentes. Cela implique une newsletter plus ou moins ciblée.

D'autre part, certaines entreprises, dont la veille fait partie intégrante de leur méthode de travail, expliquent davantage leurs offres. **FLA-Consultants** est par exemple une entreprise spécialisée en veilles de tous types dont, entre autres, en veille juridique. Cela comprend une définition du cadre légal d'un client pour ensuite être capable de bien définir les besoins en veille juridique de ce client et finalement, la mise en place d'un plan de compliance.

**Bourmendil & Consultants** se spécialisent quant à eux dans le droit de l'emploi en France et proposent une veille juridique dans ce domaine. Ils évoquent des « Experts en ligne » qui traitent les questions adressées. Parallèlement, ils publient une newsletter générale et proposent une newsletter ciblée.

**Valauris Consulting**, cabinet de conseil avec services de veille réglementaire, propose trois différentes offres plus ou moins développées selon les besoins de leurs clients.

Le Tableau 2 ci-dessous reprend les points majeurs de chacune de ces veilles législatives mentionnées.

Tableau 2 : Caractéristiques des veilles législatives équivalentes

	Définition du cadre légal	Newsletter ciblée	Newsletter	Accès à un recueil (plateforme)	Spécificité
<b>AlienorEU</b>			X	X	
<b>Athenora Consulting</b>	X	X			Information sur les personnes nommées aux postes clés
<b>DR2 Consultants</b>	X	X	X		Statuts à jour des procédures législatives
<b>FLA-Consultants</b>	X	X			Délivrable sur mesure pour le support opérationnel
<b>Bourmendil &amp; Consultants</b>	X	X	X		Legal Hotline : questions et réponses en 72h
<b>Valauris Consulting</b>	X	X	X	X	Différents packages + Planning de mise en conformité

Légende :

<b>Veille législative à des fins de lobbying</b>	<b>Veille législative à des fins opérationnelles</b>
--	--

Les X désignent le fait qu'une veille offre le service. L'absence de X signifie qu'elle n'y répond pas ou que l'information disponible n'est pas suffisante pour l'affirmer.

Lorsque ce tableau a été réalisé, une distinction a été faite entre les veilles faites dans le cadre d'activités de lobbying et les autres qui visent le support opérationnel. Cette distinction est pertinente pour savoir si l'on désire offrir ce service afin d'aider les clients à se conformer ou à les informer des diverses positions prises au niveau européen.



Par ailleurs, à la vue de ce tableau, on constate immédiatement que deux points sont quasi inévitables. Le premier est la définition d'un cadre législatif sur mesure et le second, une newsletter ciblée. Autant le premier point mentionné demande plus de main-d'œuvre, autant le second est tout à fait faisable dans le cadre de cette gestion de projet.

## ii. Potentiels clients

Pour rappel, lorsque qu'un SI entre en opération au niveau central, c'est-à-dire au niveau européen, les autorités des états membres doivent faire en sorte de mettre en place des SI nationaux et qu'ils soient compatibles avec le SI central. Par exemple, avec la solution SIB (2.4.2.ii), Sopra Steria est très compétente pour aider un pays à se connecter au système central SIS (Schengen Information System). Cependant, cela ne reste principalement que du développement, de l'installation, de la maintenance et du support. Ces autorités nationales ont également des besoins de conseil pour leur gestion quotidienne.

Dès lors, les clients potentiels sont les preneurs de décisions au sein de ministères, de services de contrôles aux frontières ou de forces de l'ordre. Ils doivent en effet être informés des implications que ces législations pourraient avoir sur leur travail. Comme cela, Sopra Steria apparaîtra non seulement comme une entreprise de développement, mais également comme un partenaire capable de soutenir les besoins d'acquisition d'information et de prise de décisions. *EU Monitoring* peut servir de service « d'approche » pouvant entraîner le désir d'un service plus personnalisé. Cela aborde le second constat de départ, qui était le besoin de reconnaissance sur le marché ESN.

Dans le même temps, *EU Monitoring* pourrait appliquer le *cross-selling* abordé ci-avant. En plus des informations récoltées à propos des législations en devenir, *EU Monitoring* servirait à partager les connaissances acquises tout au long des contrats avec l'UE. Cela aborde le troisième constat de départ, à savoir la mise à profit des compétences existantes.

Des contacts existent déjà avec de potentiels utilisateurs d'*EU Monitoring*. Par exemple *DGNum*, direction générale du numérique et des systèmes d'information et de communication, qui vise à assurer la cohérence des SI au sein du ministère français des Armées, a déjà exprimé son intérêt pour un tel service.

## 5.4. Personnes

Les personnes sont celles qui entrent (créent), traitent ou utilisent l'information produite dans le système. (Sombach, 2019)

Si l'on considère le tableau reprenant les acteurs du projet (Figure 16), les personnes qui entrent et traitent l'information dans le système sont les stagiaires successifs qui travaillent sur le projet. Les collaborateurs du projet peuvent également être amenés à éditer du contenu pour *EU Monitoring*. Ces membres de l'équipe *EU Monitoring* sont responsables de la gestion des utilisateurs. C'est-à-dire que ce sont eux qui doivent ajouter et retirer les utilisateurs. Afin d'évaluer le taux d'utilisation de la plateforme, ces membres doivent pouvoir

consulter les statistiques de visites. La plateforme qu'il a été décidé de développer doit donc donner la possibilité à ces différents acteurs d'exécuter ces tâches.

Quant à l'utilisation, il y a d'une part les employés de Sopra Steria ayant besoin de cette information et d'autre part les clients. Il est possible que le contenu stocké sur *EU Monitoring* ne puisse pas être accessible par des personnes externes à l'entreprise pour diverses raisons. C'est pourquoi le système doit permettre de créer plusieurs groupes de privilèges.

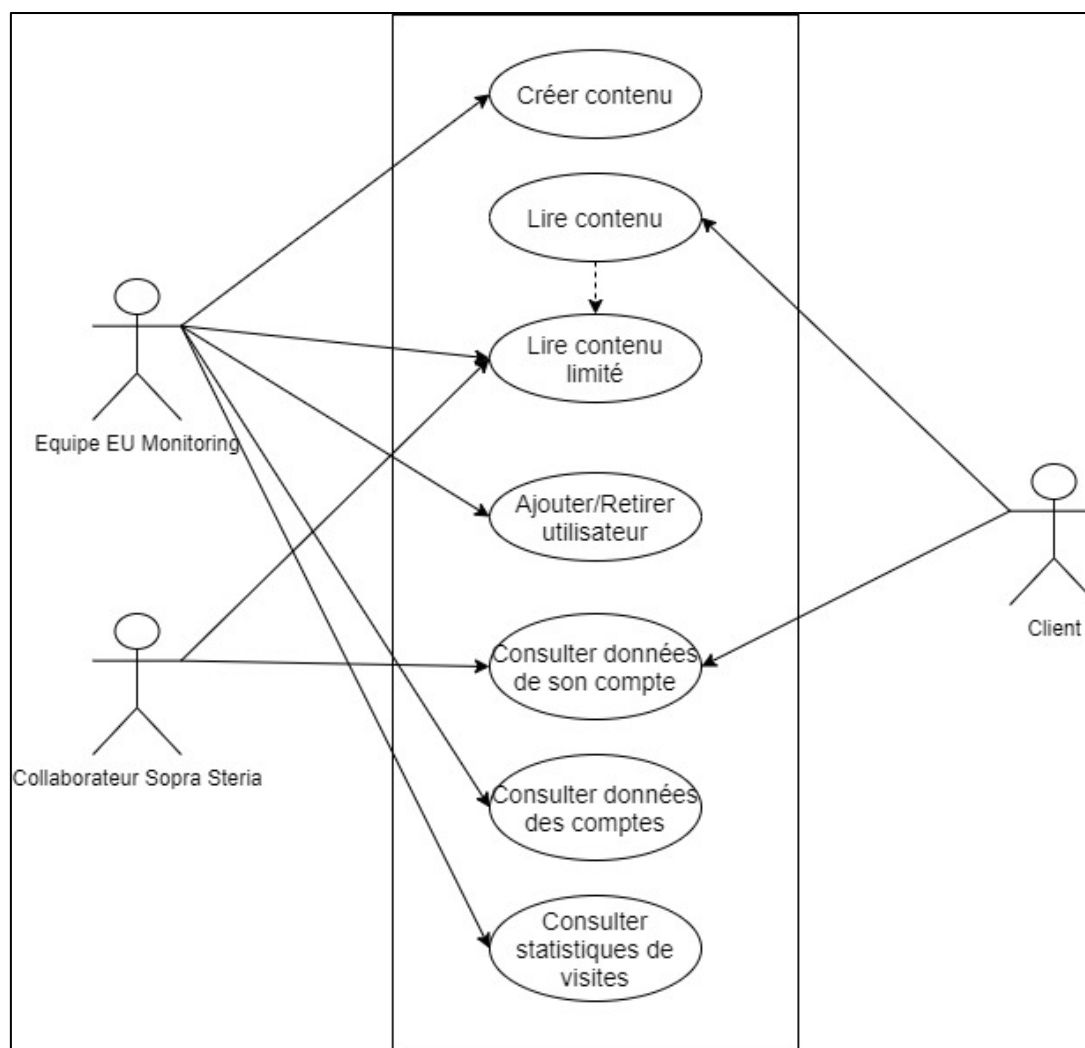


Figure 16 : Use case d'EU Monitoring

## 5.5. Technologie

La technologie se compose des logiciels, du matériel et des autres outils et équipements utilisés par les personnes dans l'accomplissement de leurs tâches. La technologie est considérée comme faisant partie de l'infrastructure si elle est partagée entre plusieurs systèmes.

Il y a deux types d'infrastructures. Premièrement, il y a le réseau qui constitue l'infrastructure technique. Deuxièmement, il y a l'infrastructure de l'information qui détermine comment l'information est partagée au sein de l'entreprise. (Sombach, 2019)

*EU Monitoring* doit être accessible sur internet selon les protocoles sécurisés (HTTPS) et avoir accès à du stockage pour y entreposer sa base de données. Le système doit rendre possible l'envoi d'e-mail directement depuis la plateforme. Cela sera possible grâce au protocole SMTP.

Notons que cette phase de développement d'*EU Monitoring* est surtout une preuve de concept. Par conséquent, la plateforme n'est pas hébergée directement sur internet. On héberge la plateforme sur une machine installée dans les bureaux de Sopra Steria. Cette machine permet d'avoir accès à la plateforme sur n'importe quel ordinateur connecté au réseau local virtuel (VLAN) de Sopra Steria.

Lorsque que l'on souhaite y accéder depuis un lieu éloigné, en dehors des installations de l'entreprise, on peut utiliser le *virtual private network* (VPN) de l'entreprise. Pour des raisons de facilité et de compétence, le système d'exploitation de la machine hôte est Microsoft Windows.

La Figure 17 montre à quoi devrait ressembler l'architecture d'*EU Monitoring*.

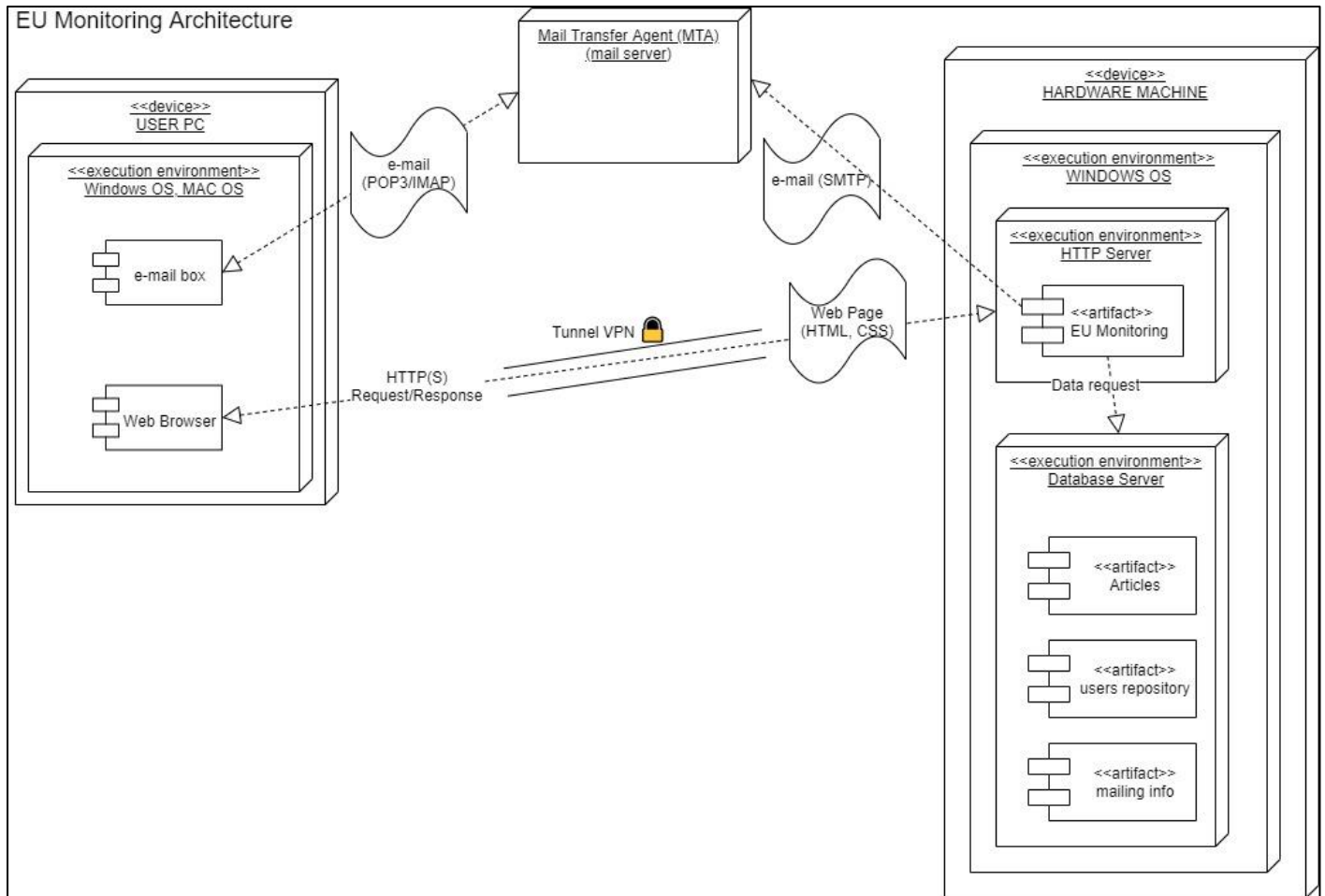


Figure 17 : Architecture d'EU Monitoring

Cependant, lorsqu'il faudra migrer le système vers des infrastructures permanentes, il faudra aussi avoir recours à des solutions de back-up afin de ne pas perdre les données. Actuellement, ce back-up doit être fait manuellement. Ensuite, pour assurer la disponibilité de la plateforme, il faudra prévoir également des sources d'électricité redondantes et indépendantes du système principal. Tout cela doit faire partie d'un plan de continuité d'activité.

## 6. Objectifs intermédiaires

Maintenant que nous connaissons les grandes lignes des exigences du système, nous pouvons déterminer les objectifs intermédiaires qui serviront d'étapes lors de l'exécution du projet. La série d'objectifs présentée ci-après et résumée dans le Tableau 3, reprend les éléments identifiés comme points importants lors de l'approche théorique ainsi que certains points que nous avons considérés comme nécessaires pour l'exécution de notre système de veille législative.

### Objectif 1 : Trouver suffisamment d'information

Ce premier objectif consiste à mettre en place des outils afin de trouver suffisamment d'information à traiter et à partager à travers *EU Monitoring*. Idéalement, il faudrait pouvoir accéder à une quantité d'informations permettant d'occuper deux jours pleins par semaine.

### Objectif 2 : Mettre en place une plateforme pour le partage et le stockage des données

Cette étape est sans doute celle qui demandera le plus d'effort et de temps. La plateforme devra respecter plusieurs exigences fonctionnelles et non fonctionnelles, notamment décrites lors de l'analyse ADEPT.

### Objectif 3 : Mettre en place des outils pour simplifier le *Business Process*

La volonté est que l'apport de travail humain soit le moins important possible. Pour cela, il faut pouvoir se reposer sur des outils afin de simplifier le *Business Process* au maximum. Idéalement, il faudrait avoir un outil pour chaque étape.

### Objectif 4 : Établir une manière standardisée de rapporter les résultats

Les données stockées sur la plateforme devront toutes avoir la même forme, comprenant des éléments comme la source ou la date de publication.

### Objectif 5 : Production de contenu d'*EU Monitoring*

Cet objectif est transversal aux autres dans le sens qu'il n'a pas de fin et doit se faire en parallèle des autres étapes. Afin de pouvoir observer si les autres éléments du *business process* mis en place sont adéquats, il faut bien entendu pouvoir les éprouver avec du contenu. À la fin de la période de stage, il faudrait qu'en moyenne, six documents (articles, communications de l'UE, texte législatifs ou vidéo) soient résumés par semaine.

### Objectif 6 : Valoriser le contenu d'*EU Monitoring*

La valorisation du contenu se fera à travers l'envoi de newsletters. Idéalement, il faudrait pouvoir envoyer une newsletter hebdomadaire contenant les six documents produits par semaine.

## Objectif 7 : Harmoniser les outils du *business process*

Au fur et à mesure de l'exécution du projet, différents outils seront ajoutés au *business process*. À la fin du stage, tous les outils doivent être intégrés clairement et harmonieusement. Ainsi, la reprise du service *EU Monitoring* par un futur stagiaire sera la plus facile possible.

