

# AMELIORER LA DIFFUSION ET PRESENTATION D'INFORMATION NUMERIQUE COMME ASPECTS QUI FACILITENT L'INTERPRETATION DANS UNE ACTIVITE DE VEILLE STRATEGIQUE

Alex Fernando BUITRAGO-HURTADO (\*), Marie-Laurence CARON-FASAN (\*), Humbert LESCA (\*)

[afbuitragoh@unal.edu.co](mailto:afbuitragoh@unal.edu.co), [marie-laurence.caron@upmf-grenoble.fr](mailto:marie-laurence.caron@upmf-grenoble.fr), [humbert.lesca@upmf-grenoble.fr](mailto:humbert.lesca@upmf-grenoble.fr)

(\*) CERAG-CNRS-UMR 5820, Université Pierre Mendès France Grenoble BP 47 38040 Grenoble Cedex 9 France

## Mots clefs :

Veille Stratégique, information numérique, surcharge d'informations, présentation des informations, diffusion des informations et Internet

## Keywords:

Strategic intelligence, digital information, information overload, presenting information, disseminating information and Internet

## Palabras clave:

Inteligencia estratégica, información digital, sobrecarga informacional, presentación de información, difusión de Información e Internet

## Résumé

Internet est désormais une source importante d'information pour la veille. Toutefois, utiliser ce media dans un processus de veille peut conduire à une surcharge informationnelle dommageable. La surcharge d'information est un concept multi dimensionnel très souvent abordé sous l'angle de la quantité d'information. Nous proposons de l'aborder sous un angle encore peu évoqué de la « facilité d'interprétation ». Ainsi, lutter contre la surcharge d'information c'est d'une part être en mesure de présenter des informations concises, cohérentes et compréhensibles et d'autre part être capable de les diffuser via des canaux appropriés.

Face aux difficultés rencontrées par les managers pour utiliser les informations issues d'Internet dans leur processus de réflexion anticipatif, nous cherchons à répondre à la question de comment présenter et diffuser les informations issues d'Internet pour une activité de veille stratégique. Dans le cadre d'une recherche ingénierique, nous avons conçu et testé l'outil Aproxima dont l'objectif est de collecter, présenter et diffuser des informations utiles au processus de prise de décision dans le cadre d'une activité de veille stratégique. Utilisé au sein du ministère colombien de l'économie<sup>1</sup>, les premiers résultats montrent que les managers abonnés au dispositif Aproxima et qui reçoivent des informations pertinentes sous un format court, sont en mesure de prendre leurs décisions avec une plus grande réactivité.

---

<sup>1</sup> Il est précisé que le contenu de cette communication, publié ou diffusé n'engage que la seule responsabilité de ses auteurs et ne représente pas l'opinion du ministère de l'économie colombien

# 1 Introduction

Les technologies de l'information et de la communication ( TIC ) sont considérées comme une des causes importantes du problème de surcharge d'informations et au même temps sa solution [6] , [32] . Internet a bouleversé le paysage de l'information et de la communication en offrant un accès quasi illimité à l'information [13] et aujourd'hui, une abondance de données (presque une infinité) est disponible. Il est possible d'accéder automatiquement en quelques clics à un univers de publications (livres et journaux) à travers de nombreux appareils électroniques [22] . Toutefois le volume et la diversité croissants des informations numériques est devenue difficile à gérer [21] et les décideurs sont dans une situation permanente de surcharge d'informations exacerbée par une production frénétique de données [21] .

L'Internet a permis une réduction systématique des obstacles de communication remplaçant les intermédiaires qui traditionnellement filtraient les informations et produisaient les résumés pour les décideurs [32] . En conséquence, les décideurs sont obligés de faire face à une charge de plus en plus importante d'information mais dans le même temps, leur capacité innée du traitement de l'information ne change pas, et, il est donc pratiquement impossible de prendre connaissance de toutes les informations reçues [32] . Ainsi, les décideurs se méfient de l'information fournie pour Internet et se mettent à douter de son utilité pour la prise de la décision stratégique [23] .

Les chercheurs du domaine de la veille stratégique reconnaissent Internet comme une source précieuse d'informations pertinentes pour l'anticipation [12] , [14] , [24] , [36] , source qui reste encore largement sous exploitée. Utiliser Internet comme source d'information pour la veille suppose de chercher et sélectionner des informations susceptibles de pallier les asymétries d'information dans la recherche de la performance, c'est-à-dire fournir une information pertinente et compréhensible, au moment approprié dans le processus de prise de décision stratégique [29] . Toutefois, le volume, la variété et la fiabilité des informations issues d'Internet ne sont pas toujours au rendez-vous et cela suppose de recourir à des processus de traitement et transformation numériques [24] , [36] . L'utilisation d'Internet comme source d'information pour la veille semble donc parfois problématique [20] .

Toutefois, l'évolution des technologies de l'information (p.ex. le web 2.0) et à la création des services spécialisés de l'information, contribuent à rendre les informations numériques de plus en plus pertinentes [24] , [35] . Les techniques automatisées de recherche et de sélection des informations ont montré qu'il est possible de filtrer et d'obtenir des données apparemment pertinentes [12] , [35] , [36] , mais, le problème se situe aujourd'hui dans la transformation et la diffusion de ces données afin qu'elles puissent servir à déclencher des actions chez les managers [14] , [35] .

La plupart des dirigeants préfèrent interpréter les données et leur donner un sens de façon autonome. Ils veulent aussi combiner la précision donnée par des intermédiaires humains avec l'opportunité de l'information en ligne [35] . En conséquence, Il semble important d'essayer de leur fournir des informations courtes et concrètes qui donnent l'impression de provenir d'une personne et non d'une machine [35] . La présentation des informations représente donc un enjeu majeur si on veut que l'information issue d'Internet soit utile à la prise de décision stratégique. Les formats appropriés pour les décideurs semblent être de plusieurs ordres : (1) visuel, (2) sonore et (3) messages texte envoyés en ligne [14]

La problématique qui est proposée dans la présente communication est donc la suivante : « *Comment présenter et diffuser les informations issue d'Internet pour une activité de veille stratégique ?* » , et ce compte tenu de l'existence d'outils informatiques qui sont capables d'extraire, de filtrer et de mémoriser des données

apparemment pertinentes [12], [35], [36]. L'apport de cette recherche vise à améliorer la perception des décideurs sur l'utilisation d'Internet comme une source d'informations pertinentes pour l'anticipation leur permettant de bénéficier de cette information sans intermédiaire.

## 2 Cadre conceptuel : concepts d'ancrage de la recherche

### 2.1 La surcharge d'information

Le concept de surcharge d'information est très souvent abordé sous l'angle de la quantité d'information. Ainsi, selon Schick la performance d'un individu varie en fonction de la quantité d'informations auxquelles il a été exposé et du temps dont il dispose [31]. De même, pour Nelson [30], la surcharge d'information est l'impossibilité d'extraire de la connaissance d'un volume considérable d'informations sachant que le *volume* s'entend comme « *a greater number of data, more materials, more items, and more detail. Users are provided with information, both useful and useless. Users must learn quickly how to sort out and choose the useful information* » [34, p.169]

D'autres auteurs, abordent la notion de surcharge d'information sous l'angle conjugué de la quantité d'information et de sa capacité de traitement. Ainsi, selon les études de Schultz [32] la surcharge d'information s'entend comme l'écart existant entre l'information reçue et les limites liées à son traitement. Les limitations du traitement sont associées à la capacité de recherche, de sélection, de classement et d'organisation de l'information. [32]. C'est cette dernière définition que nous retenons pour notre étude ce qui suppose d'aborder la surcharge d'information non pas uniquement sous l'angle de la quantité d'information mais également de la capacité de traitement des informations.

