

Solutions logicielles | Un guide pour y voir clair

Comment choisir le logiciel de veille et/ou d'IE adapté à ses besoins ? Et si on commençait par le guide de recensement des outils de collecte, de traitement et de visualisation de l'information, mis à jour par le GFII

Pour contribuer à une meilleure visibilité des acteurs évoluant sur le marché français des solutions de veille et d'IE, le groupe de travail « *Intelligence économique et économie de la connaissance* » du Groupement Français de l'Industrie de l'Information (GFII) a été mandaté par le HRIE pour mettre à jour le « *Guide de recensement des outils de collecte, de traitement et de visualisation de l'information* » réalisé en 2005 et publié par le Cigref en Janvier 2006.

Le guide des fournisseurs de solutions de veille et d'intelligence économique présents en France répertorie les fournisseurs de solutions logicielles et fournisseurs de contenus proposant des fonctionnalités de veille et/ou d'analyse, ayant une activité commerciale en France. L'accès à ce guide a été facilité via la mise en place d'une base de données accessible librement sur le site internet du GFII (<http://guideie.gfii.asso.fr>) et consultable en français et en anglais.

Cette initiative a été mise en place en partenariat avec les fournisseurs référencés puisqu'ils ont été invités à renseigner eux-mêmes les fiches de présentation de leur société et de leurs produits. Une phase de validation a

tout de même été opérée par les membres du groupe de travail du GFII avant chaque mise en ligne afin d'éviter les discours trop partisans.

L'accès aux fiches se fait par le nom d'une société, ou par une recherche en texte intégral. Un accès par fonctionnalités est également proposé : analyse de l'information non structurée (text-mining), analyse de l'information structurée (datamining), collecte - alerte, filtrage et diffusion sur profil, recherche et indexation, traduction, traitement de l'image, traitement de la parole, travail collaboratif, visualisation et cartographie.

Les fiches contiennent à la fois des données signalétiques sur l'entreprise, sur ses spécificités dans le domaine de la veille, sur ses partenaires et ses clients. Chaque produit fait également l'objet d'une description détaillée mentionnant ses caractéristiques technologiques et sa couverture fonctionnelle. La base de données intègre pour le moment près de cinquante fournisseurs et ce chiffre devrait s'étoffer au fil du temps avec le recensement d'autres d'acteurs dont de jeunes sociétés innovantes récemment créées.

Il est également possible que le périmètre de la base soit dans l'avenir étendu, notamment avec l'intégration des centres de recherche dont certains étaient présents dans le guide initial réalisé en 2005.

Présentation des données entreprise

Selon une étude réalisée par SerdaLAB et intitulée « *Les évolutions du marché de l'information électronique professionnelle en France* », le marché français était en 2007 de 42 millions d'euros avec une progression de 48 %. Cette progression a été la plus importante du marché de l'industrie de l'information électronique professionnelle. De plus, selon cette même étude, le segment est caractérisé par la présence de nombreuses PME dont le chiffre d'affaires moyen est de 2 millions d'euros.

Si l'on reprend les données extraites de la base de recensement, l'écho est identique à celui présenté dans l'étude. Ainsi, on observe que 21 % des entreprises possèdent un chiffre d'affaires allant de 0 à 500 000 euros et que 38 % des entreprises réalisent

un chiffre d'affaires entre 1 à 5 millions d'euros. La faible proportion d'entreprises ayant un chiffre d'affaires supérieur à 50 millions d'euros (seulement 9 %) montre bien la jeunesse de ce marché.

Si l'on regarde ensuite les effectifs des dites sociétés, on est bien loin des grands acteurs du secteur de la « business intelligence » tel SAP avec ses 51 000 salariés et ses 11,5 milliards d'euros de chiffre d'affaires. Le secteur des fournisseurs de solutions de veille et d'IE est quant à lui essentiellement constitué de PME innovantes avec 21 % des entreprises ayant entre 1 à 10 salariés et 45 % ayant entre 11 et 50 salariés. La jeunesse de ce marché essentiellement constitué d'entreprises ayant moins de 10 ans d'existence explique en grande partie ces chiffres.