## Objectif 8 : Identifier les besoins à court terme pour la finition du projet

Il est possible que le projet ne soit pas complètement terminé à la fin de la période de stage. Il faudra alors identifier les éléments qu'il reste à mettre en place afin qu'*EU Monitoring* soit opérationnel.

Tableau 3 : Listes objectifs intermédiaires avec leurs mesures de réalisation

Objectif n°	Titre	Métrique ou livrable
1	Trouver suffisamment d'informations	Contenu suffisant pour 2 jours de travail complets
2	Mettre en place une plateforme pour le partage et le stockage des données	Plateforme opérationnelle
3	Mettre en place des outils pour simplifier le Business Process	1 outil/étape
4	Établir une manière standardisée de rapporter les résultats	Tous les documents ont la même forme
5	Production de contenu	6 résumés/semaine
6	Valoriser le contenu d' <i>EU Monitoring</i>	Mettre en place une newsletter
7	Harmoniser les outils du Business Process	Document reprenant les outils, et la façon de les utiliser
8	Identifier les besoins à court terme	Document de retour d'informations à l'attention de l'entreprise

## **7. Méthodologie de gestion de projet**

Avant de pouvoir exécuter le projet en tant que tel, nous avons besoin de définir une méthodologie de projet adéquate à ce type de projet.

Après avoir passé en revue de multiples méthodologies de gestion de projets, notre choix s'est porté sur l'*Adaptive Project Framework (APF)*. Cette méthode introduite par K.

Wysocki en 2008 repose sur l'hypothèse que tous les projets sont différents. (ThinkThyme, 2020)

L'APF est une méthode qui convient bien au développement d'*EU Monitoring* car, comme son nom l'indique, elle est adaptative. En d'autres mots, cela permet de gérer un projet en ayant un but clairement défini, mais des exigences relativement mouvantes, qui peuvent se définir en cours de projet. Cette méthode est également itérative. En cela, elle se rapproche d'une méthode de gestion de projet de type agile classique.

À la fin de chaque itération, les exigences sont sujettes à des modifications plus ou moins importantes. C'est le cas du projet qui nous occupe. En effet, notre but est défini et très peu sujet au changement, mais les exigences peuvent, quant à elles, évoluer. Par exemple, la récolte d'information a un objectif quantitatif, mais il n'y a pas d'exigence claire en début de projet quant à la marche à suivre pour atteindre l'objectif.

De plus, l'APF requiert un soutien et un suivi important du sponsor. Pour rappel, nous avons identifié deux sponsors principaux dans le cadre de ce projet : l'agence Consulting et l'agence HLS, Defence & Aeroline, toutes deux internes à Sopra Steria. Le fait que ces sponsors fassent partie de l'entreprise satisfait largement ce besoin de soutien et confirme un peu plus le bien-fondé de l'APF pour notre sujet. C'est sur cet aspect que l'APF se distingue des méthodes agiles classiques. Alors que l'expérience client n'est qu'un aspect parmi d'autres dans une méthode agile, l'APF repose essentiellement sur le retour du client sponsor et le met au cœur du développement du projet (ThinkThyme, 2020).

Concrètement, le projet est développé en itérations et respecte le principe de « *plan-do-check-act* » de la roue de Demming illustrée à la Figure 18 (Nolleveaux, 2018). Chaque itération dure une semaine. En chaque fin d'itération, un plan est dressé pour la semaine qui suit (*plan*). Lors de cette itération, le plan est suivi jusqu'au vendredi (*do*). Le vendredi de chaque semaine a lieu une réunion avec l'équipe du projet *EU Monitoring* afin de comparer ce qui a effectivement été fait par rapport à ce qui a été prévu (*check*). Si besoin, le projet peut être redirigé dans la bonne direction (*act*) et un plan est de nouveau dressé pour l'itération suivante. La roue de Demming a fait un tour.

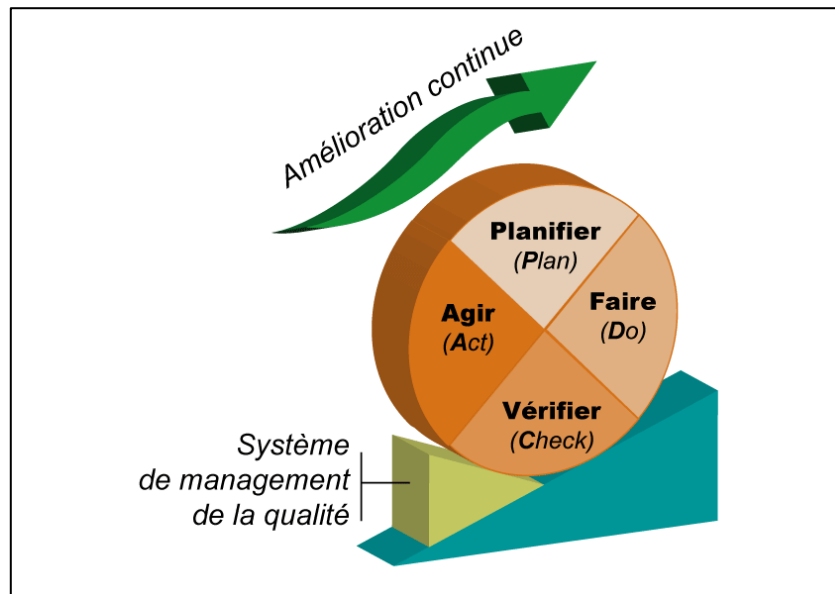


Figure 18 : Roue de Demming

Source : Nollevaux, G. (2019). Gestion de projets informatiques Gestion de projets digitaux (Notes de cours). Consulté à l'adresse [https://moodle.ichec.be/pluginfile.php/26769/mod\\_resource/content/22/ICHEC\\_Gestion\\_Projet\\_Informatique\\_Digit\\_2019\\_2020\\_v43.pdf](https://moodle.ichec.be/pluginfile.php/26769/mod_resource/content/22/ICHEC_Gestion_Projet_Informatique_Digit_2019_2020_v43.pdf)

La réunion du vendredi remplit plusieurs rôles. Tout d'abord, elle sert à la « gestion de la qualité » comme indiqué sur l'illustration de la roue de Demming. En effet, grâce à cela, nous sommes sûrs de garder un cap qui est le bon en vue d'atteindre l'objectif principal.

Ensuite, cette réunion permet de partager la connaissance acquise par la création de contenu. Rien ne sert de faire des recherches et de résumer des documents pour qu'ils restent uniquement sur une plateforme sans pouvoir être lus. En attendant d'avoir une manière efficace de partager ces éléments, un passage en revue des découvertes de la semaine est utile pour le partage de connaissance.



## Chapitre 3 : Mise en Œuvre du Projet

---

Nous avons à présent passé en revue toute une série d'éléments qui vont nous permettre de nous atteler à la mise en œuvre du projet. Dans ce chapitre, nous allons nous intéresser à chaque étape du *business process* et ce qui a été mis en place.

### 1. Récolte d'information

Une veille législative demande beaucoup d'informations, mais surtout de l'information pertinente. C'est pourquoi il est crucial de gérer les différents flots d'information, souvent importants, afin de perdre le moins de temps possible.

Cette récolte d'information se compose de trois sources principales ayant des niveaux d'efficacité variables.

#### 1.1. Boîte e-mail dédiée

Le premier système de récolte d'information mis en place est une boîte e-mail dédiée. Cette boîte a pour but de recevoir toute *newsletter* d'acteurs agissant dans le domaine visé par *EU Monitoring*. Le travail de configuration de la boîte n'est pas un travail fini, car il est perpétuel étant donné qu'il est toujours possible de trouver de nouvelles *newsletters* intéressantes.

De plus, après avoir abonné l'adresse e-mail à un nouveau contenu, il faut encore prendre le temps d'évaluer la qualité de celui-ci. Par exemple, la *newsletter* proposée par l'agence européenne Europol ne fait état que de ses opérations de police et pas, ou peu, de l'utilisation et de la gestion de données par divers systèmes. Il convient alors de se désabonner d'une telle *newsletter*.

L'efficacité de la boîte e-mail dédiée est discutable. Comme exposé au paragraphe précédent, elle a le défaut de fournir une trop grande proportion de contenu peu pertinent. Il est vrai que lorsqu'on souscrit à la *newsletter* du Parlement Européen par exemple, les sujets traités sont très variés et, bien qu'ayant fait un premier tri parmi nos préférences, la pertinence ne s'était que légèrement améliorée. Nous ne pouvons pas non plus trop circonscrire le flux d'information entrant, au risque de passer à côté d'éléments cruciaux. Par conséquent la consultation de la boîte e-mail peut parfois se révéler fortement consommatrice en temps pour n'avoir finalement que des résultats maigres.

#### 1.2. Agrégateur de flux RSS

Le second système de récolte d'information que nous utilisons est un agrégateur RSS (*Really Simple Syndication*).

Concrètement, en fonction des mises à jour d'un site internet, un contenu va être produit automatiquement, dont le format est un flux RSS. Il s'agit d'une convention de structuration garantissant que le flux de titres de pages web soit diffusé dans un format compréhensible par tous. Cette syndication du contenu permet aussi bien à un développeur web d'afficher le contenu des flux sur sa page web qu'à une personne lambda de recevoir ce même contenu dans sa boîte e-mail ou dans le cas qui nous occupe, un agrégateur.

Techniquement, un flux RSS est un fichier XML. À ce titre, il est composé de balises qui déterminent lorsqu'il s'agit d'un titre, d'un texte ou d'autre chose. Lorsque les pages web produisent des flux RSS, elles utilisent des scripts dans les langages tels que PHP ou encore Java. Les agrégateurs de flux RSS à l'inverse, doivent convertir ces fichiers XML en HTML (JDN, 2019).

L'agrégateur de flux RSS dont nous nous servons est **Feedly**. Cela nous permet de récolter de l'information provenant des divers sites internet des institutions et des agences de l'UE ou des sites d'information qui proposent ce service de flux RSS.

En plus de cela, Feedly permet de ne voir que le contenu qui semble le plus pertinent selon les dernières visites et le contenu précédemment consulté sous la rubrique « Today », comme sur la capture d'écran de la Figure 19.

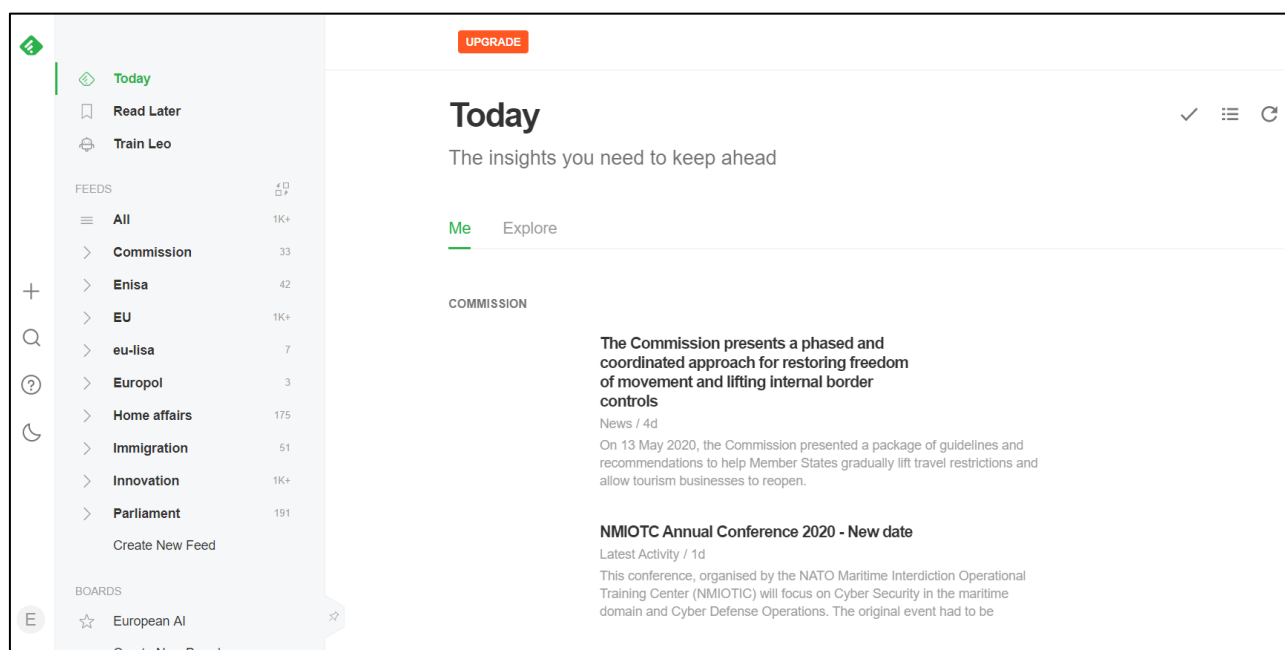


Figure 19 : Rubrique "Today" de Feedly

S'il ne semble pas que ce soit suffisant, il est possible d'utiliser la rubrique « All » dont la pertinence est moindre, mais le contenu beaucoup plus important. Dans ce cas, cela s'apparente à une boîte e-mail. La capture d'écran de la Figure 20 en est l'illustration.

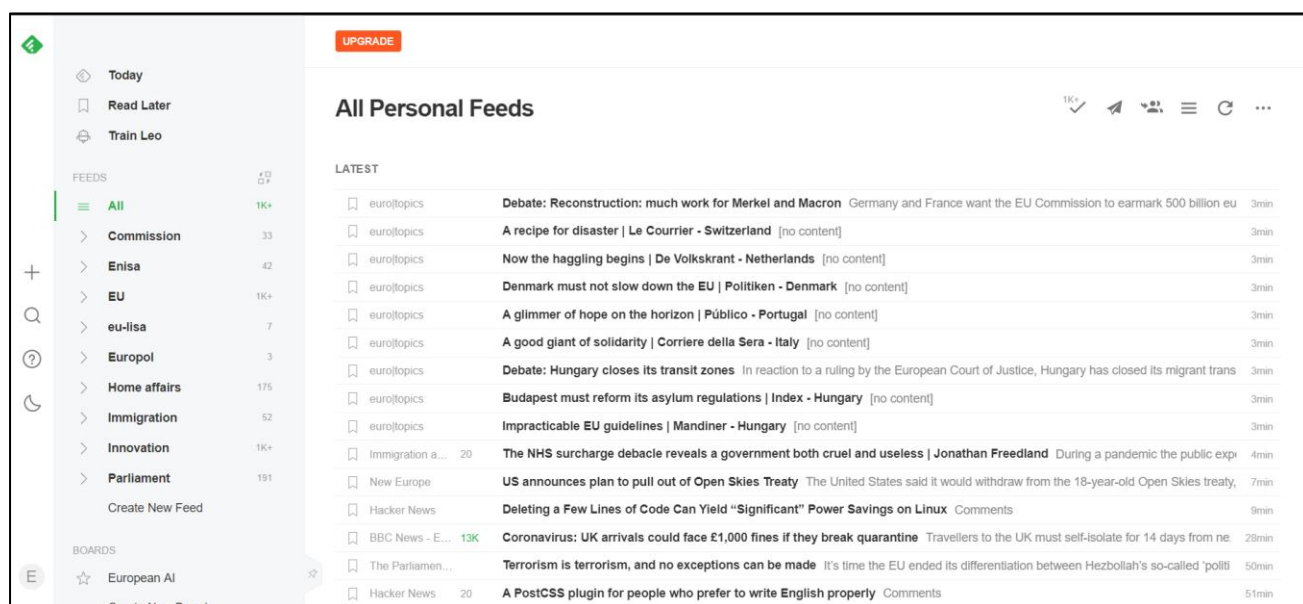


Figure 20 : Rubrique "All" de Feedly

### 1.3. Pipe de lectures

À plusieurs reprises au cours du projet, nous avons constaté que certains collaborateurs du projet recevaient des documents qui représentaient un intérêt certain pour *EU Monitoring*, mais qu'ils n'avaient pas toujours suffisamment de temps pour traiter ces informations. Il a donc été décidé de les partager dans un pipe de lectures. Ce pipe de lecture sert donc aux collaborateurs à entreposer des documents qu'ils n'ont pas le temps de lire. Ainsi, les documents peuvent être traités dans le cadre d'*EU Monitoring* et ainsi être valorisés. Lorsque que quelqu'un affiche un document dans le pipe de lecture, une personne travaillant pour *EU Monitoring* en prend la charge, le résume et le publie sur la plateforme comme si le document venait de Feedly ou de la boîte e-mail.

Sopra Steria utilise **Microsoft Teams** pour ses communications internes. Dans Microsoft Teams, nous avons un channel dédié à *EU Monitoring*. Dans ce channel, nous avons ajouté l'extension « *Planner* ». Cette extension nous permet de lister les éléments à analyser et résumer. À présent, grâce à cet outil, ce contenu n'est plus perdu.

L'application offre certaines fonctionnalités très intéressantes. Il permet par exemple de donner un degré d'importance à l'information ainsi qu'un responsable du traitement de l'information. Aussi, *Planner* permet de suivre l'évolution du traitement de l'information, autrement dit si le texte doit encore être traité, est en cours de traitement ou est traité comme l'illustre la capture d'écran de la Figure 21.