Le traitement de l'information intègre des étapes d'extraction, d'analyse, de présentation, de diffusion et d'interprétation [15] et la surcharge peut être présente à chacune de ces étapes. C'est donc un concept multidimensionnel composé de [5] : (1) la quantité ("*quantity*"), (2) la fiabilité ("*quality*"), (3) la diversité ("*variety*") et (4) les facilités d'interprétation ("*equivocality*") des informations à traiter par un individu, un groupe, système ou une organisation. Cependant, la perception de la surcharge d'information est communément associée aux deux premiers aspects que sont « la *quantité* » et « la *fiabilité* » [16], [30]. Les deux autres aspects de « *diversité* » et des « *facilités d'interprétation* » ont été moins explorés.

Le concept de la « *diversité* » de la surcharge des informations est lié aux différentes sources d'information disponibles [5]. La problématique du nombre de sources nécessaires pour explorer l'environnement est une question sans réponse, compte tenu du nécessaire et subtil équilibre entre diversité et quantité d'informations [11]. En effet, si la cible est trop focalisée, il est possible de perdre des informations pouvant être importantes [25], et par contre, si l'information est trop diverse, il est probable de collecter beaucoup trop d'informations non pertinentes parmi lesquelles, et c'est paradoxale, pourraient tout de même se cacher des informations intéressantes.

Le concept de « *facilité d'interprétation* » de la surcharge d'information, sur lequel notre recherche se focalise, est quant à lui associé à trois éléments (1) la qualité de l'information reçue, (2) l'opportunité [31] et (3) les multiples possibilités d'interprétation [5].

Le premier élément qu'est la « *qualité de l'information* » suppose une information concise, cohérente et compréhensible pouvant améliorer la capacité de traitement individuelle car un décideur est en mesure d'utiliser plus efficacement des informations de bonne qualité que des informations non structurées et/ou peu claires et/ou peu agrégées [15]. La présence d'une information non structurée, implique les mêmes effets négatifs qu'une surcharge d'informations par la quantité. L'information numérique présente cette difficulté car les informations qui sont extraites et filtrées sont généralement peu ou pas structurées. Il est alors nécessaire de les transformer et de les présenter de manière adéquate en fonction des besoins du décideur [35], [36].

Le deuxième élément est « *l'opportunité de l'information* ». La surcharge informationnelle peut subvenir en cas de mauvais choix du canal de communications pour transmettre des informations apparemment pertinentes [10]. De plus, si l'information arrive trop tard, elle est inutilisable et il est donc nécessaire d'éviter que des informations périmées viennent surcharger les individus parce que le tri sur leur obsolescence n'a pas été correctement effectué [14]. Le concept d'opportunité de l'information est également associé à la notion de diffusion de l'information : il n'est pas seulement important que l'information arrive au bon moment, mais aussi qu'elle arrive au bon destinataire [2], [37].

Le troisième élément est la problématique des « *multiples possibilités d'interprétation* ». Une des caractéristiques de l'information stratégique est la difficulté de son interprétation. La surcharge d'information peut provenir de l'existence de multiples explications valides de la même information [19], [20].

Notre recherche se focalise sur les deux premiers éléments du concept de « *facilité d'interprétation* » à savoir la « *qualité de l'information* » et « *l'opportunité de l'information* ».

## 2.2 La veille stratégique

Aguilar [1, p.1] définit la veille comme « *l'acquisition et l'utilisation d'informations à propos d'événements, de tendances et de dynamiques de l'environnement externe dont la connaissance aiderait les managers à orienter le cours de leurs actions futures* ».

La veille est un processus informationnels qui comporte différentes phases mais tous les auteurs ne s'accordent pas à définir le même nombre de phases. Il existe donc plusieurs modèles de veille et nous suivons dans cette recherche celui proposé par Lesca [25, p.10] : la Veille Anticipative Stratégique – Intelligence Collective (VAS-IC). C'est un « *processus collectif et proactif, par lequel des membres de l'entreprise (ou des personnes sollicitées par elle) traquent (perçoivent ou provoquent, et choisissent), de façon volontariste, et utilisent des informations pertinentes concernant leur environnement extérieur et les changements pouvant s'y produire. Dit en d'autres termes, la VASIC est une façon, pour l'entreprise, d'organiser activement sa curiosité vis-à-vis des changements de son environnement dans le but de renforcer sa compétitivité durable* ».

C'est un système d'information d'aide à la décision qui a comme objectif de fournir des informations aux décideurs et de les assister dans leur processus de décision [14]. Cependant, il ne s'agit pas de rechercher des informations pour répondre à des questions posées préalablement, car on ne dispose pas d'un problème identifié à priori [25].

Le modèle VAS-IC se démarque des autres modèles de veille par son caractère anticipatif supposant la détection de changements à venir et notamment d'éventuelles ruptures ou discontinuités susceptibles de survenir dans l'environnement de l'entreprise. Il prend appui sur des informations à caractère anticipatif [25]. Il s'agit

d'informations qui doivent évoquer quelque chose qui n'est pas encore arrivé. « *Elles doivent fournir des éclairages sur le futur, et non pas sur le passé ou le présent. Plus spécifiquement, elles doivent constituer des manifestations précoces d'éventuelles ruptures (ou discontinuités)* ». [27, p.4].

Pour désigner ces informations à caractère anticipatif et les distinguer des autres types d'informations Ansoff [3] parle de « *signaux faibles* ». En détectant de manière précoce les signaux faibles – les faits porteurs d'avenir, il serait possible : (1) d'agir longtemps avant que la menace ne devienne tangible et concrète [3] ; (2) de saisir des opportunités avant que d'autres ne les perçoivent ; (3) de se distinguer et de surprendre les autres [29] .

À l'origine, Ansoff centre la détection des signaux faibles dans un processus systématique d'identification basé sur des sources externes et plus particulièrement des sources humaines (experts, et/ou personnel ayant une grande quantité de contacts externes) [3] . Des recherches récentes [14] , [35] , [36] affirment qu'il est possible d'utiliser des sources alternatives aux sources humaines pour la détection des signaux faibles. Haase [18, p.1652] cite , Internet comme l'une des sources les plus utilisées pour la veille stratégique : “*Overall, we found that the three most frequently employed sources are “customers and suppliers”, “Internet” and “specialized publications”*”.

L'utilisation d'Internet pour la détection des signaux faibles soulève toutefois de nombreux questionnements [20] , [26] . Il semble difficile de détecter des signaux faibles dans la masse d'informations d'Internet car le volume, la variété et la qualité des informations de ce media rendent difficile leur utilisation et requièrent des processus de traitement et de transformation numériques [24] , [35] , [36] . Les nombreuses technologies de l'information disponibles aujourd'hui sur le marché ne répondent pas encore à cette difficulté.

Il semble donc devoir être nécessaire d'effectuer une pré-étape de préparation des informations si l'on veut que celles-ci puissent être utiles et utilisées dans la phase d'interprétation du processus de la VAS-IC. L'objectif est d'être en mesure de présenter et diffuser sur des canaux adéquats des informations qui soient peu nombreuses (en nombre de documents), peu volumineuses (en nombre de mots) et en même temps plus riches en signification (compréhensibles) et plus stimulantes.

## **2.3 La présentation et la diffusion des informations numériques sur Internet**

Les informations collectées via Internet se présentent sous la forme de textes intégraux. Souvent trop nombreux ces textes doivent être triés via une étape de filtrage numérique par mots-clés [12] , [36] afin d'obtenir uniquement des textes pertinents en relation avec le sujet de la recherche [7] , [17] , [33] . Toutefois, ces textes ne peuvent pas encore être considérés comme des informations directement utilisables dans la phase d'analyse de la veille car ils peuvent encore être volumineux (quantité de mots), non codés ou non formatés [8] . Il s'avère donc nécessaire de les présenter dans une forme plus adéquate.