Présentation des données fonctionnelles

Les fonctionnalités sont beaucoup plus diversifiées parmi les entreprises et en moyenne, chaque entreprise a déclaré avoir trois des fonctionnalités que le GFII a rattaché à une des fonctionnalités inhérentes aux solutions de veille et d'IE. Parmi les fonctionnalités qui ont été identifiées comme majoritairement représentées on retrouve la recherche et l'indexation,

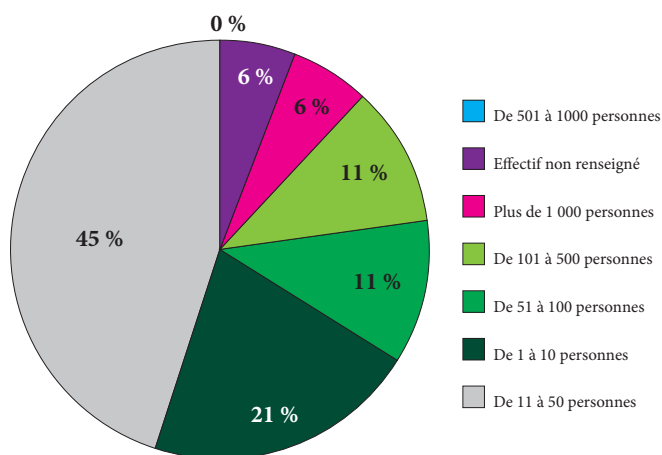
la collecte et l'alerte, l'analyse de l'information non structurée, puis à parts égales, le travail collaboratif et les outils d'aide à la visualisation et la cartographie.

On retrouve, en toute logique, au cœur des fonctionnalités, la recherche et l'indexation. Toute information, qu'elle soit à l'origine structurée ou non structurée, nécessite une réindexation afin de la rendre accessible de façon aisée aux utilisateurs des solutions de veille et d'intelligence économique. Par ailleurs, les solutions d'indexation actuelles ne se contentent plus de la création d'un « simple » index inversé, mais contribuent à la création de valeur autour de cette information en y ajoutant un premier niveau d'analyse statistique et/ou sémantique. Ainsi l'on voit clairement émerger, en plus des listings « traditionnels » de visualisation des résultats, des fonctionnalités de clustering, d'extraction d'entités nommées, d'identification de relation entre ces entités voire même une surcouche de structuration de l'information générée par les utilisateurs eux-mêmes (Tags, commentaires, qualification d'importance, de fiabilité, pertinence de la source ...)

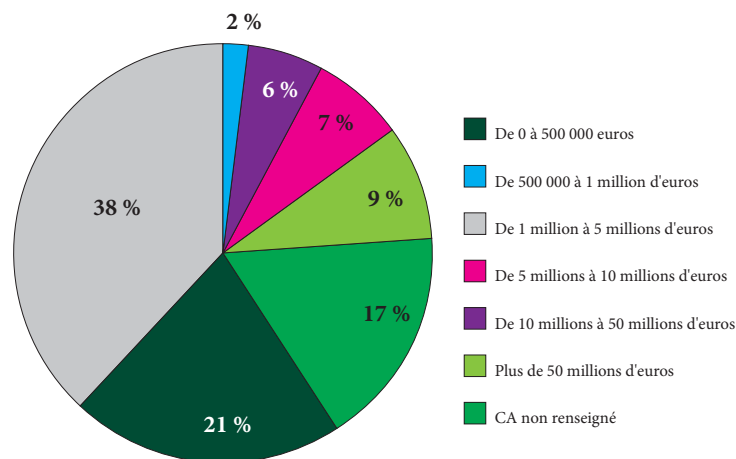
La collecte et l'alerte, cœur important des logiciels de veille et d'intelligence

économique, reste elle aussi au cœur des solutions intégrées dans la base de données du GFII. Avant même son intégration au sein d'une base de données, avant même sa structuration, les logiciels ont besoin de matières premières et pour cela nombre d'elles intègrent des fonctionnalités de collecte, qu'il s'agisse de collecte d'informations ouvertes sur le web de masse ou au contraire ciblées, de dispositifs de remontées d'information terrain ou de partage de documents de travail, ou bien encore de connexion à des bases de données professionnelles. La valorisation des données internes via la mise en place d'un processus de collecte sur ces informations (serveurs de fichiers, bases de données, etc.) est également un élément clé pour tout projet de veille.

La croissance de la masse d'informations sur Internet a dernièrement clairement fait ressentir le besoin de structurer cette information. Les métabalises traditionnelles détournées de leur usage initial à des fins commerciales ont amené les acteurs du search et de la veille à imaginer d'autres façons de structurer de l'information monobloc que peuvent livrer les pages web. Les reconnaissances de liens implicites et explicites, l'exploitation des mises en formes spécifiques et des schémas récurrents à