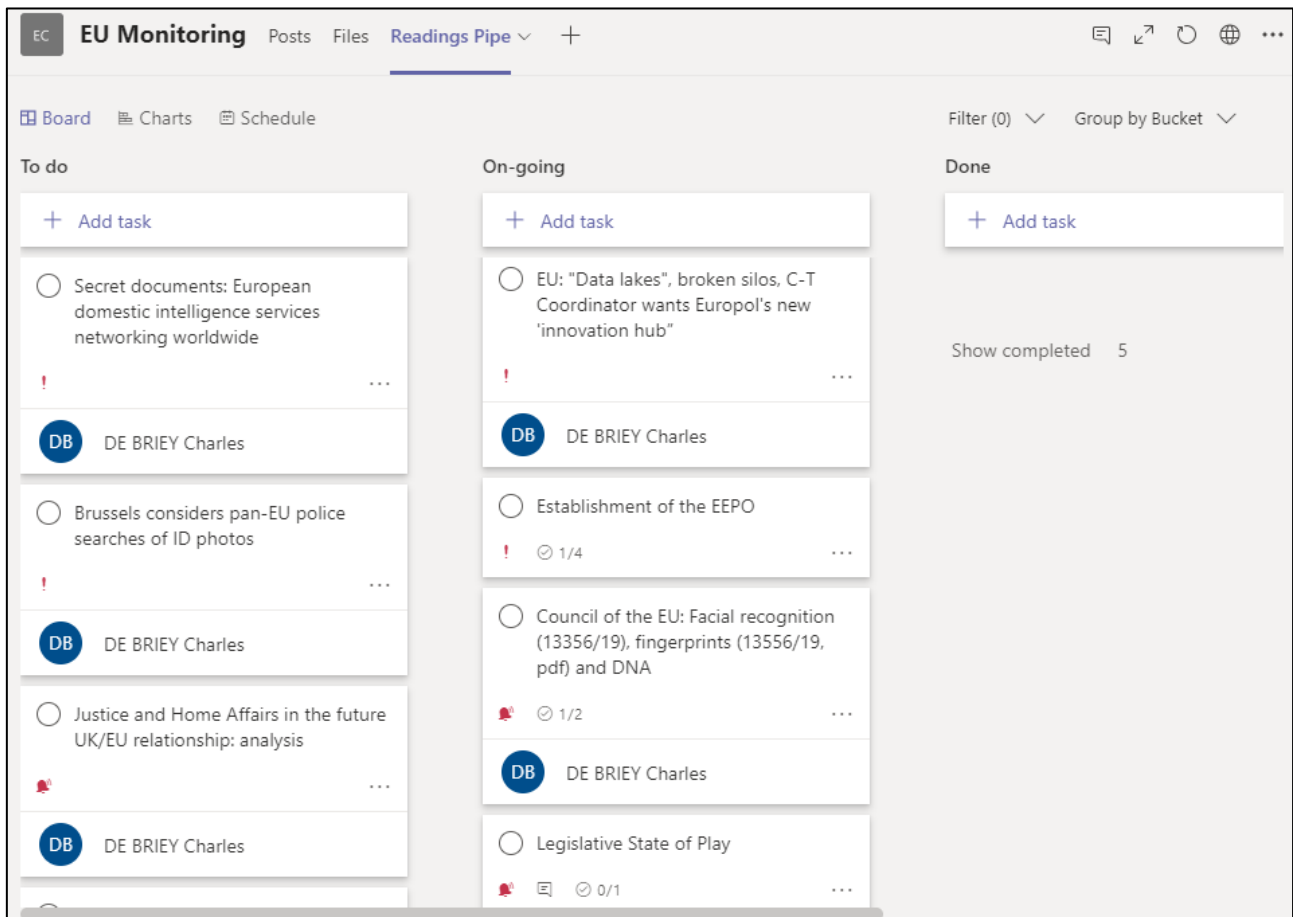


Figure 21 : Pipe de lecture

## 2. Évaluation de la pertinence

Nous avons désormais mis en place des moyens qui nous aident à nous procurer suffisamment de documents pour alimenter notre veille législative. Toutefois, bien que ces outils permettent de cibler grossièrement les sujets qui nous intéressent, un tri supplémentaire n'est pas superflu. À l'heure actuelle, ce tri doit encore se faire manuellement par un opérateur humain.

Pour toute personne initiée au projet, ce tri n'est pas compliqué. Nous avons malgré tout dressé une liste dans le Tableau 4 des sujets et de leur importance respective afin que toute nouvelle personne amenée à travailler sur le projet puisse faire les choix de documents les plus avisés possibles.

Tableau 4 : Degré d'importance des sujets à traiter

Haute importance							
Border	Schengen	Security	Migration	Defense	Brexit	Home Affairs	Justice
Importance moyenne							
Aviation	Customs	Transport	Digital Economy	Data Management	Law	Research & Devlpt (IT)	
Faible importance							
Banking	Antitrust	Trade	Cooperation & Development	Taxation	Competition	Social Affairs	Space

Nous avons trié ces sujets en trois catégories car nous avons ratissé large. Il est cependant très important de ne pas se contenter des sujets de haute importance. Comme la première étape du cycle de la veille législative de *Manager Go* exposée lors du premier chapitre le suggère, il faut pouvoir rassembler des informations plus largement que l'activité précise visée par la veille.

### 3. Résumer l'information

L'information récoltée peut parfois prendre la forme d'un long texte ou d'une longue vidéo dans lesquels ne se trouvent que quelques éléments intéressants à souligner. De plus, *EU Monitoring* se doit d'aller à l'essentiel pour que ses utilisateurs trouvent ce dont ils ont besoin dans un laps de temps le plus court possible. Il faut donc résumer ces documents.

La première façon de faire est forcément de tout lire ou de tout écouter et de retenir les points les plus importants. Ce travail est primordial car la crédibilité de la veille législative dépend du fait que l'information soit complète et correcte.

Une seconde façon de faire serait de s'aider d'une application qui résume les textes. Il en existe plusieurs et notre choix s'est porté sur **Resoomer**. Cette application fonctionne sur n'importe quel navigateur web et nous permet de résumer des textes beaucoup plus rapidement.

*Resoomer* fonctionne dès qu'un texte avec une bonne structure lui est présenté. En effet, le texte doit être structuré en présentant des balises claires. Ces balises peuvent être des connecteurs logiques (de plus, cependant, toutefois, etc), des dates ou encore des lieux. Les romans, les dialogues ou interviews sont dès lors inexploitable (Resoomer, 2020). Par contre, les articles de presse ainsi que les documents scientifiques et législatifs, auxquels nous avons le plus souvent recours, sont des textes structurés. *Resoomer* correspond dès lors très bien à nos besoins.

L'application peut s'utiliser de quatre manières différentes mais nous n'en retiendrons que trois.

Premièrement, *Resoomer* propose un résumé « automatique ». D'expérience, nous pouvons affirmer qu'une très grande majorité des éléments à souligner se retrouve dans ces résumés. C'est donc une très bonne base de travail. Considérons la Figure 22 reprenant le résumé fait de la présentation d'eu-LISA qui se trouve plus en avant de ce document.



Figure 22 : Fonction "Automatique" de *Resoomer*

Comme on peut l'observer, la longueur du résumé est d'environ un tiers du texte original et reprend des passages importants du texte. Cependant, il pourrait être nécessaire d'y ajouter quelques mots pour contextualiser davantage le résumé et malgré tout raccourcir largement l'original. On peut alors faire un copier-coller du résumé créé et y rajouter uniquement certains éléments mineurs.

Si toutefois, la version automatique ne convient pas au texte, car il a été trop abrégé, on peut utiliser l'onglet « Manuel » qui permet de choisir la proportion de texte à garder (Figure 23). Continuons avec l'exemple d'eu-LISA.

The screenshot displays the Resoomer web application interface. On the left, there is a text input area with a placeholder 'Uniquement textes argumentatifs'. Below it, a sample text is provided, discussing the European Agency for the operational management of information systems (eu-LISA). The text is partially highlighted in red. Above the input area are buttons for 'Exemple de texte', 'Effacer le texte', and 'Resoomer'. On the right, there is a control panel with a 'Résumé' section containing buttons for 'Automatique', 'Manuel' (which is selected), 'Optimisé', 'Analyse', and 'Aide'. Below these buttons, a slider allows users to choose the size of the summary, currently set at 60% (180 mots / 304). The main content area on the right shows the resulting summary of the input text, which is a condensed version of the original text.

Figure 23 : Fonction "Manuel" de *Resoomer*

Dans cet exemple, seulement un peu plus du tiers a été enlevé. On se rend vite compte que le résumé donne plus de détails et que les besoins de fournir une contextualisation est moindre.

Finalement, *Resoomer* peut aussi servir à souligner les phrases importantes du texte dans sa version « Analyse ». C'est sûrement la manière d'utiliser *Resoomer* qui prend le plus de temps mais aussi celle qui permet d'omettre le moins de points importants. Dans ce cas, le système met en évidence les phrases les plus importantes en gardant le reste du texte. Cela permet de pouvoir mettre en perspective plus facilement le contenu sélectionné. Cependant, la fonction « copier-coller » ne peut pas s'utiliser, étant donné que le choix de ce qui va se trouver dans le résumé doit se faire manuellement.

La façon d'utiliser *Resoomer* dépend de l'importance du document. S'il est de moindre importance, portant sur un sujet secondaire, le résumé automatique est suffisant. Toutefois, si le sujet revêt une haute importance comme des spécifications techniques d'ETIAS (présentée au point 2.4.2.) par exemple, il ne faut rater aucun point important et utiliser l'analyse. Il se peut aussi qu'un document soit suffisamment condensé pour qu'il ne nécessite aucun résumé.

### 3.1. Comment une application de résumés fonctionne-t-elle ?

Ce genre d'application est un bon exemple d'utilisation des principes de *Natural Language Processing* et de *Text Mining*, pour l'extraction d'information évoquée dans l'approche théorique.

En général, lorsqu'un humain résume un texte, il lit l'ensemble de la sélection pour bien la comprendre, puis rédige un résumé en soulignant les points principaux. Comme les ordinateurs n'ont pas encore les capacités linguistiques des humains, il faut envisager



d'autres méthodes. L'une des stratégies les plus utilisées par les outils de résumé de texte est l'extraction de phrases, qui se fait sur base de pondérations statistiques des phrases.

Par exemple, les outils de résumé peuvent extraire les phrases qui suivent une mention clé comme "en conclusion", après laquelle se trouvent généralement les points principaux du document. Les outils de résumé peuvent également rechercher les titres et autres marqueurs de sous-sujets afin d'identifier les points clés d'un document.

De nombreux outils de résumé de texte permettent à l'utilisateur de choisir le pourcentage du texte total qu'il souhaite extraire sous forme de résumé comme c'est le cas pour *Resoomer*. La fonction de résumé peut fonctionner avec des outils de suivi des sujets ou des outils de catégorisation afin de résumer les documents qui sont extraits sur un sujet particulier.

Un processus de résumé automatique peut être divisé en trois étapes :

- 1) L'étape de prétraitement qui donne une représentation structurée du texte original.
- 2) L'étape de traitement où un algorithme transforme la structure du texte en une structure de résumé.
- 3) L'étape de génération qui fournit le résumé final à partir de la structure de résumé.

Les méthodes de résumé peuvent être classées, en fonction du niveau linguistique, en deux grands groupes.

Premièrement, les approches superficielles, qui se limitent au niveau syntaxique de représentation et tentent d'extraire les parties saillantes du texte de manière pratique.

Deuxièmement, les approches plus profondes, qui supposent un niveau sémantique de représentation du texte original et impliquent un traitement linguistique d'un certain niveau.

Les applications comme *Resoomer* fonctionnent selon la première approche, où l'étape de prétraitement a pour but de réduire la dimensionnalité de l'espace de représentation, et peut comprendre également trois étapes à son tour:

- 1) L'élimination des mots vides, des mots communs sans sémantique et les informations agrégées pertinentes pour la tâche (par exemple, "le", "une") sont éliminés.
- 2) Le pliage de la case : consiste à convertir tous les caractères en un même type de lettres majuscules ou minuscules.
- 3) La racine : les mots syntaxiquement similaires, tels que les pluriels, les variations verbales, etc. sont considérés comme similaires. L'objectif de cette procédure est d'obtenir la racine ou le radical de chaque mot, qui met en valeur sa sémantique.

Un modèle de texte fréquemment utilisé est le modèle vectoriel. Après l'étape de prétraitement, chaque élément de texte - une phrase dans le cas d'un résumé de texte - est considéré comme un vecteur à N dimensions où chaque mot représente une dimension. Il est donc possible d'utiliser une certaine métrique dans cet espace pour mesurer la similarité entre les éléments de texte.



La métrique la plus utilisée est la mesure du cosinus, définie comme  $\cos q = (\langle x, y \rangle) / (|x| \cdot |y|)$  pour les vecteurs  $x$  et  $y$ , où  $(\langle, \rangle)$  indique le produit scalaire, et  $|x|$  indique le module de  $x$ . Par conséquent, la similarité maximale correspond à  $\cos q = 1$ , tandis que  $\cos q = 0$  indique la divergence totale entre les éléments de texte.

Pour mettre en œuvre la synthèse de texte basée sur la logique floue, on utilise généralement **MATLAB**, car il est possible de simuler la logique floue dans ce logiciel. Il faut ensuite sélectionner des caractéristiques d'un texte telles que la longueur de la phrase ou encore la similarité avec le mot clé comme entrées d'un système flou. Ensuite, toutes les règles nécessaires à la synthèse sont entrées dans la base de connaissances de ce système. Cela étant, une valeur de 0 à 1 est obtenue pour chaque phrase dans la sortie en fonction des caractéristiques de la phrase et des règles disponibles dans la base de connaissances. La valeur obtenue dans la sortie détermine le degré d'importance de la phrase dans le résumé final (Gupta et Lehal, 2009).

La figure 24 présente ce processus de résumé.

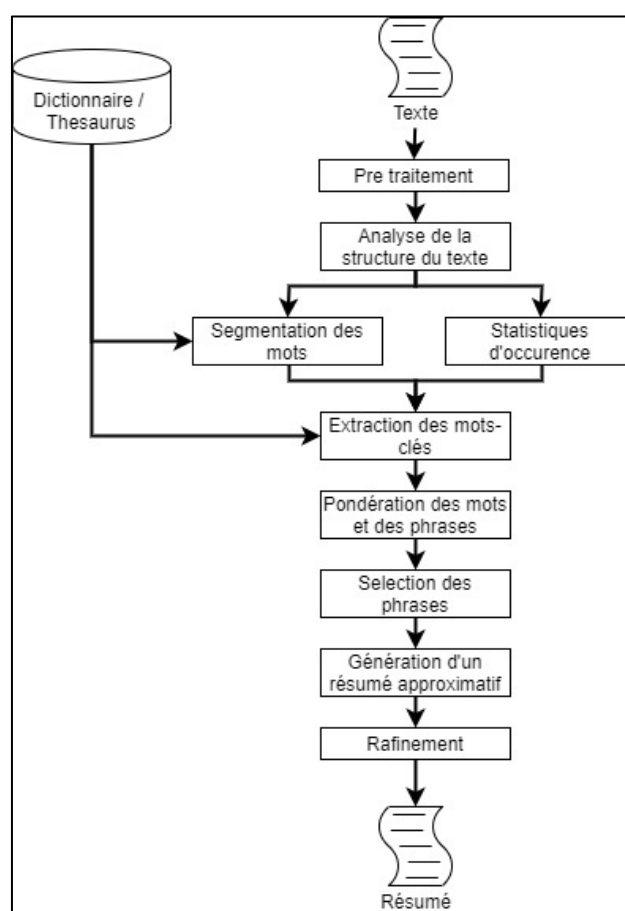


Figure 24 : Processus de résumé de texte

Source : Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60-76. <https://doi.org/10.4304/jetwi.1.1.60-76>

Il va de soi que ce genre d'application ne fonctionne que pour les textes. Néanmoins, il arrive qu'une vidéo ou un enregistrement audio présente du contenu de qualité utile pour *EU Monitoring*. Par exemple, les réunions de commissions parlementaires sont très intéressantes, mais les procès-verbaux sont très rudimentaires. Dans le chapitre 5, nous explorerons de potentielles solutions pour pouvoir automatiser l'exploitation de ce genre de contenu.

## 4. Installation de la plateforme

Ce point ne concerne aucune étape du *business process*, mais il est important de le développer, car il sera utile aux étapes qui suivent et qui dépendent toutes de la forme que la plateforme prendra.

Comme nous l'avons constaté lors de notre analyse des systèmes équivalents déjà existants, il serait opportun de disposer d'un système qui fasse office de recueil du contenu.

Nous avons donc décidé de mettre en place une plateforme. Cette plateforme doit répondre à certaines exigences citées notamment dans l'analyse ADEPT. Comme nous ne possédons pas les compétences suffisantes pour coder un système à partir de rien, nous nous sommes orientés vers une solution de **Wiki open source**.

La section suivante décrit la démarche que nous avons suivie pour déterminer quelle solution de wiki est la plus adaptée à *EU Monitoring*.

### 4.1. Choix de la solution Wiki

Nous avons commencé par chercher quelles étaient les solutions de wiki les plus accessibles et répandues.

- MediaWiki – Solution Wiki de Wikipedia
- DokuWiki
- PmWiki
- Tiki
- La fonctionnalité Wiki de Microsoft Teams

Nous avons également listé les exigences auxquelles la plateforme doit répondre. Cette liste est le résultat de deux choses. D'une part, l'analyse ADEPT. D'autre part, nous avons constaté lors de nos recherches de Wiki's, que des aspects proposés par certains seraient pertinents pour *EU Monitoring*.

Voici la liste des exigences :

- i. Gestion des utilisateurs : Pouvoir ajouter et retirer facilement des utilisateurs et pouvoir les classer dans différentes catégories.
- ii. Personnalisation : Pouvoir donner l'apparence de Sopra Steria à la plateforme.