Pour Lesca et Decker et al [12] , [25] les informations de veille et/ou de type signaux faibles doivent se présenter sous la forme de fragments d'informations courts (en nombre de mots) mais en même temps riches en signification et suffisamment stimulants pour permettre aux managers de passer à l'action. Cette transformation de textes ou parties de textes en informations courtes est toutefois considéré comme un processus complexe [35] car il suppose que le système informatique soit en mesure d'interpréter des règles linguistiques (sémantiques et/ou grammaticales) et heuristiques (basées dans l'expérience d'un experte) pour rendre possible la « recherche intelligente » de fragments d'informations noyés dans un texte.

Albright [2] suggère quant à lui quelques points intéressants au sujet de la diffusion de l'information. Tout d'abord, il est nécessaire de s'assurer que les informations auront une destination correcte malgré la difficulté de la dispersion des destinataires dans l'organisation. D'autre part, l'information doit être diffusée par les canaux et dans des formats appropriés aux préférences des destinataires et de leurs habitudes de travail. Enfin, la présentation de l'information doit correspondre aux exigences de présentation de son destinataire. En d'autres termes, l'information de veille doit être pertinente, envoyée à la bonne personne, au bon moment et sous une forme utilisable par son destinataire [37].

Les NTIC ont permis une communication plus efficace, grâce à des outils comme le courriel, la messagerie instantanée, les réseaux sociaux et les outils du web 2.0 [37]. Ces nouveaux canaux de communications représentent une opportunité pour améliorer la diffusion des informations numériques. Prenons l'exemple des alertes d'information. Ce sont des messages courts et pertinents (envoyés uniquement aux abonnés qui le souhaitent) et qui sont diffusés automatiquement. Les alertes sont donc à considérer comme un moyen adéquat pour faciliter la diffusion d'informations numériques [4], [14], [35]. Toutefois, les alertes ne peuvent pas arriver de façon incontrôlée [14] car il n'est pas souhaitable de déclencher trop souvent la réflexion chez les décideurs, ce qui peut provoquer une désaffection dans l'usage [14]. Il est donc nécessaire de savoir choisir le bon canal de diffusion des informations et dans des formats appropriés en tenant compte des préférences de l'utilisateur et de ses habitudes de travail [2].

## **Question de recherche / Problématique de la recherche**

Internet est désormais une source importante d'information pour la veille. Toutefois, utiliser ce media dans un processus de veille peut conduire à une surcharge informationnelle dommageable. La surcharge d'information est un concept multi dimensionnel très souvent abordé sous l'angle de la quantité d'information. Nous proposons de l'aborder sous un angle encore peu évoqué de la « facilité d'interprétation ». Ainsi, lutter contre la surcharge d'information c'est d'une part être en mesure de présenter des informations concises, cohérentes et compréhensibles et d'autre part de les diffuser via des canaux appropriés.

Face aux difficultés rencontrées par les managers pour utiliser les informations issues d'Internet dans un processus de réflexion anticipatif, notre question de recherche est la suivante :

<b>Comment présenter et diffuser les informations issues d'internet pour une activité de veille stratégique?</b>
--

## **3 Méthodologie de la recherche**

### **3.1 Méthode de recherche ingénierique**

Le projet de recherche s'inscrit dans une épistémologie constructiviste autour de la méthode de recherche ingénierique [9]. Cette méthode de recherche est, par certains de ses aspects, proches de la recherche action pris dans une acceptation large. Elle a comme double objectif de faire avancer les connaissances fondamentales des chercheurs en résolvant des problèmes des utilisateurs par la construction de connaissances actionnables. Toutefois, la recherche ingénierique se différencie de la recherche action par l'activité d'ingénierie et de construction qu'elle suppose [9]. Le chercheur, envisagé comme un « chercheur ingénieur »,

conçoit son modèle conceptuel, construit l'outil support de sa recherche et agit à la fois comme évaluateur et animateur dans sa mise en œuvre dans les organisations. Il contribue ce faisant à une meilleure connaissance des processus organisationnels complexes et à l'émergence de connaissances scientifiques nouvelles.

L'outil support à notre recherche est l'outil Aproxima que nous présentons dans les paragraphes suivants. Conçu par un des auteurs de ce texte, il a été utilisé comme outil de collecte de données dans le cas d'une expérimentation en entreprise. Nous présentons dans cette communication la version V2 de l'outil Aproxima qui avait tout d'abord été construite dans sa version V1 comme outil de collecte et de filtrage de données puis amélioré pour devenir également un outil d'interprétation des informations préalablement filtrée.

### 3.2 AproximaV1: outil de collecte et de tri de données numériques

AproximaV1 est né en réponse de savoir « *comment automatiser la recherche de données FULL text à caractère anticipatif sur l'Internet, ainsi que le passage de ces données brutes vers des informations plus concises et plus brèves utilisables dans la phase d'analyse du processus de veille anticipatif* » [8, p.5] .

AproximaV1 permet la mise en œuvre d'une procédure adéquate pour la recherche automatique d'informations courtes (dites *brèves*) au sein d'un gros volume de données numériques brutes [8] . On entend par « brèves » [28, p.223] « *une information ramenée à ses mot essentiels de façon à être très courte. Cette contrainte de taille résulte du fait qu'une brève est destinée à être projetée sur un écran en compagnie d'autres brèves. Ces brèves ainsi rapprochées aident à faire émerger du sens au cours d'une séance de création collective de sens, au moyen de la méthode Puzzle®. Une brève résulte de l'extraction de quelques mots essentiels à partir d'une donnée généralement documentaire pouvant compter parfois plusieurs pages* ».

Grâce à l'extraction d'informations de type brèves, AproximaV1 a pour finalité de réduire la surcharge de données engendrée par l'Internet. En d'autres termes, l'objectif d'AproximaV1 est de réduire la surcharge d'information liée au concept de « *quantité* ». [16] , [30] .

AproximaV1 est constitué d'un ensemble d'outils informatiques hétéroclites (agrégateurs de flux de données, logiciel d'aspiration de contenus et d'analyses sémantique) mais qui ont été intégrés. Il permet de collecter des données à caractère anticipatif textuelles numériques sur Internet (pages Web des journaux locaux et nationaux, Blogs et sites) grâce à des techniques associées aux flux RSS et à l'aspiration de contenus. Les données sont stockées dans une base centrale et unique dans laquelle sont effectuées les recherches. L'exploitation des informations se fait par des analyses linguistiques [8] .

AproximaV1 a été mis en œuvre à plusieurs reprises et sur différents sujets et par exemple sur les sujets de « *la chimie verte* » [8] , le « *CO<sub>2</sub>/valorisation* » et « *DANONE* » [28] .

La mise en œuvre d'AproximaV1 sur les sujets précédents a permis d'obtenir deux résultats principaux. D'une part, il est possible d'automatiser la recherche de brèves et d'autre part, il est envisageable d'identifier le caractère anticipatif d'une brève [28] .

Plusieurs difficultés ont cependant émergées lors de l'utilisation d'AproximaV1. Tout d'abord, le problème *d'intermédiation* pour utiliser le logiciel : des interventions humaines encore trop nombreuses sont nécessaires et il s'avère que la mise en œuvre de l'outil suppose un niveau minimum de connaissances informatiques de la part de l'utilisateur afin de pouvoir gérer les interfaces de communication. D'un autre côté, la problématique de la diffusion des informations à

fait son apparition. Il s'est avéré incontournable de devoir réfléchir à la problématique de la diffusion des *brèves* à plusieurs personnes, au moment opportun et sur différents sujets.

Ce dernier élément a limité et freiné l'utilisation d'AproximaV1 et a motivé l'ajout de nouvelles fonctionnalités. Ces dernières reprennent les recommandations de Duan [14] sur le logiciel « *idéal* » pour supporter un processus de la prise de la décision.