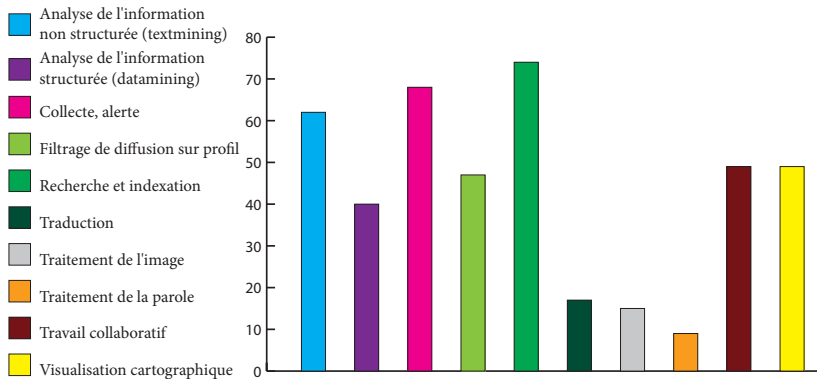


Chiffres d'affaires (en %)



Effectifs des sociétés (en %)

Les fonctionnalités couvertes (en %)



l'intérieur des pages ou bien encore l'analyse du modèle DOM des pages web, sont tout autant d'exploitations intelligentes et d'innovation reposant sur les usages en vigueur et les modes de conception et de programmation des pages web.

A parts égales on trouve les fonctionnalités permettant le travail collaboratif des équipes et les outils d'aide à la visualisation et les outils de cartographie.

Le travail collaboratif est une part essentielle du processus de veille et d'intelligence économique puisqu'il permet de dégager les aspects fonctionnels de la veille (métier de veilleur) des aspects d'expertises nécessaires à la qualification et à l'enrichissement humain de l'information.

Dans la même optique, on constate une présence importante de l'offre en outils d'aide à la visualisation et en cartographie. Face à la masse d'information ouverte aujourd'hui accessible sur Internet et dans d'autres sources numériques, les travailleurs du savoir ne se retrouvent plus à gérer la pénurie informationnelle mais plus face à la problématique d'extraire du sens par rapport à l'information collectée. Les outils de cartographie et de représentation visent à offrir une meilleure lisibilité de l'information

présente dans le système d'information et d'aide à la décision, et permettent de faire émerger du sens du magma informationnel, en se reposant sur la structuration de l'information initiale ou des couches de structuration intermédiaires ajoutées tout au long du processus de veille. Ces outils peuvent ainsi détecter automatiquement sur une masse d'information les concepts clés de type noms de personnes, lieux géographiques, noms d'organisations ou des concepts généraux.

Les solutions d'aide à la représentation et de cartographie de l'information peuvent être ainsi proposées de façon autonome à toute solution de veille via l'utilisation des formats d'entrées/sorties généralistes (souvent standards, XML en tête). Certaines solutions logicielles de veille et d'intelligence économique peuvent par ailleurs embarquer nativement ce module d'aide à la lecture et à la compréhension afin d'offrir une offre globale, couvrant l'intégralité du processus inhérent au cycle de la veille.

Plus marginalement que les fonctionnalités précédemment évoquées, on trouve aussi sur le marché français des solutions de traduction, de traitement de l'image et de traitement de la parole.

On peut expliquer que ce segment fonctionnel soit aujourd'hui moins

présent sur le panel des offres, car la démocratisation de l'accès à l'image et au son via le réseau Internet reste tardive. En effet ce n'est qu'avec la généralisation des accès haut débits que la masse d'information multimédia disponible a nécessité d'imaginer le développement de solutions logicielles afin d'en offrir un traitement systématique et rapide. Aujourd'hui, les fonctionnalités de traduction, de speech to text, et de reconnaissance automatique de formes et d'images sont au cœur des préoccupations des entreprises et les programmes de recherche et de développement français se multiplient en France afin de faire évoluer ces technologies en en augmentant la fiabilité tout en en diminuant les coûts.

Par la mise à disposition de ce guide de recensement, le GFII souhaite avant tout illustrer le dynamisme et la richesse du tissu industriel français dans le domaine du traitement avancé de l'information. Le nombre important d'acteurs référencés en est une parfaite illustration.

Mais il a également vocation à être un outil pratique pour tous les professionnels désireux d'opérer un choix de solution de veille et d'intelligence économique en leur fournissant des informations à jour sur le positionnement des sociétés et les fonctionnalités proposées.

Si vous n'avez pas encore été contacté par le GFII pour renseigner votre fiche, n'hésitez pas à nous adresser un message à l'adresse suivante : [gfii@gfii.asso.fr](mailto:gfi@gfii.asso.fr)

■ **Gabrielle Denis,**
GFII.

Christophe Marnat,
Animateur du groupe IE du GFII, AMI
Software.

Frédéric Martinet,
membre du groupe IE du GFII, Spotter.