- iii. Système de verrouillage : Système interdisant à deux utilisateurs d'éditer un même article en même temps.
- iv. Calendrier
- v. Chargement de fichiers
- vi. Canaux : Pouvoir trier le contenu dans différentes catégories
- vii. Type de stockage des données

Avec ces deux listes, nous avons dressé le Tableau 5 afin de définir quel Wiki était le plus complet.

Tableau 5 : Sélection du Wiki

	Gestion des utilisateurs	Person°	Système de verrouillage	Calendrier	Chargement de fichiers	Channels	Type de stockage des données
<b>MediaWiki</b>		X				X	?
<b>DokuWiki</b>	X	X	X			X	?
<b>PmWiki</b>	X	X	Similaire à google doc	X	X	X	Flat file
<b>Tiki</b>	X	X	X	X	X	X	MySQL
<b>Microsoft Teams Wiki</b>			Similaire aux fichiers Teams	X	X (pas adapté)	X (pas adapté)	Teams

Les X désignent le fait qu'une solution répond à une exigence. L'absence de X signifie qu'elle n'y répond pas ou que l'information disponible n'est pas suffisante pour l'affirmer.

Après l'analyse de ce tableau, deux solutions se démarquent, car elles répondent le mieux à nos exigences, PmWiki et Tiki. Nous avons dès lors téléchargé les deux Wiki's pour déterminer lequel nous garderions.

Pour cette étape, nous n'avions bien sûr pas besoin d'un serveur en ligne, mais simplement de pouvoir les héberger en local. Pour cela, nous les avons déployés sur **WampServer** (Apache, MySQL, PHP for Windows) qui permet justement d'avoir un site internet en local host.

Nous analyserons *WampServer* plus en détails plus loin dans ce document.

Une fois les deux solutions déployées, nous avons fait le choix de garder la solution **Tiki**. Ce choix repose sur trois raisons principales.

Premièrement, après les avoir succinctement essayés, Tiki semble plus User-Friendly. La personnalisation, la création de contenus ou encore la gestion des utilisateurs étaient plus instinctives.

Deuxièmement, PmWiki semble demander plus de compétences de programmation pour la personnalisation, compétences que nous ne sommes pas confiants d'avoir dans l'équipe du projet.

Troisièmement, nous avons plus d'expérience et étions plus confortables avec le stockage et l'extraction de données de type MySQL.

#### 4.2. Utilisation de *WampServer*

La machine que nous utilisons pour héberger la plateforme n'est au départ qu'un ordinateur fonctionnant avec le système d'exploitation Windows. Il faut alors la rendre capable d'héberger un site et plus tard rendre ce site accessible par d'autres ordinateurs.

Pour cela, nous avons téléchargé et installé le logiciel *WampServer*. Ce logiciel imite les services identifiés dans l'analyse ADEPT, à savoir serveur HTTP ainsi que serveur et gestion de base de données. En effet, WAMP sont les initiales de Windows (fonctionne pour Windows), Apache (technologie pour administrer le serveur http), MySQL (administration de la base de données) et PHP (administration de la base de données et du serveur web). Nous pouvons à présent mettre à jour l'architecture d'*EU Monitoring*, comme illustré à la Figure 25.

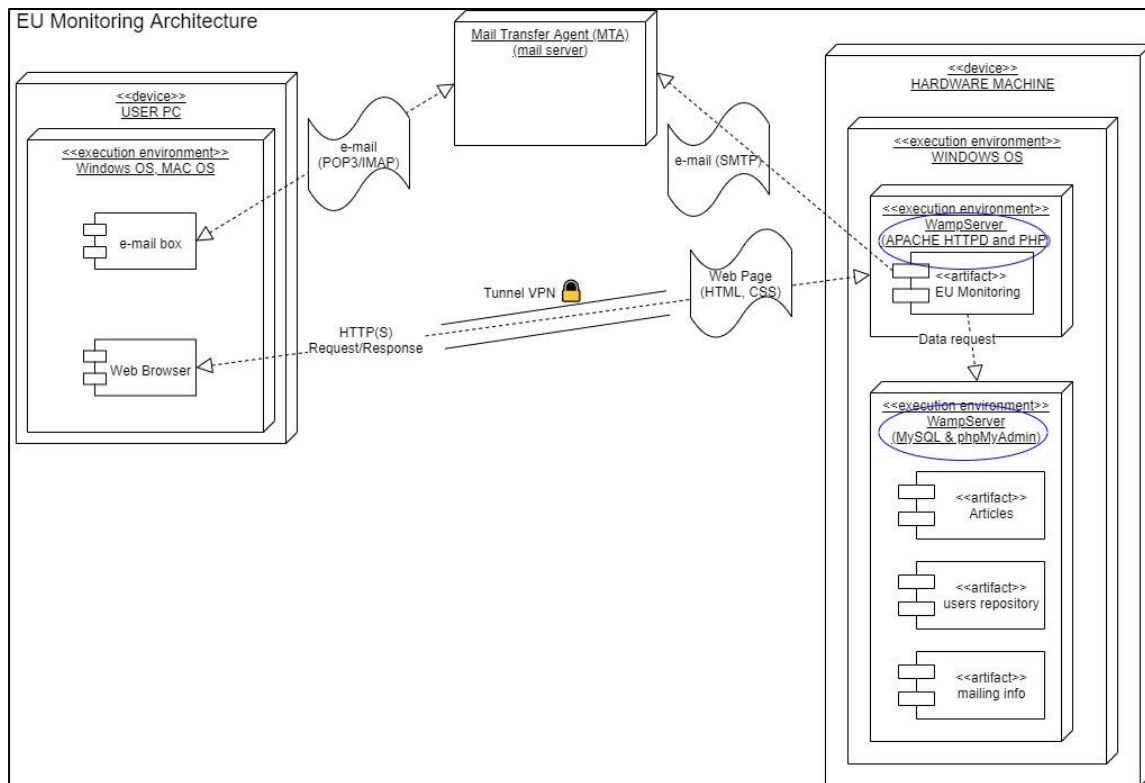


Figure 25 : Architecture d'EU Monitoring mise à jour

Tiki se télécharge sous la forme d'un fichier zip qu'il faut extraire et ensuite déployer dans l'emplacement de *WampServer*. Une fois cette action exécutée, on peut visualiser la plateforme dans un navigateur web. Cependant, le contenu est toujours en « localhost », ce qui signifie qu'il n'est visible que depuis la machine où il est hébergé. Nous avons besoin de le rendre visible au reste du réseau de Sopra Steria.

La machine hébergeant *EU Monitoring* a une adresse IP publique la rendant accessible sur tout le réseau. Il faut alors modifier les paramètres Apache de *WampServer* et remplacer « localhost » par l'adresse IP publique en tant que ServerName. La plateforme est désormais accessible par tous les ordinateurs connectés au réseau de Sopra Steria.

### 4.3. Accès à la machine hôte à distance

Bien entendu, la machine qui héberge le code de la plateforme est fixe. Cependant, une application appelée « Remote Desktop Connexion » permet d'accéder au bureau de cette machine à condition d'en connaître l'adresse IP. Cette application est particulièrement utile, car elle permet de pouvoir modifier les paramètres de *WampServer* ou le code de la plateforme sans devoir être physiquement au contact de la machine. Cela s'est révélé très opportun au moment où nous avons dû rester à domicile durant le confinement lié à la pandémie de Coronavirus.

## 5. Définition du champ d'application

Une fois un document résumé, il faut le publier sur la plateforme mais avant cela, il faut en définir le ou les champs d'application pour que le référencement soit le plus pertinent possible. La plateforme Tiki permet de donner un sujet et ainsi de référencer un article. Dès lors, il est possible pour les utilisateurs de retrouver tous les articles concernant un sujet en particulier.

## 6. Stockage de l'information

### 6.1. Contenu permanent vs Actualité

Maintenant qu'un texte est résumé et référencé, il faut pouvoir le stocker sur la plateforme.

Une des premières constatations que nous avons faites est qu'un résumé tout seul n'est parfois pas suffisant. Dans certains textes, on fait références à des sujets qu'il convient d'expliquer davantage. Par exemple, si on évoque les liens qu'EES aura avec ETIAS et l'expertise que FRONTEX pourrait apporter, un lecteur non initié à cet écosystème européen pourrait estimer qu'il y a trois éléments qui devraient être expliqués dans cette phrase.

Pour cette raison, nous avons deux parties principales à notre plateforme. D'une part, la partie « Wiki » qui permet de regrouper le contenu permanent qui consiste en des fiches descriptives sur les agences, les systèmes ou encore les régulations et d'une autre part, la partie « Articles » qui regroupe le contenu d'actualité.

Grâce à cela, nous pouvons inclure des liens dans les articles renvoyant vers des fiches descriptives. De plus, dans ces fiches descriptives, nous pouvons inclure du matériel conçu par Sopra Steria ayant ses couleurs (PowerPoint ou autres).

Ce matériel est le résultat de la connaissance de l'entreprise et *EU Monitoring* est un bon endroit où le partager pour démontrer les compétences de Sopra Steria.

Ce matériel ainsi que des textes et des images peuvent se stocker dans les « File Galleries ».

Les Figure 26 et Figure 27 montrent respectivement la partie Wiki et la partie Articles.

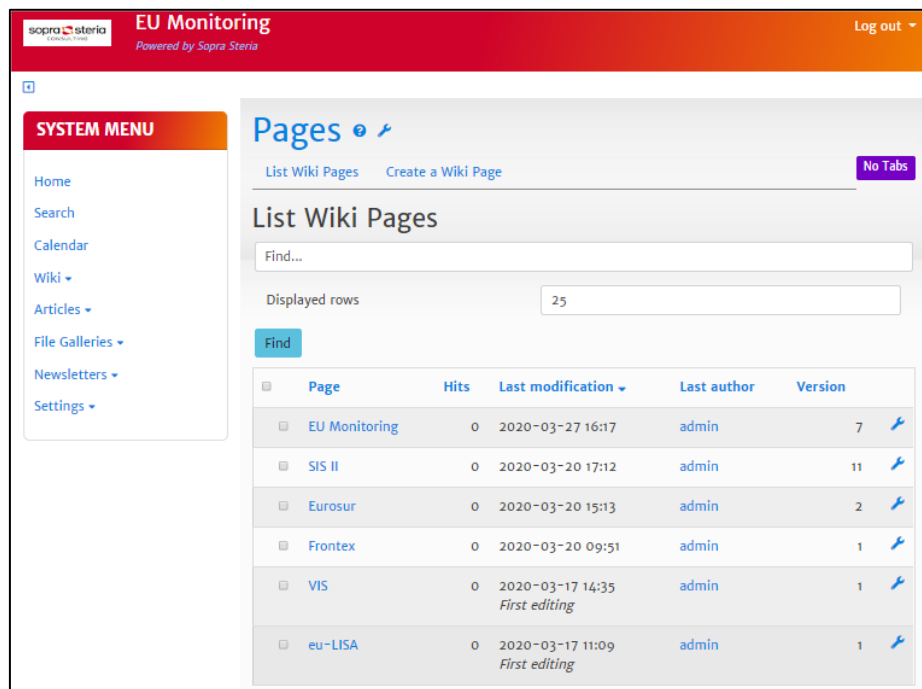


Figure 26 : Partie Wiki d'EU Monitoring

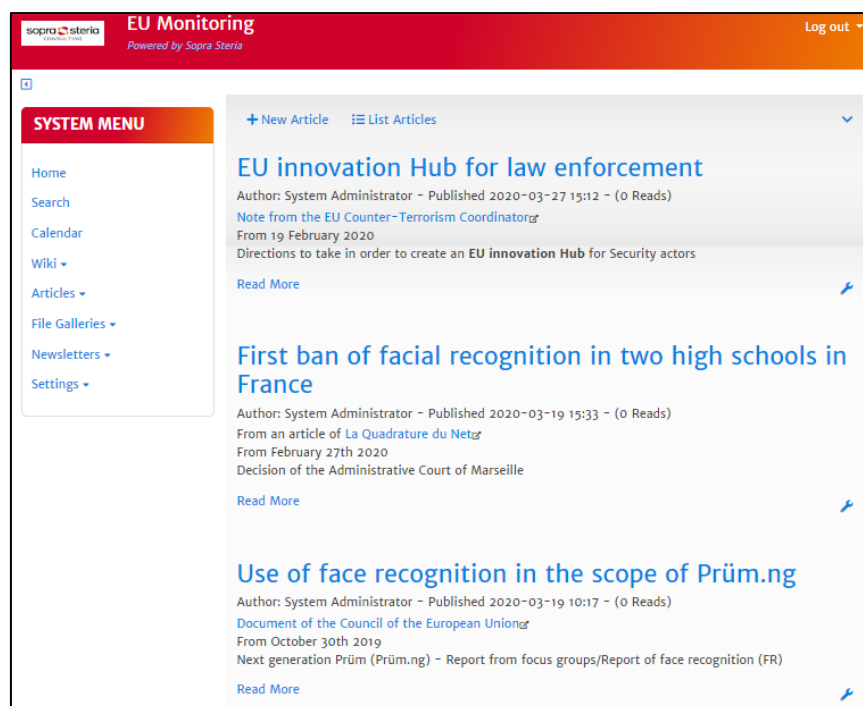


Figure 27 : Partie Articles d'EU Monitoring

Les deux types de contenu ont une forme standardisée. En effet, chaque élément du contenu permanent comporte la source de la documentation ainsi que la base légale lorsque c'est applicable. Chaque élément d'actualité comporte le type de document (article,

communication du parlement, etc), la date de publication du document original et une courte description du sujet traité.

## 6.2. Gestion des utilisateurs

Selon la source, nous ne connaissons pas toujours l'origine d'un document que nous analysons. Par exemple, **Statewatch**, un site d'information, parvient à se procurer des documents qui ne sont pas ou pas encore accessibles à tous. Cela peut parfois être un document fuité d'une institution européenne. Statewatch partage ces documents mais ce n'est pas le rôle d'*EU Monitoring* de les relayer.

Toutefois, ces documents peuvent présenter un intérêt certain pour l'agence Consulting de Sopra Steria et ont dès lors leur place sur *EU Monitoring*. C'est pourquoi il est primordial de pouvoir contraindre les privilèges de chaque utilisateur. Cette dichotomie est également d'application pour le matériel conçu par Sopra Steria, car tout ne peut pas être partagé avec ses clients.

Pour le moment, nous avons quatre différents groupes. Le premier est **Anonymous**, qui est un visiteur non connecté qui n'a accès qu'à la page d'accueil. Le second, **Registered**, donne accès au contenu de la plateforme sauf à certains éléments. Pour avoir accès à ces éléments, il faut être dans le groupe **Sopra Steria**, qui est le troisième groupe. Le quatrième et dernier groupe est **Admin** qui lui donne accès à toutes les fonctionnalités de la plateforme.

Le point fort de la gestion des utilisateurs sur Tiki est sa simplicité et sa modularité. Il est possible de définir les privilèges aussi bien de chaque page individuellement que d'un ensemble de pages. Pour illustrer ce propos, la Figure 28 reprend deux captures d'écrans. L'une montre l'accessibilité de la page d'accueil que tout le monde peut voir, tandis que la seconde montre une page que seuls les utilisateurs enregistrés dans le groupe Sopra Steria peuvent voir.



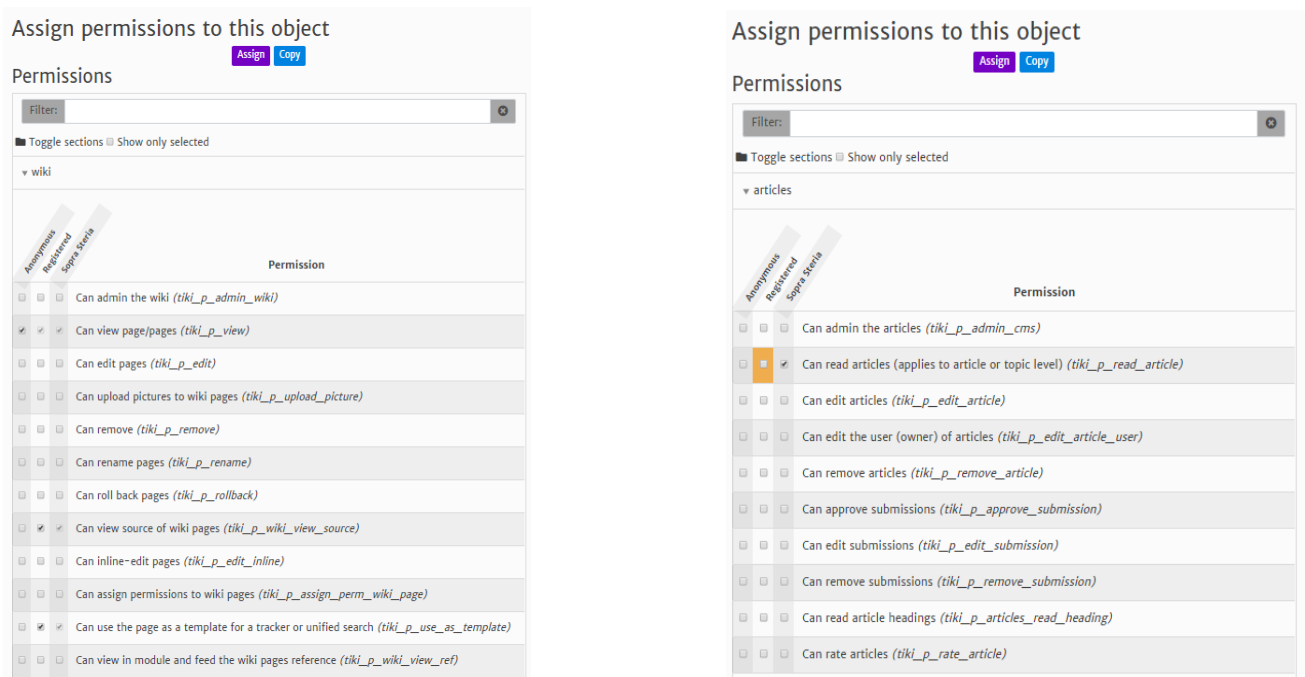


Figure 28 : Privilèges maximum et minimum d'une page

Imaginons que dans le futur, le champ d'action d'*EU Monitoring* s'étende à davantage de sujets. Au lieu de devoir installer plusieurs plateformes, nous pouvons utiliser la gestion d'utilisateurs pour limiter l'accès de chacun à ce qui est pertinent pour lui.