### **3.3 AproximaV2 : outil de collecte, de présentation des informations pertinentes et de diffusion aux destinataires.**

#### **3.3.1 Caractéristiques d'un logiciel « idéal » de support à la décision**

Une problématique fréquente dans la conception des logiciels d'aide à la décision est le manque de critères pour que la présentation des informations stratégiques soit concise, courte et significative, compte tenu du peu de temps des décideurs et du fait qu'il n'existe pas d'intermédiation humaine [35]. Selon l'étude de Duan [14], il existe trois groupes de caractéristiques essentielles à considérer pour l'acceptation d'un système d'information informatisé pour les décideurs : (1) l'usage, (2) l'autonomie et (3) l'adaptabilité.

Le concept **d'usage** ramené au contexte de la présentation de l'information est une mesure de la corrélation existant entre les informations fournies et les besoins spécifiques de l'utilisateur. En conséquence, ce que les décideurs veulent : ce sont des règles simples permettant de chercher et de trouver des informations pertinentes facilement compréhensibles et adaptées au contexte de l'organisation.

Le concept **d'autonomie** est associé à la mise en jour permanente de l'information. Le système doit permettre l'implémentation de procédures automatisées pour obtenir une information constamment à jour compte tenu des changements permanents de l'environnement de l'organisation. L'autonomie peut procurer deux avantages : (1) la spontanéité qui est la capacité de percevoir les changements de l'environnement et d'y répondre librement et sans contraintes ; (2) la proactivité qui est la capacité à anticiper les changements et agir de manière opportune.

Enfin, le concept **d'adaptabilité** qui suppose que le système soit capable d'apprendre les préférences et les habitudes d'un utilisateur et de s'adapter à l'évolution de ses besoins. Ce concept d'adaptabilité mentionné par Duan n'a pas d'utilité au stade actuel de notre recherche. Il pourra devenir une piste de recherche ultérieure.

Nous présenterons ci-dessous une nouvelle version dite AproximaV2 implémentant des caractéristiques d'usage et d'autonomie. L'objectif est tester ces nouvelles caractéristiques permettant d'améliorer la qualité et l'opportunité de l'information numérique pertinente.

#### **3.3.2 Présentation d'AproximaV2**

AproximaV2 est conçu pour fournir et distribuer des brèves de façon autonome et en temps réel. Grâce à la combinaison de plusieurs technologies de type Web qui fonctionnent harmonieusement, le logiciel peut travailler de façon distribuée pour envoyer des informations ciblées aux destinataires clés directement sur des supports électroniques (par exemple *les smartphones*). AproximaV2 est représenté dans la Figure 1. Les deux premières étapes sont des étapes de paramétrage de



l'outil qui nécessitent une intervention humaine. Cependant, cette intervention n'est pas d'ordre technique, au contraire, le paramétrage peut être effectué par un utilisateur non expert en informatique. Ensuite, les étapes 3, 4 et 5 réalisées par l'ordinateur s'effectuent périodiquement et de façon autonome. Au final, AproximaV2 est capable de proposer en « output » des informations compréhensibles et concises. La diffusion de ces informations est possible grâce à l'utilisation d'alertes programmées. Le destinataire de l'alerte peut avec ces données faire une première interprétation des informations. L'outil AproximaV2 devrait ainsi être en mesure de stimuler et d'aider les managers à donner du sens aux informations ainsi collectées.

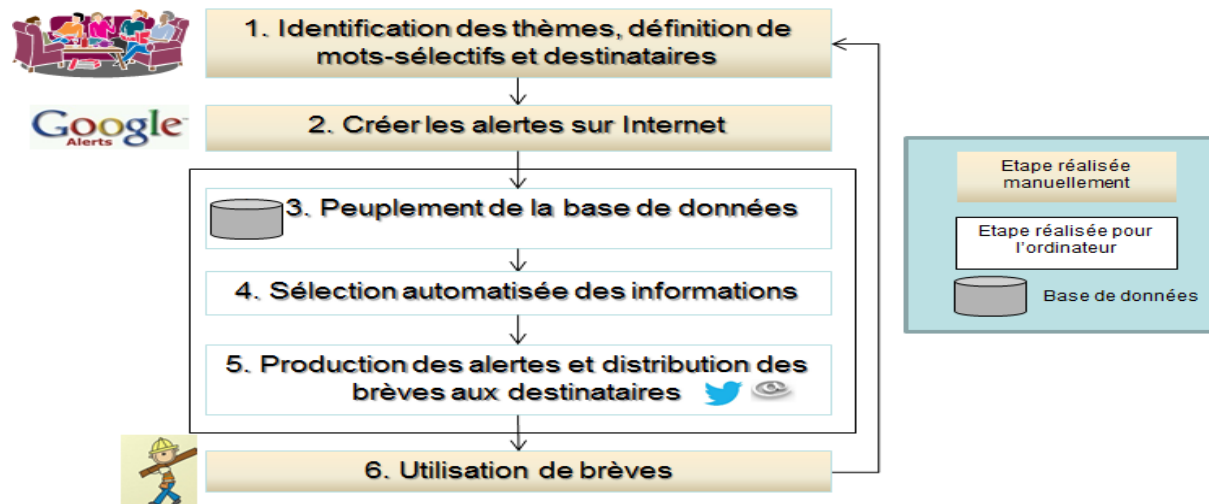


Figure 1 AproximaV2 : outil de collecte, de présentation des informations pertinentes et de diffusion aux destinataires

AproximaV2 s'exécute en 6 étapes comme présentées ci-dessous :

### Etape 1. Indentification des thèmes, définition de mots sélectifs et destinataires (humain)

La première étape est le paramétrage d'AproximaV2. Il est nécessaire d'enregistrer dans le logiciel les informations suivantes :

- Le ou les thème(s) de la recherche c'est-à-dire un ou plusieurs centre d'intérêt relatif au devenir de l'entreprise.
- La définition des mots sélectifs permettant de spécifier un thème de manière précise [8].
- Les bénéficiaires de l'information c'est-à-dire des personnes physique ou morale décisionnaire interne (parfois externe) à l'organisation dont les décisions et les actions sont susceptibles d'avoir, dans le futur proche ou plus éloigné, une influence (positive ou négative, selon le cas) sur le devenir de son organisation ou des certaines de ses activités [25]

## Étape 2. Créer les alertes sur Internet (humain)

La deuxième étape du paramétrage est la création d'alertes sur Internet. Cette activité est externe à AproximaV2 car elle utilise le service d'alertes de Google<sup>2</sup>. Les alertes de Google permettent d'avertir par différents moyens électroniques (par exemple le courriel ou les flux RSS) l'existence de nouveaux résultats correspondant aux mots sélectifs préalablement identifiés. Les résultats des alertes sont récupérés pour AproximaV2 grâce aux services d'Internet construits pour cette finalité. La création de l'alerte est simple et ressemble à une recherche sur le moteur de Google, il n'est donc pas nécessaire de posséder une connaissance informatique pour réaliser cette activité. Les alertes de Google sont créées une seule fois puis elles envoient automatiquement les données collectées au service d'AproximaV2. Elles restent en vigueur tant que les thèmes ne changent pas.

## Étape 3. Peuplement de la base données (automatique)

AproximaV2 utilise un service d'Internet pour recouper les alertes fournies par Google. Ce service est autonome et fonctionne sur une infrastructure d'administration de tâches de Windows<sup>3</sup>. Le planificateur de tâches est un logiciel du système d'exploitation de Windows permettant de créer une tâche qui exécute un programme périodiquement en fonction d'un planning. Les données récupérées sont stockées dans la base de données. Ces données ne sont pas encore les textes intégraux, ce sont des données permettant l'identification de la source et des liens d'accès au texte intégral.

## Étape 4. Sélection automatisée d'Information (automatique)

Après avoir stocké les données, il est possible de les filtrer. Cette étape est automatisée grâce au planificateur de tâches de Windows. L'étape de sélection comprend les activités suivantes :

**Extraction de textes intégraux.** L'extraction consiste en l'aspiration du contenu puis le nettoyage des données. Pour cela, AproximaV2 traite chaque donnée afin d'isoler le contenu textuel porteur d'informations pertinentes des autres contenus de la page internet comme par exemple la publicité, les liens aux données ou aux autres sections [8].

