

OUTILS DE LA VEILLE DOCUMENTAIRE DANS « LA VEILLE SANITAIRE DE DEFENSE »

TANTI M (*), HUPIN C (*), BOUTIN J-P (*), HASSANALY P (**)

imtssa.veilledoc@wanadoo.fr, imtssa.bedouin@wanadoo.fr,

imtssa.desp@wanadoo.fr, parina.hassanaly@lvi.u-3mrs.fr

(*) Unité de veille sanitaire- Institut de médecine tropicale du service de santé des armées, BP 46-13998 Marseille Armées (France),

(**) Laboratoire de valorisation de l'information et de la communication- Groupe d'étude et de recherche sur l'information et la connaissance- Université Aix-Marseille III- Paul Cézanne - Faculté des sciences et techniques de St Jérôme - Avenue Escadrille Normandie Niémen - 13397 Marseille Cedex 20 (France).

Mots clefs :

Armées françaises, veille sanitaire, veille documentaire, document scientifique, base de données bibliographiques, collecte d'informations

Keywords:

French forces, medical intelligence, documentary-awareness, scientific document, bibliographic-database, information gathering

Palabras clave :

Fuerzas armadas francesas, vigilancia-sanitaria, vigilancia-documental, documento científico, base-datos-bibliográfica, reunir de información

Résumé

Une des missions du Service de santé des armées (SSA) est de préserver la santé de ses personnels projetés hors territoire. Dans cette optique, il a confié à l'Institut de médecine tropicale du service de santé des armées (IMTSSA) une mission de veille sanitaire de défense. Pour répondre à cette mission, trois bases documentaires ont été mises à sa disposition. La mise à jour de ces bases, la détection rapide de nouveautés et de « signaux faibles en émergence » nécessite un processus de veille documentaire en plusieurs phases : collecte de documents, analyse et diffusion. Le document scientifique, validé, polymorphe est l'objet de ce processus. Les phases de collecte et d'analyse, cruciales, sont décrites dans cet article. La phase de collecte nécessite la mobilisation d'outils spécifiques selon trois modalités: le « pull », le « push » et le fil RSS (Really Simple Syndication). Le « pull » est la méthode de consultation classique et directe des documents mais exige une rigueur dans les recherches. Dans le « push », le document est poussé, de façon automatique, vers l'utilisateur, par courriel, au risque de le submerger de documents sans intérêts. Le fil RSS constitue la solution d'avenir car il évite à l'utilisateur le pollu-postage comme c'est souvent le cas avec le « push ». Les outils de l'analyse soutiennent le veilleur mais ils ne remplaceront jamais l'analyse qu'il doit faire à froid. Dans ce processus chronophage, un matériel informatique performant se révèle indispensable.

Introduction

Pour maintenir l'état de santé de ses personnels projetés, le Service de santé des armées (SSA) a confié à l'Unité de veille sanitaire (UVS) de l'Institut de médecine tropicale du service de santé des armées (IMTSSA) une mission de veille sanitaire de défense définie en 2004 comme « collecte, analyse et diffusion à ceux qui en ont besoin, d'informations sanitaires polymorphes, provenant de populations extérieures à celle de l'organisme de veille, afin d'identifier et de prévenir les risques sanitaires potentiels pour la population sous la responsabilité du veilleur dans le domaine de la défense » (1, 10). Pour répondre à cette mission, trois bases documentaires servant de socle de documentation scientifique ont été mises à sa disposition. La première base diffuse des documents scientifiques bruts sur les agents du risque biologique naturel et provoqué, mais également sur les toxiques chimiques industriels et militaires. La seconde met à disposition des documents de synthèse sur les recherches et développements pouvant être en relation avec une volonté de militarisation d'agents biologiques et chimiques. La dernière fournit à la fois des renseignements médicaux, épidémiologiques mais aussi historiques, géographiques, économiques, écologiques sur les différents théâtres extérieurs actuels ou possibles des personnels. Dans les armées, elle est dénommée BEDOUIN pour Base épidémiologique de données sur l'outre-mer et l'intertropical. La mise à jour de ces bases, la détection rapide de nouveautés et de « signaux faibles en émergence » nécessite un processus de veille documentaire en plusieurs phases : collecte de documents scientifiques, analyse et diffusion.