La liste des utilisateurs et leurs privilèges étant une table de données comme d'autres, la gestion peut également se faire via **phpMyAdmin**, développé au point suivant.

### 6.3. Gestion de la base de données

Pour administrer la base de données, *WampServer* donne accès à un service s'appelant *phpMyAdmin*. Ce service permet de supprimer des tables et d'en créer de nouvelles. Cela permet également d'exporter une base de données. Dans le cadre d'*EU Monitoring*, cela pourra s'avérer utile au moment de migrer la plateforme vers un autre système de stockage. En effet, lorsque qu'*EU Monitoring* ne sera plus un PoC, mais bien un service opérationnel, il faudra le migrer vers une solution de type cloud par exemple.

## 7. Partage de l'information

### 7.1. Newsletter

La plateforme Tiki rend possible l'envoi de newsletters comportant le contenu créé sur *EU Monitoring*. Toutefois, il faut que le serveur qui n'est pas connecté à internet soit capable d'envoyer des e-mails. Comme indiqué sur l'architecture d'*EU Monitoring*, nous utilisons le protocole simple de transfert de courrier (SMTP – *Simple Mail Transfer Protocol*). C'est un

protocole permettant de transmettre le courrier électronique vers des serveurs de messagerie électronique.

Afin de permettre à *WampServer* d'envoyer des e-mails, nous avons suivi plusieurs tutoriels indiquant qu'il fallait télécharger une extension appelée **sendmail**. Cette extension comprend plusieurs scripts dont un où il faut remplir les références de l'adresse e-mail que l'on souhaite utiliser pour envoyer les e-mails. Cette façon de procéder fonctionne sur un réseau domestique avec une adresse e-mail commerciale comme G-mail de Google.

Toutefois, le procédé semble plus compliqué lorsqu'on souhaite l'exécuter sur un réseau d'une grande entreprise avec les critères de sécurité que cela implique et en utilisant une adresse e-mail dépendant de l'entreprise.

À la fin du temps imparti pour le projet, l'envoi d'e-mails depuis la plateforme n'était pas opérationnel. Lorsque le projet sera repris, il faudra discuter l'utilité de la *newsletter* et s'il en ressort que c'est effectivement une bonne solution, il faudra explorer d'autres manières d'actionner l'envoi d'e-mails. Des contacts ont malgré tout eu lieu avec la DSI, service responsable du support informatique au sein de l'entreprise, afin de créer un compte SMTP.

## 8. Harmonisation du *Business Process*

Pour chaque étape, un ou plusieurs éléments ont été mis en place afin de réaliser le travail d'*EU Monitoring*. Afin que ce soit facile pour quiconque devrait reprendre le projet en main, nous avons voulu rendre le *business process* clair et indiquer les outils correspondants à chaque étape.

- Récolte d'information :
  - ⇒ Boîte e-mail dédiée ;
  - ⇒ **Feedly** ;
  - ⇒ Reading Pipe dans Microsoft **Teams**.
- Evaluation de la pertinence :
  - ⇒ Liste de sujets.
- Resumer l'information :
  - ⇒ **Resoomer**.
- Stockage de l'information :
  - ⇒ Remote Desktop Connexion pour l'accès à la machine hôte à distance ;
  - ⇒ **phpMyAdmin** pour la gestion de la base de données ;
  - ⇒ La rubrique « Articles » d'*EU Monitoring* pour le contenu d'actualité ;
  - ⇒ La rubrique « Wiki » d'*EU Monitoring* pour le contenu permanent.

Nous pouvons à présent mettre la modélisation du *Business Process* à jour à la Figure 29.

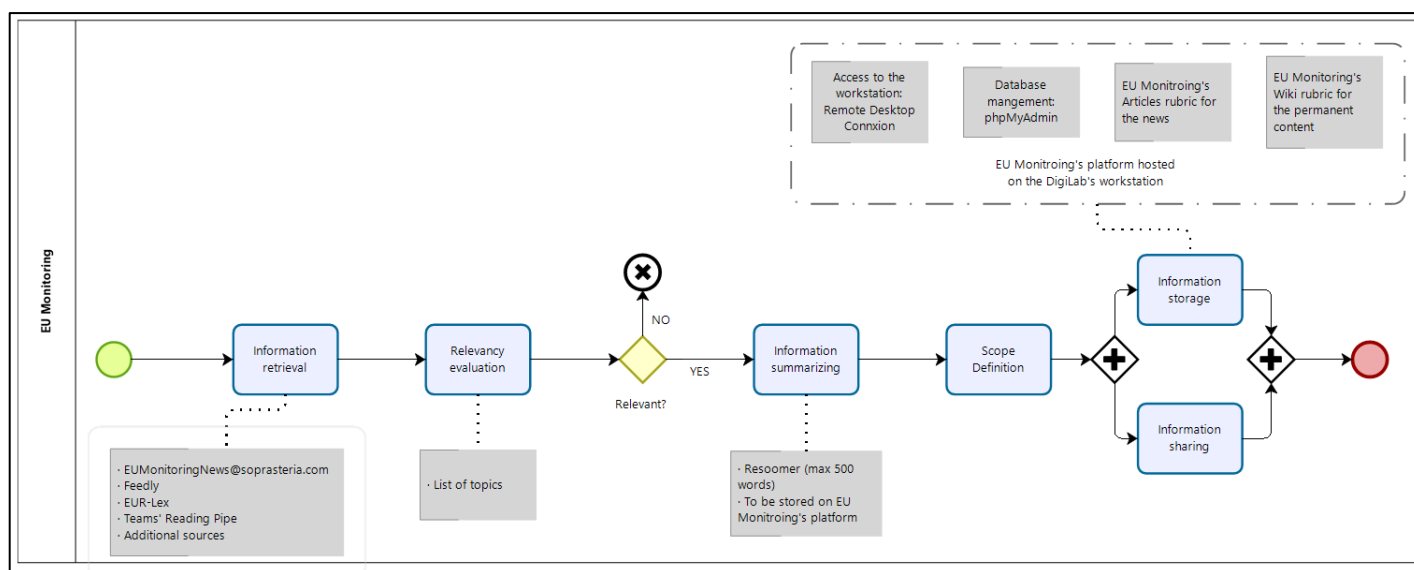


Figure 29 : Business Process mis à jour avec les outils utilisés

Dans l'optique d'automatiser au maximum le *business process*, un passage en revue de ce qui est automatique et de ce qui ne l'est pas est repris ci-dessous.

Tout d'abord, la récolte d'information est centralisée en trois endroits principaux. Il ne faut donc plus rechercher manuellement des informations sur le Web.

Ensuite, les résumés peuvent être créés automatiquement grâce à Resoomer mais nécessitent malgré tout l'attention d'un opérateur humain.

Pour le reste, un opérateur humain doit superviser complètement les autres tâches. Posons-nous dès lors la question de savoir s'il serait pertinent d'utiliser l'automatisation des procédés (*Robotic Process Automation, RPA*) ? La RPA sert principalement à confier à la machine des tâches répétitives et demandant peu de compétences intellectuelles. Ainsi, le temps nécessaire à un humain pour répéter ces tâches peut être économisé.

Dans le cadre d'*EU Monitoring*, on peut facilement imaginer l'automatisation de plusieurs tâches et principalement la transition entre les tâches. Si nous devons avoir recours à la RPA, la machine pourrait prendre un article de Feedly, pour le copier dans Resoomer et finalement enregistrer le résumé sur la plateforme.

Pour plusieurs raisons, l'utilité de cette automatisation peut être discutée. Comme expliqué plus en avant, la pertinence du contenu disponible sur Feedly n'est pas assurée et une extraction automatique aurait comme résultat que trop d'articles sur la plateforme n'y auraient pas leur place. Il faudrait alors qu'un opérateur humain aille faire le tri et les enlève. Ce problème peut être solutionné par un système qui améliorerait la pertinence de la récolte d'information. Le *Web Mining* peut par exemple aider dans ce sens.

Ensuite, un autre obstacle à l'automatisation est le fait que les résumés créés grâce à Resoomer ne sont pas parfaits. Il faudrait donc qu'un opérateur humain retouche les résumés présents sur la plateforme.

Finalement, le référencement ne peut pas se faire automatiquement car pour cela, la machine devrait déterminer le sujet principal de l'article, ce qui semble difficile à ce stade. Il faudrait donc que l'opérateur humain se charge du référencement manuellement.

Nous constatons donc que le temps économisé par un humain ne serait finalement pas si important que nous pourrions l'espérer. Cela aurait toutefois un avantage qui est le fait que le travail est centralisé en un endroit qu'est la plateforme.

Aussi, pour le moment, le *business process* d'*EU Monitoring* ne se compose pas d'énormément d'étapes. Ce faible nombre d'étapes rend gérable le fait de réaliser tout ou presque manuellement. Dans le cas où ce nombre passerait à 15 ou 20 étapes, il deviendrait alors très opportun voire quasi primordial d'automatiser une partie voire l'ensemble du *business process*.

Néanmoins, cette solution n'est sûrement pas à mettre définitivement de côté. Avec le soutien du *machine learning* permettant à la machine de comprendre les éléments qui valent la peine d'être pris en considération, l'automatisation pourrait sans nul doute faciliter grandement le travail d'*EU Monitoring*. Il est peut-être juste encore trop tôt pour pouvoir l'utiliser.

## 9. Identification des besoins à court terme

Avant que l'on puisse considérer *EU Monitoring* comme un service mature et partageable avec plus d'utilisateurs autant internes qu'externes à l'entreprise, certains points doivent encore être adressés. La plupart des éléments qui seront cités dans cette portion sont des propositions ou des recommandations provenant d'un document rédigé par le stagiaire à la fin de son stage à destination de l'équipe du projet. Lorsque c'est applicable, nous proposerons également les avis du chef de projet (Cédric Genin) à certaines de ces propositions et recommandations.

### 9.1. Parvenir à partager l'information

Au point 7 de ce chapitre, nous avons constaté que le système de *newsletter* n'était pas au point. La dernière étape du *Business Process* consistant à partager l'information n'est actuellement pas prise en charge. La recommandation est donc de faire appel à des collaborateurs ayant de l'expérience de configuration de serveurs SMTP et une compréhension extensive des mesures de protection mises en place au sein de l'entreprise.

L'avis de M. Genin est qu'il serait intéressant de discuter l'intérêt d'utiliser des *newsletters*. En effet, sans forcément mettre les *newsletters* de côté, il existe d'autres façons de partager l'information. Des flux RSS sortants pourraient par exemple être une solution de remplacement ou une solution complémentaire.

## 9.2. Migration du système

Il faut aussi que la plateforme soit plus facilement accessible, ce qui passerait par une migration de celle-ci vers un autre serveur dont l'accès est possible depuis internet. La solution actuelle est tout à fait pertinente pour l'utilité présente et le restera encore pour l'évolution prochaine de la plateforme. L'infrastructure actuelle a permis de réaliser tout au moins une bonne maquette voire un PoC acceptable. Toutefois, le moment de la migration n'est pas très lointain. Cette migration devra se faire dans un environnement comprenant un plan de continuité d'activité (PCA). Ce type de plan en informatique permet de garantir la poursuite des prestations d'un système (Alarmit, 2020). Cela comprend par exemple un back-up des données et un système électrique de secours.

## 9.3. Calibration des moyens humains

Le choix du profil de stagiaire pour s'occuper du projet est un profil d'ingénieur commercial. Il semble que cela ait été opportun pour les premiers pas du projet, tel que la définition des besoins et la mise en place des premiers outils. Toutefois, pour certains aspects, un support d'autres types de profils pourrait s'avérer utile. Voici les principaux points pour lesquels un ingénieur commercial pourrait atteindre son seuil de compétence :

- Un très bon niveau d'anglais est attendu de tels profils mais il n'est pas acquis que le niveau dans cette langue soit tout à fait irréprochable. Bien que le choix d'éditer le contenu en anglais soit le bon, il serait intéressant d'avoir un support pour la relecture. Des fautes d'orthographe ou de grammaire nuiraient grandement à la crédibilité du service.
- La programmation informatique ne fait pas partie des points forts d'un ingénieur commercial. Cependant, afin de pouvoir personnaliser encore davantage la plateforme basée sur Tiki, de telles compétences seraient les bienvenues. Le fait qu'elle soit codée dans des langages aussi répandus que PHP, CSS et SQL permet de pouvoir le personnaliser relativement facilement. C'est pourquoi il faudrait qu'*EU Monitoring* puisse bénéficier de l'appui ponctuel d'un développeur désigné.
- La recherche d'informations utiles à *EU Monitoring* s'apparente parfois à un travail journalistique avec les compétences que cela requiert. Posons-nous alors la question de savoir s'il ne faudrait pas considérer ce genre de profils pour s'occuper de la production de contenu. Cédric Genin abonde dans ce sens en affirmant qu'une personne avec un parcours journalistique, voire même juridique ou en sciences politiques, pourrait être plus adéquat lorsque l'ensemble des fonctionnalités seront opérationnelles.

#### 9.4. Claire définition de la portée d'*EU Monitoring*

Il faudrait définir clairement si *EU Monitoring* doit être une source d'information ou bien un outil d'analyse légale. Les deux rempliraient l'objectif principal d'aider les utilisateurs à se retrouver dans les législations européennes.

Néanmoins, l'approche n'est pas tout à fait la même. Présentement, nous avons pris le parti d'en faire une source d'information plus ou moins brute, mais il faudrait cependant analyser la possibilité d'y présenter des cadres légaux. Cela rejoint dès lors le point précédent, dans le sens où la définition de cadres légaux est un travail qui relève des juristes.

Le chef de projet propose une perception nuancée en affirmant que sans aller jusqu'à de l'analyse légale telle qu'un juriste pourrait le faire, fournir une mise en contexte et une explication des conséquences des décisions prises au niveau européen serait pertinent et apporterait de la valeur ajoutée à l'information collectée.

Nous avons maintenant passé en revue tout ce qui a été mis en place et ce qui devrait être mis en place dans un futur proche. Nous pouvons à présent passer au chapitre suivant, où l'on comparera ce qui a été mis en place avec les objectifs que nous nous étions fixés.

## Chapitre 4 : Bilan

---

Dans ce chapitre, nous allons analyser à quel point ce qui a été fait correspond à ce qui était prévu. Nous nous poserons également la question de ce qui pourrait être amélioré dans la gestion de projet et de ce qui, au contraire, a bien fonctionné. Les pistes pour les potentielles améliorations fonctionnelles du système seront abordées au chapitre suivant.

### 1. Accomplissement des objectifs

#### Objectif 1 : Trouver suffisamment de sources d'information

Avec les trois outils que sont Feedly, le pipe de lecture et la boîte e-mail, la quantité d'informations récoltées est importante. Lors des deux semaines qui ont suivi la mise en place du dernier de ces trois outils, nous avons suffisamment de contenu pour travailler pendant deux jours complets.

Cependant, après cela, l'actualité a été submergée par les nouvelles concernant la pandémie de COVID-19. Le matériel pertinent pour *EU Monitoring* est passé au second plan et s'est donc fait plus rare par la suite, ne permettant plus de remplir l'objectif.

Si nous estimons que sans l'apparition du virus, nous aurions pu continuer au même rythme que les deux premières semaines et qu'il s'agissait de la seule période pour juger de l'efficacité de la récolte d'information, l'objectif est rempli à 100 %.

#### Objectif 2 : Mettre en place une plateforme pour le partage et le stockage des données

La plateforme Tiki a été installée et bénéficie d'un espace de stockage. Pour les besoins actuels d'une preuve de concept, la plateforme est suffisamment opérationnelle. Néanmoins, pour devenir une plateforme à des fins commerciales, il serait nécessaire de migrer vers un environnement assurant les systèmes redondants adéquats.

Le but était d'avoir une preuve de concept et pour cela, cet objectif est rempli à 100 %.

#### Objectif 3 : Mettre en place des outils pour simplifier le Business Process

La Figure 29 montrant le *Business Process* et les outils mis en place pour chaque étape nous permet de constater que toutes les étapes n'ont pas un outil associé. En effet, trois des six n'en ont pas. L'objectif n'est donc rempli qu'à 50 %.

## Objectif 4 : Établir une manière standardisée de rapporter les résultats

Le contenu est divisé en deux catégories (articles et wiki). Chacune de ces catégories présente une manière standardisée que tous les éléments respectent. L'objectif est dès lors rempli à 100 %.

## Objectif 5 : Production de contenu d'*EU Monitoring*

L'observation faite pour l'objectif 1 est également d'application ici, dû à leur lien évident. Ne pouvant pas trouver suffisamment d'information à cause de la pandémie, nous ne pouvions pas non plus produire autant de contenu par rapport à ce qui était originalement prévu. Cependant, comme pour l'objectif 1, nous avons pu produire pendant deux semaines précédant le confinement autant de contenu que ce que nous souhaitions, voire même plus que les six éléments fixés comme objectifs.