**Filtrage des données.** Le filtrage des données utilise des techniques d'analyse linguistique [8]. AproximaV2 s'appuie sur le logiciel Unitex<sup>4</sup> qui grâce à l'utilisation de patrons linguistiques, est capable d'identifier les phrases dans les textes intégraux. Ces phrases nécessaires à la réalisation de l'étape suivante, sont sélectionnées et isolées grâce aux mots sélectifs définis dans l'étape 1.

Ainsi, le filtrage des données se fait d'une part sur la détection de mots anticipatifs (qui concernent le futur, par exemple les mots « *projet* », « *recherche* », « *nomination* » etc.) et sur l'identification des verbes conjugués au futur. A l'issue de ce travail, AproximaV2 est capable de fournir des brèves qui sont ensuite classées par thème puis stockées dans la base de données. De même, les textes intégraux contenant de brèves sont identifiés pour faciliter leur future récupération dans les étapes ultérieures

---

<sup>2</sup> Service gratuit de Google accessible sur <http://www.google.fr/alerts>

<sup>3</sup> Le fonctionnement pour planifier une tâche sur Windows est décrit dans la page <http://windows.microsoft.com/fr-ca/windows-vista/schedule-a-task>

<sup>4</sup> Unitex est un ensemble de logiciels libres permettant de traiter des textes en langues naturelles en utilisant des ressources linguistiques. Ces ressources se présentent sous la forme de dictionnaires électroniques, de grammaires et de tables de lexique-grammaire. (<http://www-igm.univ-mlv.fr/~unitex/>)

### Etape 5. Production d'alertes et diffusion des brèves aux destinataires (automatique)

AproximaV2 transforme l'information présente dans la base de données textuelle en deux types de documents : (1) des messages courts (144 caractères maximum) et (2) des résumés. Ces documents sont stockés grâce à un service Web autonome et sont formatés pour leur diffusion comme l'illustre les deux exemples ci-dessous.

**Exemple 1 : Messages Courts.** L'information est limitée à 144 caractères afin de faciliter sa diffusion à travers un réseau social auquel les décideurs peuvent s'abonner. Nous utilisons le format proposé par le réseau Twitter comme l'illustre la Figure 2



Figure 2 Message court sous format Twitter

- Les thèmes. Il est possible que le message soit associé à un ou plusieurs thèmes. Le thème est identifié dans le message par le symbole # appelée *hashtag*. Créé par Twitter comme moyen permettant de catégoriser les messages<sup>5</sup>, ce symbole permet la catégorisation des messages d'AproximaV2 de façon chronologique.
- La date. La date affichée correspond à l'heure exacte de publication du message court sur la plateforme de Twitter. Il est utilisé pour faciliter l'ordre chronologique des messages.
- La brève. La brève résulte de la sélection automatisée (étape numéro 1).
- La source. La source est le lien de la page du site internet contenant le texte intégral.

La diffusion des messages courts sur Twitter utilise un service Internet. Suite à la sélection automatisée, les brèves stockées sont agrégées par thème. AproximaV2 construit le message court de la Figure 2, et grâce à une application développée sur Twitter ce message est publié sur la plateforme d'alertes. Twitter avise automatiquement l'arrivée d'un nouveau message aux abonnés.

**Exemple 2 : Résumé.** Un résumé comporte un certain nombre d'informations (date, titre, source et message) pour un thème donné et dans une période de temps programmé. La structure du résumé est préparée par être envoyée automatiquement par courriel comme le montre la Figure 3.

<sup>5</sup> L'explication sur l'utilisation des *hashtag* est disponible dans la page <https://support.twitter.com/articles/231414-en-quoi-consistent-les-hashtags-symboles#>

Le thème

Resultados sobre el tema de investigación Findeter · Proyectos Fecha de Publicación: 25/04/2013 12:53:21 PM

Palabras Selectivas: findeter(Obligatorio), [urrá], \*alerta, \*corrup, \*denun, \*despilf, \*desviac, \*evaluac, \*finanza, \*incorrec, \*indebid, \*infi, \*malver, \*recurs, \*regalia, \*sospecho, \*transfere, acueducto, agua, agua potable, amenaza, apoyo fiscal, avaluo, cedelca, ciudad, compensac, contingencia, contrabando, contralor, credito, crisis, DAF, decreto 028, decreto 28, defensa judicial, deficit, departament, descontento, educa, electric, elusion, embargo, endeuda, evasion, excedente, financia, fiscal, fondo, fonpet, giro, hacienda, hospital, impuesto, infraestruct, inveral, ley 358, ley 550, litigio, loteria, mineria ilegal, monitoreo, municip, obligacion, pasivo, pension, plan, plan de desarrollo, predial, prestamo, presupuesto, programa, proyecto, queja, quiebra, reclam, riesgo, salud, saneamiento fiscal, sentencia, SGP, sistema general, sostenib, territori, universidad, viabilidad, vigencia futura, viviend

Por favor no contestar al presente correo. Consulte la base de datos de Noticias en [www.minhacienda.gov.co/news](http://www.minhacienda.gov.co/news) o suscribase el periódico de la Dirección General de Apoyo Fiscal del Ministerio de Hacienda y Crédito Público <http://paper://SentalesDebiles?1338483399>.

Palabras Clave: Manizales, TIM, Cortés Carrillo, Findeter, SETP, Cereté, Colpatría, Armenia, Luis Fernando, Arboleda, BID, Roberto Boenheim Bernal

Frases

25/04/2013

Título(s) : Concejo de Manizales negó liquidar TIM, Temas destacados

Autor(es) : GOOGLE NOTICIAS

También estuvo de acuerdo Cortés Carrillo, quien manifestó que el proyecto deben correrlo, pero expuso que si se liquida ¿quién asumirá las deudas de TIM y quién quedará cumpliendo con el objeto social?, esto por las acreencias, solo con Findeter en de \$13 mil millones, y por las demandas que se vendrán de los actores del Sistema, que no recibieron lo que les propusieron ganar con el SETP.

Texto Completo:

[GOOGLE NOTICIAS - Noticia Temas destacados](#) , [GOOGLE NOTICIAS - Noticia Concejo de Manizales negó liquidar TIM](#)

24/04/2013

Título(s) : Ya se ven los edificios

Autor(es) : El Meridiano de Córdoba

Dicho proyecto es adelantado por Findeter v la Alcaldía de Cereté, siendo el constructor la firma Colpatría. "Las primeras 112 viviendas tentativamente serán entregadas el 26 de junio.

Texto Completo:

[El Meridiano de Córdoba - Noticia Ya se ven los edificios](#)

24/04/2013

Título(s) : Armenia estará en el proyecto ciudades competitivas

Autor(es) : La Cronica del Quindío

Luis Fernando Arboleda, presidente del Financiera de Desarrollo Territorial Findeter, aseguró en un medio nacional que a mediano plazo la capital quindiana entrará a hacer parte del programa ciudades competitivas que coordina la entidad con el apoyo del Banco Interamericano de Desarrollo, BID.

Según Roberto Boenheim Bernal, gerente de la zona Eje Cafetero de Findeter, la vinculación al proyecto representará un gran paso para la capital quindiana. "

Texto Completo:

[La Cronica del Quindio - Noticia Armenia estará en el proyecto ciudades competitivas](#)

Textos Tratados: 8

Noticias de interés: 3

La date

Le mots sélectifs

Le mots clés

Le message

La source

Figure 3 Exemple d'un « Résumé »

- Le thème. Le résumé concerne un seul thème.
- Les mots sélectifs. Liste des mots utilisés pour filtrer l'information
- Les mots clés. Les mots clés émergent de l'analyse linguistique et correspondent à des noms propres de personnes pouvant représenter un intérêt pour le décideur.
- Le message. : Le message est ici une métadonnée comprenant (1) la date de réception, (2) le titre, (3) l'auteur et (4) la brève
- La source. La source est le lien de la page contenant le texte intégral et la localisation de la brève dans le texte intégral.