1 Définition de la veille documentaire

Une définition de la veille documentaire dans la veille sanitaire de défense est en cours d'élaboration. La définition qui pourrait être proposée est la suivante : « processus de collecte, d'analyse et de diffusion de documents scientifiques, validés, polymorphes, afin de mettre à jour des bases documentaires qui permettront au commandement militaire d'identifier des signaux faibles en émergence et d'aider à la décision dans le cadre de la conduite et de la planification d'opérations de projection des forces armées »

La répartition mondiale des microbes, des implantations industrielles, les recherches sur les agents pathogènes et chimiques évoluent très vite.

Face à cette évolution incessante, pour répondre à la mission de veille confiée, le veilleur documentaire doit être informé très rapidement des nouveautés et des changements dans ses domaines d'intérêts.

Dans ce contexte, le document scientifique est l'objet même du processus de veille documentaire et les étapes de collecte et d'analyse sont cruciales.

Face aux flux de documents scientifiques de qualité et de quantité variable, l'étape de collecte n'est possible que par la surveillance de sources documentaires bien définies et par l'application des outils de type « push », « pull » et fil RSS (Really Simple Syndication) (7).

2 Document scientifique, objet de la veille documentaire

Par définition, le document scientifique est une publication scientifique électronique ou imprimée, objective, rédigée par des experts du domaine, validée par les pairs, et dont le contenu d'information est inédit ou original (8). Le document scientifique validé et mis à jour montre donc son caractère indispensable face à l'évolution permanente des thématiques d'intérêts de la veille sanitaire de défense.

Pour exemple, une enquête épidémiologique réalisée dans les standards de qualité de la discipline, et dont les résultats sont publiés dans une revue scientifique avec comité de lecture a valeur de référence pour confirmer la certitude d'une information et par conséquent la certitude de l'existence d'un risque, à un instant donné, dans un lieu donné.

Cependant, entre le recueil de l'information, sa validation scientifique et sa publication, un délai de plusieurs mois à plusieurs années peut s'écouler. A contrario, une rumeur peut paraître immédiatement dans la rubrique fait-divers d'un quotidien grand public mais n'a aucune valeur scientifique.

3 Typologie des documents collectés

Dans l'impossibilité de collecter des documents décrivant des données individuelles et standardisées sur des cas (règle de bonne pratique de la surveillance), le veilleur documentaire collecte des documents scientifiques et médicaux de nature très diverse (origine, langue...) et de valeur différente.

Les principales formes de documents d'intérêt pour la mission de veille documentaire sont les articles de revues de recherche et de synthèse, les revues scientifiques et médicales, les brevets d'invention, les ouvrages, rapports, mémoires et thèses au format électronique ou imprimé.

Les documents relatifs aux maladies transmissibles sont au premier plan de cette veille. Ils sont, en effet, l'objet de la demande principale en documents référencés, de par la nature duale du risque biologique et des conséquences induites par la contagiosité des agents.

Mais, les documents relatifs aux risques environnementaux et en particulier ceux relatifs aux risques d'origines industriels et chimiques sont aussi considérés avec la plus grande attention.

Le veilleur documentaire se livre donc à une catégorisation en fonction des priorités définies.

Les documents traitant de recherches et de développements concernant les agents biologiques et chimiques militarisables (ex: charbon humain, choléra, peste, ricine...) sont collectés en premier lieu.

Les documents de synthèse sur les mêmes agents et les documents opérationnels sur la conduite à tenir en cas de crise sanitaire d'origine biologique chimique ou radiologique (ex : plans BIOTOX, PIRATOX et PIRATOME) intéressent également le veilleur documentaire dans sa mission de collecte.