Si l'on considère que les deux semaines sont suffisantes pour juger de l'efficacité du système, on peut considérer que l'objectif est rempli à 100 %.

Notons par ailleurs que *Resoomer*, l'outil servant à résumer des textes, a commencé à être utilisé pendant la période creuse en termes d'information. Nous pouvons donc espérer que l'efficacité pourrait même être accrue en cas de flux normal d'actualités.

## Objectif 6 : Valoriser le contenu d'*EU Monitoring*

La piste principalement explorée pour la valorisation du contenu fut la *newsletter*. Cette dernière ne fonctionnant pas, nous n'avons pour l'heure encore aucun autre moyen de valoriser l'information.

Sans aucun moyen de transmettre le travail d'*EU Monitoring* à plus d'utilisateurs il faut considérer que l'objectif est rempli à 0 %.

## Objectif 7 : Harmoniser les outils du *business process*

À la fin du stage, une présentation PowerPoint servant de « *Handover* » a été créée. Cette présentation très complète reprend tous les outils utilisés dans l'exécution du *Business Process*, les manières de les utiliser et les références des différents comptes. Dans cette présentation se trouve par exemple la modélisation du *Business Process* mise à jour avec les outils, comme montrée par la Figure 29.

Grâce à cette présentation, toute personne reprenant en main le projet peut être rapidement opérationnelle et se rendre facilement compte de tout ce qui a déjà été mis en place. L'objectif est rempli à 100 %.



## Objectif 8 : Identifier les besoins à court terme pour la finition du projet

Un document déjà mentionné au point 9 du chapitre précédent a été écrit. L'efficacité de ce document ne peut pas encore être évaluée. Il faudrait d'abord que les recommandations faites soient appliquées pour pouvoir les considérer pertinentes ou non.

Un aperçu des objectifs et de leur réalisation respective est repris au Tableau 6.

Tableau 6 : Réalisation du projet

Objectif n°	Métrique/Délivrable souhaitée	Métrique/Délivrable observée	Réalisation
1	Contenu suffisant pour 2 jours de travail complets	Contenu suffisant pour 2 jours de travail complets	100%
2	Plateforme opérationnelle	Plateforme opérationnelle	100%
3	1 outil/étape	1 outil/étape seulement pour 3 étapes sur 6	50%
4	Tous les documents ont la même forme	Tous les documents ont la même forme	100%
5	6 résumés/semaine	6 résumés/semaine	100%
6	Mettre en place une newsletter	Rien n'a été mis en place	0%
7	Document reprenant les outils, et la façon de les utiliser	Document reprenant les outils, et la façon de les utiliser	100%
8	Document de retour d'informations à l'attention de l'entreprise	Document de retour d'informations à l'attention de l'entreprise	N/A
<b>Réalisation globale → (550% : 7)</b>			<b>78,6%</b>

## 2. Appréciations Générales

### 2.1. Fonctionnalités

Si nous comparons ce qui est fonctionnel par rapport au *business process* prévu lors de l'analyse ADEPT, nous nous rendons compte que toutes les étapes sont abordées et opérationnelles, sauf le partage d'information. Il est vrai que c'est une étape primordiale mais mis à part cela, *EU Monitoring*, en tant que tel, est opérationnel. Le système a déjà

réussi à détecter des informations pertinentes pour l'activité de Sopra Steria. Nous pouvons citer certains exemples de réussites.

- Nous avons découvert qu'e-CODEX, un projet de preuve de concept d'uniformisation des preuves dans des affaires judiciaires afin de se les partager à travers l'Union, pourrait devenir un système à part entière. La gestion de ce système pourrait être confiée à eu-LISA.

C'est évidemment intéressant pour Sopra Steria car eu-LISA étant déjà un client important, l'entreprise bénéficierait d'*aprioris* positifs en cas d'appel d'offre. Tout cela reste encore au conditionnel, car on est encore au stade des discussions.

- Nous avons pu être informés à l'avance que la Commission allait publier un livre blanc à propos de l'utilisation de l'intelligence artificielle. Lorsque ce livre blanc a finalement été présenté, nous avons pu tout de suite nous y intéresser sans perdre de temps. Ce genre de livre blanc est important, car il débouche souvent sur des propositions de régulations et finalement sur des régulations.
- Le système *False and Authentic Documents Online (FADO)* qui est, résumé brièvement, une base de données comprenant tous les documents officiels que l'on peut rencontrer dans l'Union Européenne ainsi que toutes les sortes de falsifications de ces documents déjà observées.

Ce système aide notamment les douaniers et les agents qui contrôlent les frontières à détecter les faux documents. Cependant, ce système vieillissant va voir sa gestion passer du Conseil à Frontex, ainsi que son fonctionnement mis à jour. Frontex, dont les moyens vont grandir significativement, n'est pas encore un client important de Sopra Steria. C'est pourquoi une telle nouvelle est intéressante car cela pourrait aider Sopra Steria à devenir un sous-traitant plus conséquent de Frontex.

Les deux premiers points valent la peine d'être soulignés, car ces informations ont été relevées avant même que les discussions fassent l'objet de propositions de la Commission. C'était un des aspects visés par *EU Monitoring*. Le troisième est aussi significatif, car c'est clairement ce genre d'opportunités que Sopra Steria peut exploiter.

Par ailleurs, si nous comparons les actions exécutées dans le cadre *EU Monitoring* à notre cycle théorique, nous réalisons qu'*EU Monitoring* n'exécute pas certaines de ces étapes.

Nous constatons tout d'abord que nous n'avons pas de définition systématique des besoins de la veille car les besoins de la veille sont assez clairs et ne varient que très peu. Nous n'avons dès lors pas ressenti le besoin de réaliser cette première étape du cycle.

Ensuite, comme mentionné lors de la définition des activités dans le chapitre 2, il n'y a pas de suivi automatique des projets de lois.

Finalement, la dernière étape manquante est évidemment la diffusion d'information par d'autres canaux que la plateforme que nous avons mise en place.

Pour le reste, notre système de veille correspond à notre cycle théorique.

## 2.2. Efficacité

Le projet d'EU Monitoring est encore trop jeune et trop peu implémenté dans l'entreprise pour pouvoir réellement juger de son efficacité. En effet, afin de juger d'un service, quel qu'il soit, il faut que ce dernier soit utilisé par suffisamment de personnes et depuis suffisamment de temps. Ce n'est pas encore le cas d'EU Monitoring.

Cependant, nous pouvons quand même réaliser une ébauche de *Balanced Scorecard* qui pourra être utilisée lorsque le système sera mature.

Pour rappel, *Balanced Scorecard* se compose de quatre éléments (Financier, Client, Processus, Apprentissage - Innovation). Pour chacun de ces éléments, il faut définir au moins un objectif avec un indicateur, une valeur cible et une initiative, tous liés à cet objectif.

Tableau 7: Balanced Scorecard d'*EU Monitoring*

	Objectifs	Indicateurs	Valeurs cibles	Initiative stratégique
<b>Financier</b>	<i>EU Monitoring</i> auto-suffisant	Résultat d'exploitation (Re)	Re > 0 EUR	Facturer l'utilisation d' <i>EU Monitoring</i>
<b>Client</b>	Acquérir de nouveaux clients	Nombre de nouveaux clients	2 nouveaux contrats cadres + 2 contrats par an	Proposer un essai d' <i>EU Monitoring</i> gratuitement à de potentiels clients
	Conformité des clients plus rapide	Taux de temps en moins pour atteindre la conformité	15% de temps en moins	Rien (à part l'utilisation d' <i>EU Monitoring</i> )
<b>Processus internes</b>	Partage d'information plus rapide	Taux de diffusion de l'information la première semaine	20% de diffusion supplémentaire	Rien (à part l'utilisation d' <i>EU Monitoring</i> )
<b>Apprentissage et innovation</b>	Être plus familier avec les spécificités techniques des régulations	Taux des collaborateurs familier avec ces spécificités	30% de collaborateurs familiers supplémentaires	Rien (à part l'utilisation d' <i>EU Monitoring</i> )

### 2.2.1. Explication des objectifs de la BSC (Tableau 7)

- *EU Monitoring* auto-suffisant : *EU Monitoring* ne doit pas être un service extrêmement lucratif, mais il doit cependant ramener suffisamment d'argent pour couvrir les coûts qu'il entraîne. À l'heure actuelle, les coûts sont très faibles, mais les revenus sont nuls.  
Si à l'avenir, il est possible de facturer l'utilisation de la veille, on aura une entrée d'argent clairement imputable à *EU Monitoring* et qui pourra justifier plus d'investissements. Aussi, si le service de veille permet que des clients souhaitent collaborer davantage avec l'entreprise, le résultat va augmenter, mais la portion imputable à *EU Monitoring* ne sera pas clairement identifiable.

- Acquérir de nouveaux clients : En offrant l'utilisation d'*EU Monitoring* à des entités que Sopra Steria aimerait débaucher comme client, cela pourrait déboucher sur la conclusion de nouveaux contrats-cadres. Le contrat-cadre est particulièrement intéressant, car il inscrit la collaboration dans la durée. *EU Monitoring* devrait aussi participer à l'obtention de contrats classiques.

#### Définition de contrat-cadre de *La Lettre des Réseaux*

D'après l'article 1111 du Code civil (français), « [l]e contrat-cadre est un accord par lequel les parties conviennent des caractéristiques générales de leurs relations contractuelles futures » et « [d]es contrats d'application en précisant les modalités d'exécution ». Le contrat-cadre est donc un contrat qui fixe, comme son nom l'indique, le « cadre » de la conclusion de contrats futurs, les contrats d'application ; le contrat-cadre appelle la conclusion de contrats d'application, en en fixant d'ores et déjà certaines conditions.

- Conformité des clients plus rapide : Ce service doit permettre aux clients de se conformer plus rapidement. De plus, *EU Monitoring* doit aussi aider les collaborateurs à acquérir suffisamment de compétences pour accompagner les clients vers la conformité.
- Partage d'information plus rapide : Une fois que l'entreprise acquiert de l'information, il faut qu'elle circule parmi toutes les personnes qui en ont besoin. Plus rapidement ce processus se met en place, mieux c'est.
- Être plus familier avec les spécificités techniques des régulations : Pour pouvoir offrir le meilleur service possible aux clients, il faut que les collaborateurs puissent être correctement informés des détails techniques.

## Chapitre 5 : Potentielles évolutions

Nous avons observé plusieurs éléments qui pourraient être améliorés ou simplement rajoutés au *business process*. C'est pour cela que nous allons aborder plusieurs choses qui pourraient apporter des évolutions intéressantes à *EU Monitoring*. Ces évolutions sont principalement des ajouts de technologies pour permettre une exécution du travail plus rapide ou plus précise.

### 1. Web Mining

Imaginons qu'en plus des résumés d'articles, *EU Monitoring* propose également d'informer ses utilisateurs avec des tendances que l'on peut faire en analysant le Web. Le *Web Mining* (fouille ou exploration du Web) fait référence au processus global de découverte d'informations ou de connaissances potentiellement utiles et jusqu'alors inconnues à partir des données du web (Mebrahtu et Srinivasulu, 2017).

Il existe trois sortes principales de Web Mining (Figure 30):

1. *Web Content Mining* : analyse le contenu des pages Web en tant que tel ;
2. *Web Usage Mining* : analyse les interactions que les utilisateurs ont avec le Web ;
3. *Web Structure Mining* : analyse le lien entre les différentes pages Web.

Le Web Content Mining en particulier, peut être une source de contenu pour *EU Monitoring* en offrant plus de contenu en moins de temps. Nous nous focaliserons principalement sur les données textuelles présentes sur le Web.

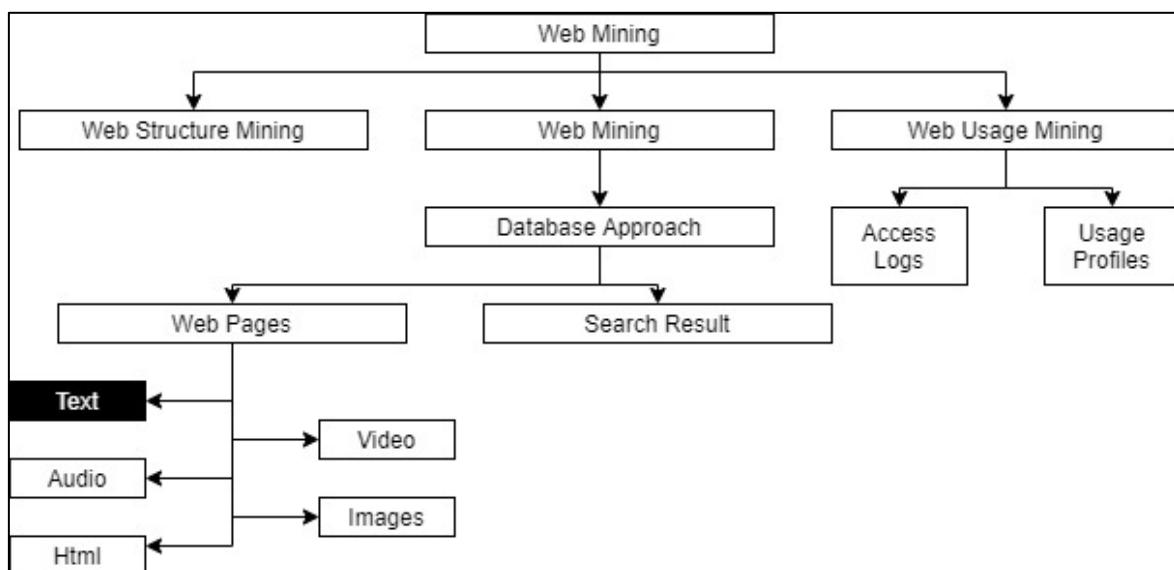


Figure 30 : Différents types de Web Mining

Source : Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60-76. <https://doi.org/10.4304/jetwi.1.1.60-76>

## 1.1. *Web Content Mining*

Le *World Wide Web* (WWW) est un média populaire et interactif dont la croissance de la quantité de données ou d'informations disponibles est énorme. Le *World Wide Web* est la collection de documents, de fichiers texte, d'images et d'autres formes de données sous forme structurée, semi-structurée et non structurée (Mebrahtu et Srinivasulu, 2017). Il est également immense, diversifié et dynamique, ce qui en augmente l'extensibilité. L'objectif principal de l'exploration du Web est d'extraire des informations et des connaissances utiles du Web. Bien que nous ayons déjà une quantité intéressante d'informations grâce aux outils que nous avons mis en place tels que Feedly, cela ne reprend finalement que les flux RSS dont nous avons pris connaissance. Or, il y a beaucoup de contenu que nous ne connaissons pas. C'est pour cela qu'il serait intéressant de « ratisser » encore plus large grâce au Web Content Mining.

Le Web devient la première source d'information pour des utilisateurs de plus en plus nombreux dans différents domaines. Pourtant, l'exploration de ses ressources devient une tâche difficile en raison de son hétérogénéité et de son manque de structure. Les utilisateurs sont donc noyés par une surcharge d'informations et cela devient un véritable défi de s'y retrouver efficacement (Mebrahtu et Srinivasulu, 2017).

Premièrement, il est difficile de trouver l'information appropriée à notre recherche. Lorsqu'un utilisateur souhaite trouver des informations spécifiques sur le Web, il saisit une simple requête par mot-clé. La réponse à la requête sera la liste des pages classées en fonction de leur similarité avec la requête. Cependant, les outils de recherche actuels présentent certains problèmes, tels que la faible précision (due à la non-pertinence des résultats de recherche) et le faible rappel (incapacité à indexer toutes les informations disponibles) (Mebrahtu et Srinivasulu, 2017). Si par exemple, dans le cas d'*EU Monitoring*, nous voulions rassembler des informations via la simple recherche dans un navigateur, les résultats seraient sûrement très décevants. Soumettre « Schengen » ou « eu-LISA » à un navigateur nous guiderait vers trop de documents de faible pertinence.

Ensuite, il faut pouvoir extraire des connaissances de ce qu'on trouve sur le Web. Ce problème est un processus dit « *data-triggered* » (déclenché par les données), alors que le précédent est un processus dit « *query-triggered* » (déclenché par les requêtes). L'internaute doit alors extraire des informations potentiellement utiles à partir d'un ensemble de contenus disponibles (Mebrahtu et Srinivasulu, 2017). Cela n'est pas toujours évident, car même si l'information est pertinente, la quantité de sources et de formes rend ce processus compliqué.

Pour répondre à ces défis, on peut s'aider du *Web Content Mining* (Extraction de contenu du Web). On appelle contenu, les données visibles sur les pages web ou tout type d'information comprenant du texte, de l'audio, de la vidéo, des images, du HTML ou encore du XML. Afin de découvrir toutes ces ressources disponibles, il faut se servir de Robots Web qui passent le WWW en revue pour ainsi donner du contenu à analyser.

### 1.1.1. Découverte des ressources

Lors de la découverte des ressources, nous effectuons une recherche automatique de tout le contenu pertinent du Web. Cette tâche est réalisée par des robots Web, appelés aussi *Web Crawlers*, *Web agents*, *Web spiders*, etc.

Un *Web Crawler* est un programme qui parcourt le *World Wide Web* de manière méthodique et automatisée afin de générer une copie de toutes les pages visitées pour qu'elles soient ensuite traitées par un moteur de recherche.

En raison de la limitation de la bande passante, des ressources informatiques et de la nature dynamique du web, les moteurs de recherche ne peuvent pas indexer toutes les pages Web, et même la partie couverte du Web ne peut pas être surveillée en permanence pour détecter les changements. Il est donc essentiel de développer des agents efficaces pour effectuer des recherches en temps réel pour les utilisateurs (Kumar et al, 2010).