Les résumés sont diffusés automatiquement (en fonction d'une périodicité fixée à l'avance) par courriel aux destinataires identifiés dans l'étape 1 de paramétrage. Plusieurs thèmes pourront être envoyés à plusieurs destinataires.

### **Etape 6. Utilisation des brèves (humaine)**

La pertinence de la brève est jugée par le décideur qui peut agir de trois manières différentes :

- retenir les brèves qui l'intéressent parmi celles qu'il vient de recevoir et agir en fonction de cette information.
- chercher à obtenir des explications. La brève peut être intéressante et nécessiter une analyse plus approfondie. Le décideur peut appeler un expert (ou groupe d'experts) pour l'aider dans l'interprétation. Il peut également être amené à demander la création d'un nouveau thème.
- prendre connaissance de la situation et ne rien faire

## **4 Expérimentation : Cas du ministère de l'économie en Colombie**

### **4.1 Contexte et enjeux de l'expérimentation**

Le ministère de l'économie de Colombie, à travers sa Direction Générale d'Appui Fiscal (DGAF), a pour tâche la coordination de la surveillance, l'exercice du suivi et le contrôle de l'utilisation de ressources financières des territoires (départements et agglomérations). Il a également été chargé d'impulser une culture de la responsabilité fiscale.

Le contexte de la Colombie est singulier avec une situation géographique particulière où les trois quarts de la population habite dans le centre du pays à côté de la cordillère des Andes. La distribution des ressources ne dépend pas de la taille de la population de chaque région et ces ressources ne sont pas réparties de manière équitable. En Colombie, la distribution des ressources financières appartient aux régions qui dépendent en grand partie de l'exploitation des ressources naturelles, comme le pétrole, le gaz naturel, le charbon, l'or, le nickel et les émeraudes. Lorsque des ressources sont exploitées, la plus grande partie du profit est censée rester dans ce territoire en tant que rétribution de cette exploitation.

Malheureusement, en regardant l'histoire de la Colombie, les ressources n'ont pas été bien réparties ni réinvesties. Le constat est que les territoires qui reçoivent chaque année la plus grande quantité de ressources, sont en même temps les plus pauvres : par exemple le « *Chocó* » dans le pacifique colombien où se trouvent les mines d'or les plus importantes de la Colombie. Dans cette région, les conditions de vie de la population sont déplorables, l'accès à l'eau potable est inexistant, l'électricité n'est un service que pour les plus riches et les services de santé et d'éducation brillent par leur absence.

Dans ce contexte, une demande permanente des cadres de la DGAF est d'obtenir des renseignements sur les possibles situations d'irrégularité dans la redistribution et l'appropriation des ressources publiques notamment dans les régions. La directrice d'appui fiscal observe que: « *En fait, nous devons travailler pour protéger les ressources publiques attribuées aux régions. Ces ressources doivent être utilisées correctement pour les services basiques concernant la santé, l'éducation, l'eau potable ainsi que les autres services qui sont sous la responsabilité directe des territoires. Mais l'information nous arrive trop tard pour agir et nous nous trouvons dans une situation de correction permanente basée uniquement sur de l'information financière* ».

Pour accomplir sa mission la DGAF dispose d'un groupe d'experts dans les différents thèmes sectoriels (santé, éducation, etc.) et juridiques ainsi que de responsables territoriaux. Ces derniers (au nombre de 40) ont comme fonction principale d'apporter une assistance technique financière aux entités territoriales dont ils ont la charge.

En Colombie, les journaux locaux représentent une source particulièrement riche d'informations d'autant plus que la plupart sont numérisées sur Internet. Dans ces journaux, on peut trouver des informations intéressantes et notamment des articles rédigés par des journalistes de terrain et proches des régions. On trouve également des informations concernant la mise en place de nouveaux projets qui pourront être qualifiés comme projet à risque à la lumière du bilan financier d'une entité territoriale (département, grande ville ou les villages).

L'enjeu pour la DGAF est donc d'utiliser ses sources électroniques en complément de celles déjà mobilisée afin d'être en mesure d'améliorer leur capacité préventive.

Pour aider la DGAF dans cette tâche, nous avons mis en place l'outil AproximaV2 entre le 14 mai 2012 et le 24 avril 2013. Nous avons identifié comme utilisateurs de l'outil 43 personnes ayant deux profils.

*Tableau 1 Les utilisateurs participant dans l'expérimentation.*

Profils	Description	Information attendue	Nombre d'utilisateurs d'ApproximaV2
<b>Cadres de la DGAF</b>	Directrice et directeurs adjoints. Ils sont les responsables de la stratégie de suivi et de contrôle de l'utilisation de ressources financières des territoires.	Information transversale sur les problèmes qui pourraient affecter une entité territoriale. L'information transversale correspond aux mégaprojets régionaux, aux projets de loi nationaux et infranationaux, aux indices de corruption ou aux problèmes liés à la mauvaise utilisation des ressources etc.	3
<b>Responsables territoriaux</b>	Le responsable territorial apporte une assistance technique financière aux entités territoriales. Ils fournissent également des informations aux cadres de la DGAF sur de possibles problèmes affectant la capacité financière d'une entité. Un responsable territorial peut avoir plus d'une entité assignée.	Informations sur les projets en cours, les plaintes des citoyens et les potentiels problèmes liés à la mauvaise utilisation des ressources.	40

## 4.2 Utilisation d'ApproximaV2

Nous reprenons ainsi les 6 étapes de la Figure 1 en les illustrant.

### Étape 1. Identification des thèmes, définition de mots sélectifs et destinataires (humaine)

Les nouvelles sont classées par thème:

- Régions : Bogotá, Antioquia, Caldas, Risaralda, Atlántico, Bolívar, Córdoba, San Andrés y Providencia, Sucre, Arauca, Boyacá, Casanare, Amazonas, Caquetá, Cauca, Chocó, Guainía, Guaviare, Nariño, Putumayo, Quindío, Valle del Cauca, Vaupés, Cundinamarca, Meta, Vichada, Cesar, La Guajira, Magdalena, Norte de Santander, Santander, Huila, Tolima
- Sources de Financement : Regalías, Sistema General de Participaciones
- Projets : Ciudades Amables, hospitales
- D'autres: MHCP, Apoyo Fiscal.

Les bénéficiaires de l'information sont les responsables et les cadres. La DGAF nomme un responsable dans chaque région pour surveiller sa solidité financière. Le responsable présente tous les six mois un rapport sur la situation de l'entité, il peut toutefois être amené à diffuser des informations plus fréquemment s'il collecte des informations montrant que la stabilité d'une région pourrait être affectée.

Pour chaque thème, des mots-sélectifs ont été définis. Par exemple, pour le thème "source de financement (*regalias*)", les mots sélectifs sont<sup>6</sup> : "regalia(Obligatorio), regalia(Obligatorio), SGR(Obligatorio), ^alerta, ^corrup, ^denun, ^despilf, ^desviac, ^distribuc, ^evaluac, ^finanza, ^incorrec, ^indebid, ^intervent, ^invers, ^malver, ^recurs, ^sospecho, ^transfere, <Hum>, <invertir>, <Region>, <Ville>, comision rectora, departamento, DNP, mesa de trabajo, MHCP, ministerio de hacienda, ministerio de minas, municipio, ocad, organo colegiado, organos colegiados, primer giro, queja"

## Étape 2. Créer les alertes sur Internet (humaine)

L'utilisation de la plateforme de Google, permet la création d'alertes en fonction de mots sélectifs. L'enregistrement utilise la même structure que celle du moteur de recherche de Google. Dans notre exemple, et pour le thème « Source de financement », la requête est : « *regalia OR SGR) AND ("comision rectora" OR departamento OR DNP OR "mesa de trabajo" OR MHCP OR "ministerio de hacienda" OR "ministerio de minas" OR municipio OR ocad OR "organo colegiado")* ».