Les documents traitant de crise sanitaire dans un pays de stationnement permanent ou de projection actuelle, probable ou possible des forces (ex : choléra à Djibouti en 1997) et les documents traitant d'émergence d'événements sanitaires inhabituels dans des territoires jusqu'alors indemnes de ces événements ou de leur réémergence inexplicquée (ex : émergence de fièvre hémorragique à Marburg en Angola en 2005) font également partie des centres d'intérêts.

4 Description des outils de la collecte

4.1 Outils de type « pull »

Le « pull » représente la méthode classique d'utilisation de la documentation autant en bibliothèque que sur l'Internet, l'utilisateur se rend directement et régulièrement sur des sources de documents scientifiques, validés, opérationnels, pour en tirer les informations les plus récentes dans les domaines d'intérêt. La fréquence de consultation et les sites consultés sont déterminés par le veilleur documentaire.

L'Unité de veille sanitaire a choisie de consulter les notices bibliographiques et les documents en texte intégral de la base bibliographique en sciences biologiques et médicales, Medline (MEDical Literature Analysis and Retrieval System on LINE) (2, 4), produite par la National Library of Medicine (NLM™) aux Etats-Unis.

Medline contient des liens vers des articles en texte intégral sur les sites des éditeurs participants, ce qui permet leur téléchargement éventuel. Elle contient des liens vers des sites tiers tels les centres de séquençage. Elle fournit également un accès et des liens aux bases de données intégrées de biologie moléculaire produites par le National Center for Biotechnology Information (NCBI). Le choix de consulter cette base a été motivé par le fait qu'elle indexe la plupart des revues nationales et internationales en médecine, en sciences vétérinaires, en sciences pré-cliniques et en sciences biologiques, domaines qui intéressent particulièrement la mission de veille documentaire. De plus, sa mise à jour est quotidienne, d'où une relative fraîcheur de l'information. Cette base contient des références bibliographiques et des résumés des auteurs de plus de 4 600 revues publiées aux États-Unis et dans 70 autres pays. Elle est constituée d'environ 12 millions de documents remontant jusqu'en 1966. Les références sont indexées avec les mots clés MeSH (Medical Subject Headings) de son thésaurus, ce qui est très avantageux pour augmenter la pertinence des résultats de la recherche. Par contre, les monographies et les résumés de congrès ne sont pas indexés. Bien que de couverture mondiale, la plupart des documents sont écrits ou résumés en anglais. L'Unité de veille sanitaire a choisi de consulter Medline à partir de PubMed car l'accès à partir de ce site est gratuit. Le lien Related Articles de PubMed s'affiche en face de chaque référence et permet, par un algorithme, de retrouver les références similaires. Cette commande permet ainsi de compléter ou d'élargir une recherche en faisant l'économie d'une recherche de mots clés. Le lien vers PubMed Central (PMC) permet de consulter les revues en libre accès dans les domaines des sciences de la vie. PubMed offre en plus la possibilité d'effectuer des stratégies de recherche plus complexes en utilisant des opérateurs booléens et avec limitation de date, de langue ou d'auteurs et par combinaison interchamps ou intrachamps. Medline ne couvre pas la littérature médicale du monde entier et selon les sujets, il donne peu de résultats.

L'Unité de veille sanitaire oriente alors ses recherches sur d'autres bases comme ToxNet, service d'information spécialisé de la National Library of Medicine (NLM), qui interroge séparément ou de façon croisée 8 bases de données complémentaires en toxicologie humaine, animale, ou environnementale concernant médicaments, produits chimiques, phytosanitaires ou industriels. Les sources de données sont : la NLM, l'Environmental Protection Agency (EPA), le National Cancer Institute (NCI).

Parmi ces 8 bases, Hazardous Substances Data Bank (HSDB) est la plus intéressante pour la veille sanitaire de défense. Facilement exploitable, elle se révèle capable de fournir des données pratiques sur la plupart des substances chimiques connues. L'interrogation du moteur de recherche est simple (nom de la substance recherchée), et la réponse est complète (et facilement paramétrable via le bouton Limits). En effet, sont détaillées et surtout référencées : les toxicités humaines, animales, les conduites à tenir urgentes, les données pharmacodynamiques et cinétiques, propriétés physiques, chimiques,

données environnementales, méthodes de dosage. Pour détaillée qu'elle soit, la présentation de la réponse reste claire et efficace.