En général, il est difficile d'utiliser un seul *crawler* pour rassembler toutes les pages. Par conséquent, de nombreux moteurs de recherche exécutent souvent plusieurs processus en parallèle pour accomplir cette tâche, c'est pourquoi le terme "crawling en parallèle" est utilisé. Cette approche peut améliorer considérablement l'efficacité de la collecte (Kumar et al, 2010)

## 2. Suivi de sujets

Comme évoqué déjà à deux reprises dans le document, nous n'avons pas de suivi automatique des projets de lois. Cela n'a jusqu'alors pas posé de problème car cela se fait facilement lorsque les découvertes sont encore limitées. Cependant, lorsque leur nombre va grandir, il sera de plus en plus difficile de tout suivre sans l'aide d'un système. Intéressons-nous dès lors au suivi des sujets, autre branche du *Text Mining* nécessitant le *Web Crawling*.

Google propose un outil gratuit de suivi des sujets, *Google Alerts*, qui permet aux utilisateurs de choisir des mots clés et les informe lorsque des informations relatives à ces sujets deviennent disponibles.

La technologie de suivi des sujets a cependant des limites. Par exemple, si un utilisateur crée une alerte pour le "*Text Mining*", il recevra plusieurs nouvelles sur l'exploitation minière des minéraux, et très peu sur le *text mining* en tant que tel (Gupta et Lehal, 2009). Nous avons essayé nous-même ce service de Google et, par exemple, pour le mot-clé « eu-LISA », nous recevions souvent de l'actualité concernant des personnes s'appelant LISA.

Néanmoins, certains outils d'exploration de texte sont meilleurs et permettent aux utilisateurs de sélectionner des catégories particulières d'intérêt où le logiciel peut même déduire automatiquement les intérêts de l'utilisateur en fonction de son historique de lecture et des informations qu'il a consultées.

Il existe de nombreux domaines dans lesquels le suivi de sujets peut être appliqué dans l'industrie. Il peut être utilisé pour alerter les entreprises chaque fois qu'un concurrent fait



l'actualité. Cela leur permet de se tenir au courant des produits concurrents ou des changements dans le marché.

Les exemples sont nombreux et dans le cas qui nous occupe, cela nous permettrait de nous tenir au courant de l'actualité concernant le contrôle des frontières et la sécurité intérieure dans l'Union.

Concrètement, cela fonctionne en détectant les mots-clés, qui sont un ensemble de mots significatifs dans un article, donnant une description de haut niveau de son contenu aux lecteurs.

L'extraction manuelle de mots-clés est une tâche extrêmement difficile et longue. En réalité, il est presque impossible d'extraire manuellement des mots-clés dans le cas d'articles d'actualité publiés en une seule journée en raison de leur volume. Pour une utilisation rapide des mots clés, nous devons établir un processus automatisé qui extrait les mots clés des articles de presse.

L'architecture du système d'extraction des mots clés est présentée à la Figure 301. Les pages d'actualité en HTML sont collectées à partir d'Internet. Les listes mots-clés candidates sont extraites à l'aide d'un module d'extraction de mots-clés. Enfin, les listes de mots-clés sont sélectionnées par un module de comparaison inter-domaines.

Le module d'extraction de mots-clés fonctionne de la manière suivante ; Nous faisons des tableaux pour "document", "dictionnaire", "occurrences des termes" et "pondération TF-IDF" dans la base de données relationnelle.

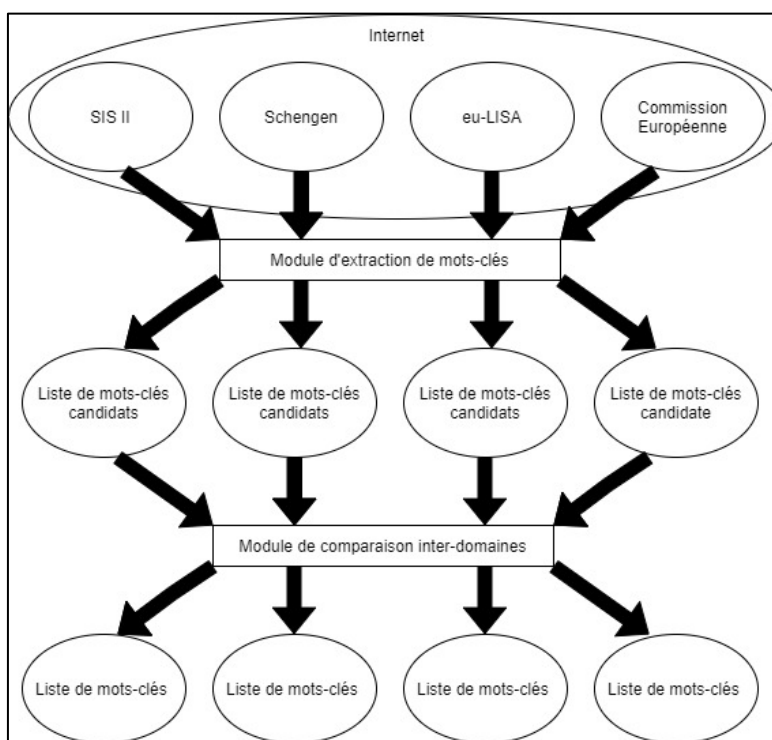


Figure 31 : Architecture du système d'extraction des mots-clés

Source : Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60-76. <https://doi.org/10.4304/jetwi.1.1.60-76>

Dans un premier temps, les documents d'actualité téléchargés sont stockés dans la table "Document" et les noms sont extraits des documents et enregistrés dans la table "Dictionnaire" (Gupta et Lehal, 2009).

Définition de TF-IDF de tfidf.com (Traduction de l'anglais)

TF-IDF est l'abréviation de "*term frequency-inverse document frequency*", et la pondération TF-IDF est une pondération souvent utilisée dans la recherche d'informations et l'exploration de textes. Cette pondération est une mesure statistique utilisée pour évaluer l'importance d'un mot pour un document dans une collection ou un corpus. L'importance augmente proportionnellement au nombre de fois qu'un mot apparaît dans le document, mais elle est compensée par la fréquence du mot dans le corpus. Les variations du schéma de pondération TF-IDF sont souvent utilisées par les moteurs de recherche comme un outil central pour noter et classer la pertinence d'un document en fonction d'une requête de l'utilisateur.

Ensuite, les mots apparaissant dans les documents sont mis à jour dans le tableau "Occurrences des termes". Après cela, les pondérations TF-IDF pour chaque mot sont calculées à l'aide du tableau "Occurrences des termes" et le résultat est mis à jour dans le tableau "pondérations TF-IDF".

Enfin, grâce au tableau "pondérations TF-IDF", nous pouvons créer des listes de mots-clés candidats pour chaque domaine d'information. Les mots qui ont les pondérations les plus hautes sont classés en tête (Gupta et Lehal, 2009). Le module d'extraction des mots clés est présenté à la Figure 35.

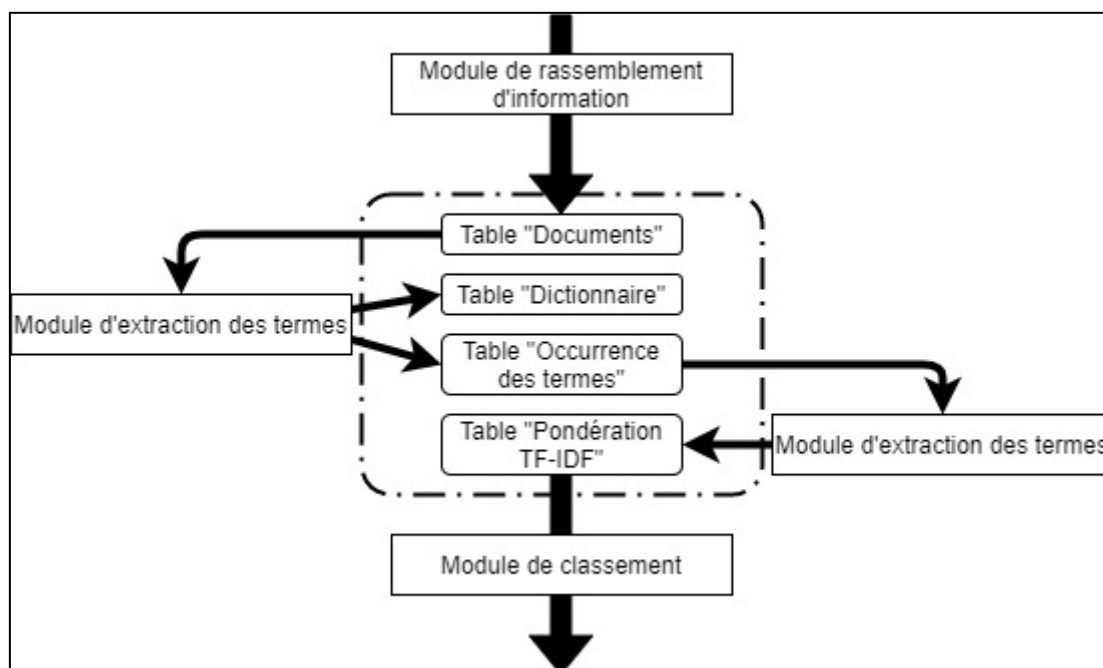


Figure 32 : Module d'extraction des mots-clés

Source : Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60-76. <https://doi.org/10.4304/jetwi.1.1.60-76>

Lorsque ces listes de mots-clés candidats sont obtenues, des règles sont appliquées pour sélectionner les mots-clés qui seront les plus pertinents et donc les documents les plus pertinents.

Pour entraîner les modèles nécessaires au suivi des sujets, il faut un nombre important de textes et c'est pour cela que le *Web Crawling* est utile pour fournir suffisamment de contenu à l'entraînement des modèles.

### 3. Analyse de sentiment/objectivité

Jusqu'à présent, nous nous sommes intéressés au fait que l'information que l'on récoltait soit pertinente et vraie. Toutefois, nous ne nous sommes pas encore posé la question de l'objectivité des documents.

Le rôle d'*EU Monitoring* est de rapporter des faits de manière neutre et sans jugement. Il arrive cependant que certains sites d'information relatent des faits en y introduisant un biais de leur ligne éditoriale. Cela ne veut en aucun cas dire que leurs articles sont dénués d'intérêt mais plutôt qu'il faille y faire attention lorsqu'on souhaite les résumer. En effet, au moment d'utiliser *Resoomer*, il ne faudrait pas que le système garde l'aspect biaisé que présentait le texte original.

Pour nous aider à prévenir les biais dans les documents, nous pouvons, par exemple, nous servir du *Natural Language Processing* (NLP).

Les applications du NLP sont nombreuses et comptent parmi elles celle qui nous intéresse, l'évaluation de l'objectivité d'un texte. Pour cela, l'ordinateur performe ce qu'on appelle l'analyse de sentiment.

Un sentiment peut être une croyance, un point de vue, une opinion ou encore une conviction. L'analyse de sentiment c'est donc se poser la question « Que ressent l'auteur à propos du sujet qu'il traite ? ». Cela peut s'utiliser aussi bien en marketing pour déceler l'avis des clients qu'en politique pour déterminer l'opinion des électeurs (Davis, 2019). Pour nous, cela serait intéressant d'identifier un biais dans un texte afin de ne pas le transmettre dans le contenu d'*EU Monitoring*.

Ce processus s'effectue en quatre étapes reprises à la Figure 36 (Davis, 2019) :

- 1) La première étape s'effectue après que le texte ait été traité et séparé en plusieurs déclarations. Chaque déclaration va donc être soumise à un test à l'aide d'un lexique pour déterminer si elle est objective (O), qu'elle relate des faits, ou si elle est subjective (S) et donc qu'elle offre une opinion. C'est la polarité O-S.
- 2) Si la déclaration est en effet subjective, il faut à présent évaluer si la déclaration est positive (P) ou négative (N). Cette évaluation est à nouveau réalisée grâce à un lexique comprenant des marqueurs (ex. : merveilleux vs insupportable) qui aident à déterminer la polarité N-P.
- 3) Ensuite, la cible de la déclaration est identifiée. La cible peut être une agence (eu-LISA, Frontex, etc), une institution, un groupe politique ou encore un responsable politique.
- 4) Une fois que les polarités O-S et N-P et les cibles de chaque déclaration du texte ont été enregistrées, elles sont agrégées, d'abord par paragraphe et ensuite sur l'ensemble du texte pour avoir le sentiment général du texte.

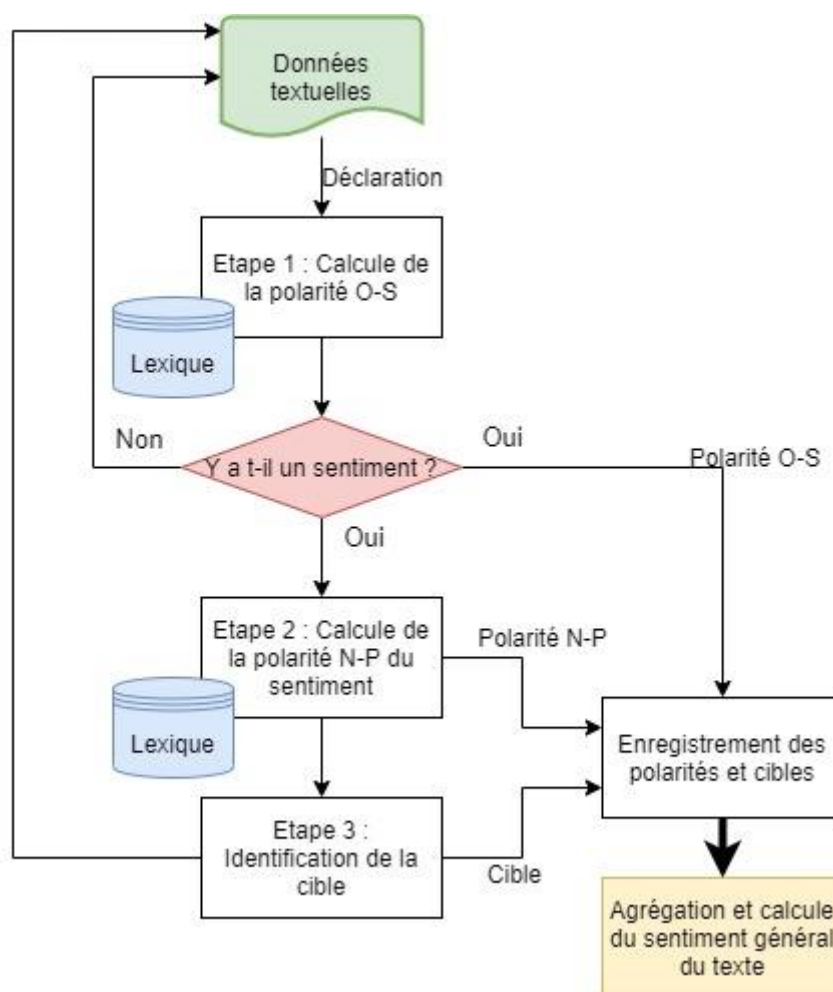


Figure 33 : Analyse de sentiment d'un texte

Source : Davis, V. (2019). *Text Mining, Sentiment Analysis, and Social Analytics* [Diapositives]. eClass.  
<https://eclass.srv.ualberta.ca/my/>

Grâce à ce processus, nous pouvons déterminer quels documents sont biaisés. Nous pouvons alors y prêter davantage attention pour supprimer ces biais et transcrire l'information de la manière la plus neutre possible.

## 4. Voice-to-Text

Plus en avant dans ce chapitre, il était question de traitement de données textuelles. Cependant, les documents analysés dans le cadre d'*EU Monitoring* ne sont pas tous des textes. Certains sont des vidéos ou des enregistrements audios. Pour cette raison, posons-nous la question de savoir en quoi la technologie de *Voice-to-Text* peut nous permettre de traiter ces types de documents comme des textes classiques.

Le principe *Voice-to-Text* s'appuie en réalité sur la technologie de reconnaissance automatique de la parole. Cette technologie, comme beaucoup d'autres branches apparentées à l'intelligence artificielle, a énormément progressé ces dernières années.

Les programmes de reconnaissance vocale font généralement appel soit à la synthèse vocale générée par ordinateur, soit à la parole humaine avec réponse vocale de l'ordinateur, soit aux deux.

La communication humaine est au cœur des développements de la reconnaissance vocale et la complexité du langage rend les approches informatiques de plus en plus difficiles. C'est là que les techniques de *machine learning* telles que le NLP (abordé ci-avant) entrent en jeu. Le NLP ouvre de nouveaux horizons pour améliorer l'interaction entre l'homme et l'ordinateur, il a été une technologie complémentaire pour les processus de reconnaissance vocale, les rendant moins longs et plus faciles (Khillar, 2020).

La reconnaissance automatique de la parole est une technologie qui permet à un ordinateur d'identifier et d'interpréter des mots et des phrases en langage parlé et de les convertir en textes par ordinateur. Il s'agit d'un sous-domaine de la linguistique informatique qui traite des technologies permettant la saisie de la parole dans les systèmes. La reconnaissance automatique de la parole devient rapidement un pilier de l'interaction entre l'homme et l'ordinateur.