## Étape 3. Peuplement de la base données (automatique)

Entre le 14 mai 2012 et le 24 avril 2013, 432 871 textes intégraux ont été collectés et stockés, avec une moyenne de 1 272 textes intégraux par jour.

## Étape 4. Sélection automatisée d'Information (automatique)

Le nombre des textes sélectionnés par AproximaV2 comme étant pertinents a été de 35 722. La moyenne de textes intégraux jugés comme pertinents est de 105 par jour, soit 8% de la totalité des textes peuplant la base de données. **En moyenne pour les 39 thèmes configurés dans AproximaV2, les destinataires de l'information ont reçu entre 3 et 4 brèves par jour et par thème.**

<sup>6</sup> Le format utilisé pour le paramétrage des mots sélectifs est le format POSIX. Ce format permet la création des filtres morphologiques à partir de la racine d'un mot. Par exemple s'il est nécessaire d'extraire tous les mots commençant par la lettre « a » nous devons utiliser l'expression ^a. Il y a des filtres morphologiques pour les verbes, les adjectifs, les prépositions les groupes de mots (par exemple toutes les villes <Ville>). L'information sur le format POSIX est disponible sous <http://www-igm.univ-mlv.fr/~unitex/UnitexManual2.1.pdf>. Page 77.

## Étape 5. Production des alertes et diffusion des brèves aux destinataires (automatique)

Concernant le thème « Régions », chaque responsable abonné a reçu par courriel et tous les jours entre midi et une heure de l'après-midi, un résumé composé des brèves d'un jour et demi sous le format présenté dans la Figure 3. Les autres thèmes ont fait l'objet de diffusion de messages courts par Twitter. A ce propos, le Ministère d'économie a implémenté un compte Twitter nommé « *señales débiles* » (signaux faibles) - accessible sur <https://twitter.com/SenalesDebiles>.

## Étape 6. Utilisation des brèves (humain)

Nous avons pu relever deux cas d'utilisation :

- 1- Celui des responsables de région. Lorsqu'une brève est analysée et si une situation inconnue ou contradictoire est décelée, ceux-ci déclenchent une action de recherche d'informations supplémentaires sur le terrain. En fonction des résultats obtenus et s'ils le jugent nécessaire les responsables prennent la décision d'alerter les dirigeants de la DGAF.
- 2- Celui des cadres de la DGAF : ils reçoivent des informations différentes de celles reçues par les responsables. En effet, les brèves envoyées aux cadres sont relatives à des thèmes transversaux. Dans le cas où un cadre de la DGAF juge une brève comme importante, la procédure est simple. En premier lieu, il se renseigne auprès des experts qui sont à disposition à la direction. La communication est généralement téléphonique et en fonction de l'avis et du degré de surprise de l'expert, le cadre de la DGAF prend une des trois actions possibles :
  - la situation est inconnue mais gérable, le cadre renvoie les brèves aux responsables concernés (le thème peut toucher plusieurs régions). L'objectif est de faire connaître les brèves et de générer un état d'alerte sur des thèmes singuliers.
  - la situation est connue par l'expert et elle est jugée par lui comme non représentative. Dans ce cas, la cadre ne fait rien.
  - La situation est complexe et l'information doit être discutée en réunion en présence d'autres cadres, experts et responsables concernés. L'objectif est de discuter des éventuelles actions à donner à l'information.

## 5 Premiers résultats et discussion sur les apports managériaux

D'ores et déjà les managers abonnés au dispositif AproximaV2 sont en mesure de prendre leurs décisions avec une meilleure réactivité. Il conviendra toutefois de mener une étude plus approfondie courant du troisième trimestre 2013 pour mesure de manière fine le gain de réactivité effectivement réalisé.

Après une année d'utilisation, nous avons pu observer que :

1. Le recours à des media comme la messagerie et Twitter permet une utilisation immédiate d'AproximaV2. Les cadres retrouvent dans ces deux outils des moyens de communication qui leur sont familiers et pour lesquels il n'existe plus d'effort lié à l'apprentissage technologique de leur usage.
2. La présentation des informations sous la forme du « résumé » a été très bien acceptée. Une fréquence journalière d'émission des « résumés » par des courriels semble être pertinente, notamment pour les informations qui concernent une région.
3. L'information transversale est appréciée des cadres. La réaction de la directrice sur le sujet en atteste puisqu'elle a décrit AproximaV2 comme : « *le bijou de la couronne que tout le monde veut* ».



4. L'information à destination des responsables territoriaux est perçue comme étant pertinente. Selon ces responsables ApproximaV2 permet d'une part de compléter les données quantitatives déjà en leur possession et d'autre part de mieux préparer leurs visites aux communes et départements. Selon eux, Approxima V2 donne un aperçu intéressant et pertinent des principaux sujets actuels de discussion, des risques et des opportunités.

L'expérimentation d'ApproximaV2 soulève également plusieurs difficultés :

1. Quelques « brèves » ont été difficilement interprétables. Le problème dans ce cas semble venir d'un manque de contexte pour comprendre le sens de l'information.
2. Les informations relatives aux régions et qui sont relayées au niveau des medias nationaux produisent « une surcharge d'information » dans ApproximaV2. En effet, lorsqu'une information est jugée importante et sensible, elle est relayée à la fois par les media locaux et nationaux. La même information se retrouve ainsi plusieurs fois dans ApproximaV2.
3. Il arrive que les informations fournies par ApproximaV2 soient déjà connues de leur destinataire via leurs réseaux informels et humains. En effet, les cadres de la DGAF peuvent mobiliser sur le terrain le réseau de leurs collaborateurs comme les leaders communautaires avec qui ils sont en contact permanent pour conseils. Les moyens de communications utilisés sont le « chat » ou le téléphone mobile ce qui permet une transmission rapide et directe des informations jugées importantes. Dans ces conditions, les informations reçues via ApproximaV2 sont d'avantage utilisées comme un moyen de confirmation que d'anticipation.

Ce dernier résultat appelle plusieurs réflexions sur les informations issues de sources numériques comme Internet. Il semblerait que les sources de terrain soient mobilisées pour leur rapidité mais aussi pour leur fiabilité difficile à évaluer. En effet, les informations de veille ont cette caractéristique d'informer sur des événements à venir ou en devenir et sont par nature peu fiables. Mobiliser des sources connues et informelles comme des collaborateurs de terrain permet de pallier pour partie cette faiblesse.

Il semblerait également que les medias synchrones soient source d'une plus grande réactivité pour les managers. Ce résultat est intéressant et questionne les pratiques des managers qui utilisent uniquement les sources Internet pour s'informer. Trop souvent, les managers ne mobilisent dans Internet que les informations issues de sites de presse ou de sites d'entreprises. Ces informations sont généralement très encadrées et contrôlées par leur(s) émetteur(s) et leur caractère novateur et surprenant est souvent rare. Trop peu utilisent les informations issues des blogs et des forums. Or, ces dernières, plus informelles et moins contrôlées peuvent être une solution à une plus grande réactivité des informations via Internet.

L'ensemble des premiers résultats obtenus nous permet de considérer ApproximaV2 comme un outil (de collecte, de tri, de présentation et de diffusion des informations numériques issues d'Internet) complémentaire des informations issues des sources informelles et de terrain. Une solution mixte semble donc intéressante qui associerait sources formelles et informelle, sources numériques et de terrain. Une piste de recherche intéressante serait dès lors de comprendre et d'évaluer comment ApproximaV2 pourrait mieux prendre en compte ou prendre en compte différemment les informations issues des blogs et des forums.