Aujourd'hui, les outils de reconnaissance automatique de la parole sont utilisés dans différents types de tâches de dictée, comme la composition d'un message textuel, la lecture de musique via un appareil connecté à la maison, ou les applications de synthèse vocale avec des assistants virtuels. Dans notre cas, cela nous servirait à pouvoir traiter une vidéo comme si c'était un texte. La reconnaissance automatique de la parole est devenue omniprésente aujourd'hui comme un excellent moyen d'interagir avec la technologie, comblant ainsi de manière significative le fossé entre l'interaction homme-machine et la rendre plus naturelle (Khillar, 2020).

De manière simplifiée, un système de reconnaissance automatique de la parole (RAP) se découpe en trois phases principales (Levis et al, 2012).

Tout d'abord, la phase *front-end*, ou traitement acoustique en français, qui permet l'extraction d'un signal vocal. À partir de ce signal, un modèle acoustique est produit et est numérisé.

Ensuite, à l'aide d'algorithmes de *machine learning*, le modèle acoustique est utilisé pour la modélisation statistique, se servant notamment des modèles cachés de Markov (HMM, *Hidden Markov Models*). Cette modélisation statistique permet d'associer les segments élémentaires de la parole et les éléments lexicaux.

Finalement, en utilisant le principe du *pattern matching*, les modèles élémentaires précédemment appris sont concaténés afin de reconstituer le discours le plus probable permettant ainsi le décodage du signal vocal par l'ordinateur.

Le décodage alors réalisé permet de transformer le signal audio en texte et le traitement de ces données textuelles peut alors être réalisé.

Concrètement, la piste pour *EU Monitoring* serait d'utiliser cette technologie pour ne plus avoir de différence entre les documents et tous les traiter de la même manière.

## Conclusion

---

Ce document est un mémoire faisant suite à un stage chez Sopra Steria Benelux et plus précisément au sein de l'agence *Consulting* de l'entreprise. L'objet de ce stage était une gestion de projet dont l'objectif était d'implémenter un service de veille législative européen qui s'appellerait *EU Monitoring*.

Au début de ce mémoire, nous nous étions alors posé la question suivante : « Comment concevoir un système de veille législative européenne efficace ? ».

Dans un premier temps, la question a été abordée de manière théorique. Au cours de cette approche théorique, nous avons appris que, peu importe son application, un service de veille est avant tout un outil d'aide à la décision. Toutefois, bien que ce point soit clair, il est difficile de quantifier clairement l'apport précis d'un service de veille sur la décision finale et donc sur les résultats de l'entreprise. Il semble cependant que les outils de type *Balanced Scorecard* permettent une évaluation relativement pertinente.

Il fallut aussi inclure dans l'approche théorique les spécificités d'une veille de type législatif. Il est bien entendu question de sources juridiques qu'il faut être capable de différencier. C'est pourquoi l'écosystème législatif européen a dû être développé et expliqué. Nous avons constaté que le système législatif européen a ses particularités qu'il est important de connaître dans le cadre d'un service de veille législative.

Une fois cette approche théorique terminée, nous nous sommes intéressés au cas précis d'*EU Monitoring*. Pour cela, des constats ont été tirés de la situation qui ont poussé Sopra Steria à vouloir s'équiper d'un tel service. Ces constats révélaient principalement un manque de visibilité que ce service pourrait combler. Nous avons alors formulé un objectif clair qui nous guiderait au cours de l'exécution du projet. Les acteurs qui seraient amenés à travailler pour le projet ont également été identifiés. Les acteurs sont finalement peu nombreux mais cela a facilité la coordination.

Afin d'avoir une définition des besoins techniques pour *EU Monitoring*, une analyse ADEPT a été réalisée. Cela a permis de déterminer le *business process*, de déterminer comment les données devraient être traitées, par qui, grâce à quels media et quelles technologies. Cette analyse nous a également permis d'aborder ce qui se fait déjà dans d'autres entreprises.

L'analyse ADEPT couplée à l'approche théorique nous a permis de définir une série d'objectifs intermédiaires qui nous permettraient finalement d'arriver à l'implémentation d'*EU Monitoring*.

Ces objectifs ont été les lignes directrices de la gestion de projet et ont été les sujets des réunions hebdomadaires prévues dans notre méthodologie APF (*Adaptive Project Framework*). Cette méthodologie hautement souple nous a permis de revoir les attentes de chaque objectif à la hausse ou à la baisse. Une réévaluation des besoins en temps a également pu être réalisée chaque semaine en concertation avec les sponsors complètement intégrés au projet.

Ensuite, l'exécution en tant que telle du projet a pu avoir lieu. Nous avons commencé par nous concentrer sur la récolte d'information car c'est l'aspect le plus important d'une veille législative.

Nous avons ensuite mis en place une liste de sujets pour améliorer au maximum la pertinence des recherches.

Une fois que nous avons assuré un flux suffisant d'information pertinente, nous pouvions nous focaliser sur l'amélioration de l'exécution des autres étapes du *business process*. En plus de la plateforme installée sur une machine de Sopra Steria le temps que durera la phase de preuve de concept, nous avons incorporé certains outils pour les autres étapes.

A la fin du stage, il a fallu déterminer ce qu'il restait à faire afin que la gestion du projet puisse continuer ensuite. Cette étape est éminemment importante pour que le travail réalisé jusqu'alors ne soit pas inutile et que d'autres collaborateurs sachent directement quelle direction prendre. En suivant les quelques points alors exprimés, *EU Monitoring* devrait être un service opérationnel.

Cependant, bien que quasi opérationnel, le *business process* pourrait être encore plus automatisé en diminuant l'apport humain. En effet, les outils utilisés jusque maintenant sont principalement des outils du Web 2.0. Néanmoins, en se servant d'outils du Web 3.0 aussi appelé Web sémantique, on devrait pouvoir aller plus loin vers l'automatisation d'*EU Monitoring*.

C'est pourquoi nous avons exploré plusieurs pistes répondant à de légères lacunes que l'on pourrait observer dans le business process actuel. Nous sommes convaincus que le travail effectué jusque maintenant couplé à ces technologies du web sémantique peuvent résulter en un système de veille très efficace.



## Bibliographie

---

Aliénor EU. (s. d.). AliénorEU. Consulté le 16 février 2020, à l'adresse <http://www.alienoreu.com/fr/>

AlarmTILT. (s. d.). *Qu'est-ce qu'un Plan de Continuité des Activités (PCA) ?* AlarmTILT®. Consulté 13 août 2020, à l'adresse <https://www.alarmtilt.com/fr/etudes-de-cas/869-qu-est-ce-qu-un-plan-de-continuit%C3%A9-des-activit%C3%A9s-pca>

AQSE-France. (2016). *Veille réglementaire HSE QSE réglementation sécurité et environnement*. aqse-france.com. [https://www.aqse-france.com/securite/fr/veille\\_reglementaire.html](https://www.aqse-france.com/securite/fr/veille_reglementaire.html)

Athenora Consulting. (s. d.). Athenora Consulting. Consulté le 16 février 2020, à l'adresse <http://www.athenora.com/>

Bastien, L. (2018, 7 novembre). *RPA ou Robotic Process Automation : qu'est-ce que c'est, à quoi ça sert ?* Le Big Data. <https://www.lebigdata.fr/rpa-robotic-process-automation-definition>

Bathelot, B. (2016, 28 juin). *Veille concurrentielle*. Définitions Marketing. <https://www.definitions-marketing.com/definition/veille-concurrentielle/#:%7E:text=La%20veille%20concurrentielle%20est%20un,concurrentielle%20consiste%20g%C3%A9n%C3%A9ralement%20%C3%A0%20surveiller%20%3A&text=%2D%20les%20actions%20marketing%20et%20publicitaires,newsletter%2C%20campagnes%20publicitaires..>)

Boumendils et consultants. (2016, septembre 28). Assistance et veille juridique. Consulté le 16 février 2020, à l'adresse <http://www.boumendiletconsultants.fr/prestations/assistance-veille-juridique/>

Commission Européenne. (2016, décembre 5). Organisation de la Commission. Consulté le 16 février 2020, à l'adresse [https://ec.europa.eu/info/about-european-commission/organisational-structure/how-commission-organised\\_fr](https://ec.europa.eu/info/about-european-commission/organisational-structure/how-commission-organised_fr)

Conseil Européen. (2020, 31 janvier). Le processus de prise de décision du Conseil. European Council. <https://www.consilium.europa.eu/fr/council-eu/decision-making/>

Davis, V. (2019). *Text Mining, Sentiment Analysis, and Social Analytics* [Diapositives]. eClass. <https://eclass.srv.ualberta.ca/my/>

Davis, V. (2019, 1 septembre). *Data Mining Process, Methods, and Algorithms* [Diapositives]. eClass. <https://eclass.srv.ualberta.ca>

Davison, L. (2001). Measuring competitive intelligence effectiveness : Insights from the advertising industry. *Competitive Intelligence Review*, 12(4), 25-38. <https://doi.org/10.1002/cir.1029>

Deloitte. (s. d.). Consulté le 16 février 2020, à l'adresse <https://fr.wikipedia.org/wiki/Deloitte>

- DR2 Consultants. (2020, janvier 14). Services. Consulté le 16 février 2020, à l'adresse <https://dr2consultants.eu/work/#policy-monitoring>
- Dutheil, C. (2006, février 20). Veille d'information : Factiva veut élargir sa cible clients. Consulté le 16 février 2020, à l'adresse <https://www.itespresso.fr/veille-dinformation-factiva-veut-elargir-sa-cible-clients-15198.html>
- eu-LISA. (s. d.). Large-Scale IT Systems. Consulté le 16 février 2020, à l'adresse <https://www.eulisa.europa.eu/Activities/Large-Scale-It-Systems>
- European Court of Auditors. (2019). EU information systems supporting border control - a strong tool, but more focus needed on timely and complete data (20). Consulté à l'adresse [https://europarl.europa.eu/meetdocs/2014\\_2019/plmrep/COMMITTEES/LIBE/DV/2020/01-20/SR\\_Border\\_control\\_EN.pdf](https://europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/LIBE/DV/2020/01-20/SR_Border_control_EN.pdf)
- EY (société). (2020, février 7). Consulté le 16 février 2020, à l'adresse [https://fr.wikipedia.org/wiki/EY\\_\(soci%C3%A9t%C3%A9\)](https://fr.wikipedia.org/wiki/EY_(soci%C3%A9t%C3%A9))
- Feedly : organize, read and share what matters to you.* (s. d.). feedly. Consulté 13 août 2020, à l'adresse <https://feedly.com/i/welcome>
- fla-consultants. (s. d.). Veille Réglementaire. Consulté le 16 février 2020, à l'adresse <https://www.flacconsultants.com/fr/veille/veille-reglementaire>
- Frontex. (s. d.). Principales responsabilités. Consulté le 16 février 2020, à l'adresse <https://frontex.europa.eu/fr/nos-activites/principales-responsabilites/>
- Garbade, M. J. (2018, 15 octobre). *A Simple Introduction to Natural Language Processing*. Medium. <https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32>
- Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60-76. <https://doi.org/10.4304/jetwi.1.1.60-76>
- Herring, J.P. (1999). Measuring effective intelligence: meeting the management communication challenge. *Society of Competitive Intelligence Professionals (SCIP)*. [http://www.cireport.com/herring/herring\\_1.html](http://www.cireport.com/herring/herring_1.html)
- Johnson, F., & Kumar Gupta, S. (2012). Web Content Mining Techniques : A Survey. *International Journal of Computer Applications*, 47(11), 44-50. <https://doi.org/10.5120/7236-0266>
- Journal du Net. (2019, 17 janvier). *RSS : qu'est-ce que c'est, comment en profiter ?* JDN. <https://www.journaldunet.com/solutions/dsi/1125572-rss-qu-est-ce-que-c-est/>
- Khillar, S. (s. d.). *Difference Between Speech Recognition and Natural Language Processing*. differencebetween.net. Consulté 13 août 2020, à l'adresse

<http://www.differencebetween.net/technology/difference-between-speech-recognition-and-natural-language-processing/>

KPMG. (2020, février 14). Consulté le 16 février 2020, à l'adresse  
<https://fr.wikipedia.org/wiki/KPMG>

Kumar Pani, S., Mohapatra, D., & Ratha, B. K. (2010). Integration of Web mining and web crawler : Relevance and State of Art. *International Journal on Computer Science and Engineering*, 2(3), 772-776.  
[https://www.researchgate.net/publication/49618606\\_Integration\\_of\\_Web\\_mining\\_and\\_web\\_crawler\\_Relevance\\_and\\_State\\_of\\_Art](https://www.researchgate.net/publication/49618606_Integration_of_Web_mining_and_web_crawler_Relevance_and_State_of_Art)

Leitzelman, M. (2009, octobre). *Etat de l'art et tendances sur le marché de veille et d'intelligence compétitive*. i3S-Sophia Antipolis. <http://www.i3s.unice.fr/~mh/RR/2009/RR-09.14-M.LEITZELMAN.pdf>

Lettre des reseaux. (s. d.). *Contrat-cadre*. Consulté 13 août 2020, à l'adresse  
<https://www.lettreseureseaux.com/P-535-678-P1-contrat-cadre.html>

Levis, J., & Suvorov, R. (2013). *The Encyclopedia of Applied Linguistics - Automatic Speech Recognition*. Blackwell Publishing Ltd. <https://doi.org/10.1002/9781405198431.wbeal0066>

Manager GO ! (2020, 10 février). *Comment faire une veille juridique - réglementaire ? Méthode et outils*. <https://www.manager-go.com/intelligence-economique/veille-juridique.htm>

Mebrahtu, A., & Srinivasulu, B. (2017). Web Content Mining Techniques and Tools. *International Journal of Computer Science and Mobile Computing*, 6, 49-55.  
<https://www.ijcsmc.com/docs/papers/April2017/V6I4201725.pdf>

Nollevaux, G. (2019). Gestion de projets informatiques Gestion de projets digitaux (Notes de cours). Consulté à l'adresse  
[https://moodle.ichec.be/pluginfile.php/26769/mod\\_resource/content/22/ICHEC\\_Gestion\\_Projet\\_Informatique\\_Digit\\_2019\\_2020\\_v43.pdf](https://moodle.ichec.be/pluginfile.php/26769/mod_resource/content/22/ICHEC_Gestion_Projet_Informatique_Digit_2019_2020_v43.pdf)

Parlement Européen. (s. d.). Pouvoirs et procédures. Consulté le 16 février 2020, à l'adresse  
<https://www.europarl.europa.eu/about-parliament/fr/powers-and-procedures>

PriceWaterhousCoopers — Wikipédia. (s. d.). Consulté le 16 février 2020, à l'adresse  
<https://fr.wikipedia.org/wiki/PriceWaterhousCoopers>

Resoomer. (s. d.). *Resoomer / Aide & Support*. Consulté 13 août 2020, à l'adresse  
<https://resoomer.com/help/>

Rouach, D. (1998). *Veille technolog.& intelligence eco. (QUE SAIS-JE ?)*. QUE SAIS JE.

- Roulet, A., Bezençon, C., & Madinier, H. (2015). Évaluation de la performance et de l'impact de la veille. *I2D - Information, données & documents*, 52(3), 70.  
<https://doi.org/10.3917/i2d.153.0070>
- Salzano, G., Ankoud, M., & Desfriches Doria, O. (2013). Apports du SOC d'entreprise Hypertagging aux activités de veille. *Document numérique*, 16(1), 31-53.  
[https://ezproxy.ichec.be:2084/revue-document-numerique-2013-1-page-31.htm?xd\\_co\\_f=YWZmODk2NWEtNTM3Ny00NGZILTgzMmQtNjhiNmUwMTYxY2Y1#](https://ezproxy.ichec.be:2084/revue-document-numerique-2013-1-page-31.htm?xd_co_f=YWZmODk2NWEtNTM3Ny00NGZILTgzMmQtNjhiNmUwMTYxY2Y1#)
- Sombach, R. (2019, 1 septembre). *Investigating Systems Requirements* [Diapositives]. eClass.  
<https://eclass.srv.ualberta.ca/my/>
- Sopra Steria. (s. d.-a). *Chiffres clés*. Sopra Steria Corporate. Consulté 13 août 2020, à l'adresse <https://www.soprasteria.com/fr/investisseurs/a-propos-de-sopra-steria/chiffres-cles>
- Sopra Steria. (s. d.-b). *Notre mission*. Sopra Steria Corporate. Consulté 13 août 2020, à l'adresse <https://www.soprasteria.com/fr/nous-connaître/notre-mission>
- Sopra Steria. (2020a, avril 14). *Histoire du groupe*. Sopra Steria Corporate.  
<https://www.soprasteria.com/fr/nous-connaître/notre-histoire>
- Sopra Steria. (2020b, juin 1). *Sopra Steria Benelux* [Diapositives]. Teams.
- Technopedia. (s. d.). *Web 2.0*. Techopedia.com. Consulté 13 août 2020, à l'adresse <https://www.techopedia.com/definition/4922/web-20#:~:text=Web%20is%20a%20term,the%20web%20decades%20from%20now.>
- Tf-idf : A Single-Page Tutorial - Information Retrieval and Text Mining. (s. d.). tfidf. Consulté 15 juin 2020, à l'adresse <http://www.tfidf.com/>
- ThinkThyme. (2018, 7 avril). *Adaptive Project Framework Methodology*.  
<https://thinkthyme.com/project-management/adaptive-project-framework-methodology#:~:text=Overall%2C%20potentially%20Adaptive%20Project%20Framework,the%20input%20of%20the%20client.>
- ZDNet. (2020, 9 janvier). *Qu'est ce que le RPA ? 4 points pour bien comprendre l'automatisation des processus métier*. ZDNet France. <https://www.zdnet.fr/pratique/qu-est-ce-que-le-rpa-4-points-pour-bien-comprendre-l-automatisation-des-processus-metier-39873151.htm>
- .