## 6 Bibliographie

[1] AGUILAR FJ. *Scanning the business environment*. New York: Macmillan; 1967.

[2] ALBRIGHT KS. *Environmental scanning: radar for success*. The Information Management Journal. May/June , juin 2004, p38-44.

- [3] ANSOFF HI. *Managing Strategic Surprise by Response to Weak Signals*. California Management Review. 75 , déc 1975, p21-33.
- [4] BERNAT J-P, BRUFFAERTS-THOMAS J, LIBMANN A-M, DESCHARMES S, LIBMANN F. *Les contours de la veille*. IDocumentaliste-Sciences de l'information. 45(4) , 2008, p32-44.
- [5] BETTIS-OUTLAND H. *Decision-making's impact on organizational learning and information overload*. Journal of Business Research. 65(6) , 2012, p814-820.
- [6] BOCK G-W, MAHMOOD M, SHARMA S, KANG YJ. *The Impact of Information Overload and Contribution Overload on Continued Usage of Electronic Knowledge Repositories*. Journal of Organizational Computing and Electronic Commerce. 20(3) , 2010, p257-278.
- [7] BUTLER T, FELLER J, POPE A, EMERSON B, MURPHY C. *Designing a core IT artefact for Knowledge Management Systems using participatory action research in a government and a non-government organisation*. The Journal of Strategic Information Systems. 17 , 2008, p249-267.
- [8] CARON-FASAN ML, LESCA H, CASAGRANDE A, BUITRAGO HURTADO AF. *Comment collecter des données numériques et textuelles, utiles à la phase d'exploitation d'un dispositif de veille anticipative : problématique et proposition d'un outil*. Veille Stratégique Scientifique & Technologique. Toulouse; 2010.
- [9] CHANAL V, LESCA H, MARTINET AC. *Vers une ingénierie de la recherche en sciences de gestion*. Revue Française de Gestion. 116 , 1997, p41-51.
- [10] CHO J, RAMGOLAM DI, SCHAEFER KM, SANDLIN AN. *The rate and delay in overload: An investigation of communication overload and channel synchronicity on identification and job satisfaction*. Journal of Applied Communication Research. 39(1) , 2011, p38-54.
- [11] DAY GS, SCHOEMAKER PJH. *Peripheral Vision: Detecting the Weak Signals that Will Make or Break Your Company [Internet]*. Press HB, éditeur. Strategy / Harvard Business School Press; 2006. Disponible sur: [http://books.google.fr/books/about/Peripheral\\_Vision.html?id=CU26U467DmAC&redir\\_esc=y](http://books.google.fr/books/about/Peripheral_Vision.html?id=CU26U467DmAC&redir_esc=y)
- [12] DECKER R, WAGNER R, SCHOLZ SW. *An internet-based approach to environmental scanning in marketing planning*. Marketing Intelligence & Planning. 23(2) , 2005, p189-199.
- [13] DESCHAMPS C, MOINET N. *L'émergence d'internet dans les outils d'Intelligence économique*. Le Temps des médias. 16(1) , 2011, p147-160.
- [14] DUAN Y, ONG VK, XU M, MATHEWS B. *Supporting decision making process with « ideal » software agents- What do business executives want?* Expert Systems with Applications. 39 , 2012, p5534-5547.
- [15] EPPLERA MJ, MENGISA J. *The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines*. The Information Society: An International Journal. 20(5) , 2004, p325-344.
- [16] FEATHER J. *The Information Society: A Study of Continuity and Change*. paperback, éditeur. Facet Publishing; 2004.
- [17] GRAEFE A, LUCKNER S, WEINHARDT C. *Prediction markets for foresight*. Futures. 42 , 2010, p394-404.
- [18] HAASE H, FRANCO M. *Information sources for environmental scanning: do industry and firm size matter?* Management Decision. 49(10) , 2011, p1642-1657.
- [19] HENRIET I, CAMPOY E, KALIKA M. *Surcharge informationnelle, urgence et TIC. l'effet temporel des technologies de l'information*. Management & Avenir. 3(13) , 2007, p149-168.
- [20] HILTUNEN E. *Good Sources of Weak Signals: A Global Study of Where Futurists Look For Weak Signals*. Journal of Futures Studies. 12(4) , 2008, p21-44.
- [21] KINGSLEY L. *8 tips for raising data from the dead. (cover story)*. Information Management Journal. 46(4) , 2012, p30-33.
- [22] KIRON D, PRENTICE PK, FERGUSON RB. *Death by Information Overload*. Harvard Business Review. 87(9) , 2009, p83-89.
- [23] LANKTON NK, SPEIER C, WILSON EV. *Internet-based knowledge acquisition: Task complexity and performance*. Decision Support Systems. 53(1) , 2012, p55-65.
- [24] LAU RYK, LIAO SSY, WONG KF, CHIU DKW. *Web 2.0 environmental scanning and adaptive decision support for business mergers and acquisitions*. MIS Quarterly. 36(4) , déc 2012, p1239-1268.
- [25] LESCA H. *Veille Stratégique. La méthode L.E. SCANing*. EMS É, éditeur. Management & Société; 2003.

- [26] LESCA H, LESCA E, LESCA N, CARON-FASAN M-L. *Gestion de l'information. 2e édition*. EMS; 2010.
- [27] LESCA H, BLANCO S. *Contribution à la capacité d'anticipation des entreprises par la sensibilisation aux signaux faibles. 6ème Congrès international francophone sur la PME*. HEC, Montreal. , oct 2002,.
- [28] LESCA H, LESCA N. *Les signaux faibles et la veille anticipative pour les décideurs*. Lavoisier, éditeur. Hermes Science; 2011.
- [29] LESCA N, CARON-FASAN M-L. *Veille anticipative. Une approche de l'intelligence économique*. Lavoisier, éditeur. Paris: Hermes Science; 2006.
- [30] NELSON MR. *We have the information you want, but getting it will cost you!: held hostage by information overload*. Crossroads. 1(1) , sept 1994, p11-15.
- [31] SCHICK AG, GORDON LA, HAKA S. *Information overload: A temporal approach*. Accounting, Organizations and Society. 15(3) , 1990, p199-220.
- [32] SCHULTZ U, VANDENBOSCH B. *Information Overload in a Groupware Environment: Now You See It, Now You Don't*. Journal of Organizational Computing and Electronic Commerce. 8(2) , 1998,.
- [33] SIDHOM S, LAMBERT P. *Information Design for « Weak Signal » detection and processing in Economic Intelligence: A case study on Health resources*. Journal of Intelligence Studies in Business. 1 , 2011, p40-48.
- [34] WELZ K. *Information Overload and its Effects on Workplace Productivity*. In: McNerney C, éditeur. Influence of Cognitive and Learning Styles on Knowledge Sharing Behavior: A Theoretical Framework [Internet]. The Knowledge Institute; 2010. p174. Disponible sur: [http://www.academia.edu/499533/Influence\\_of\\_Cognitive\\_and\\_Learning\\_Styles\\_on\\_Knowledge\\_Sharing\\_Behavior\\_A\\_Theoretical\\_Framework](http://www.academia.edu/499533/Influence_of_Cognitive_and_Learning_Styles_on_Knowledge_Sharing_Behavior_A_Theoretical_Framework)
- [35] XU M, ONG V, DUAN Y, MATHEWS B. *Intelligent agent systems for executive information scanning, filtering and interpretation: Perceptions and challenges*. Information Processing and Management. 47 , 2011, p186-201.
- [36] YOON J. *Detecting weak signals for long-term business opportunities using text mining of Web news*. Expert Systems with Applications. 39 , 2012, p12543-12550.
- [37] ZHANG X, MAJID S, FOO S. *Environmental scanning: An application of information literacy skills at the workplace*. Journal of Information Science. 36(6) , 2010, p719-732.