Toxline est une autre base internationale accessible à partir de Toxnet. Sa consultation est également très utile. Elle permet l'obtention rapide de références bibliographiques en toxicologie, pharmacologie, biochimie et effets physiologiques des médicaments et autres produits chimiques. Ses références proviennent d'articles de revues, monographies, rapports techniques, thèses, lettres et résumés de congrès. Toxline contient des références de la base Medline, Dart (Development and Reproductive Toxicology) et Biosis. Sa mise à jour est mensuelle. En moyenne, 9300 références sont rajoutées chaque mois. Elle contient 3 millions de références depuis 1965.

IRIS (toxicité environnementale et risque sanitaire humain) est une autre base également accessible à partir de Toxnet. Elle recense à peu près 500 substances chimiques ayant un risque pour la santé. Elle recense également des informations provenant de l'EPA. Les doses de référence par voie buccale et les concentrations de référence par inhalation pour les substances non carcinogènes sont incluses. Les coefficients de variations et les unités de risques pour les substances carcinogènes sont consultables. Ces différentes données ont été validées par les scientifiques de l'EPA et la mise à jour est mensuelle.

L'Unité de veille sanitaire consulte également les notices bibliographiques issues des bases de données bibliographiques multidisciplinaires Pascal et Francis produites par l'INIST-CNRS (3). Ces deux bases multilingues sont consultables en abonnement sur Internet à partir de l'interface Silverplatter. Cette consultation est motivée par le fait que ces bases indexent, depuis 1973, la plupart des revues nationales et internationales, en sciences de la vie et de la santé pour Pascal et en sciences humaines et sociales, pour Francis. Ces deux domaines intéressent la veille sanitaire de défense et ne sont pas couverts par d'autres bases de données spécialisées. Dans ces bases, une place importante est accordée à la littérature française et européenne qui représentent 45% des documents signalés. Outre les articles de périodiques, ces bases offrent un signalement renforcé de la littérature grise (rapports, actes de congrès, thèses...). Depuis 1996, toutes les affiliations de tous les auteurs sont présentes dans les notices. Elles permettent d'effectuer des études bibliométriques afin d'identifier et de localiser les experts ou de cerner les domaines dans lesquels la recherche apparaît particulièrement productive.

Pour consulter les brevets d'inventions sur les nouveaux outils diagnostics et les nouveaux vaccins du risque biologique et chimique, l'Unité de veille sanitaire a choisi d'interroger la base, esp@cenet, consultable gratuitement sur internet. Cette base référence les brevets publiés par l'Office Mondiale de la Propriété Intellectuelle, le Japon et l'Office Européen des brevets. La recherche est multilingue et par numéro de brevet. Elle peut se faire en mode simple ou avancée, avec limites de date, d'année ou de langue. Mais, le site ne permet pas de télécharger les textes intégraux. Le choix a été pris de les copier et de les coller sur un document Word®.

L'Unité de veille sanitaire consulte aussi les notices bibliographiques du fond documentaire mutualisé des centres de documentation du Service de santé des armées sur l'Intranet documentaire de la défense (INTRADOC). Ce fonds contient les ouvrages référencés, les rapports et formations internes mais également les travaux de recherche clinique. L'interface est consultable en mode simple et avancé et par mots clefs du texte. Le choix de consultation de cette base, alimentée par les documentalistes du SSA, est lié au fait qu'elle indexe la littérature grise interne.

Dans ces différents outils du « pull », le veilleur ne fait aucune inscription sur les sites qu'ils consultent pour sa recherche. Il n'y a pas d'enregistrement d'un profil de recherche. Le veilleur documentaire ne divulgue pas son adresse électronique. Ce qui est intéressant pour des raisons certaines de confidentialité des thématiques de recherche des armées. La relance de la recherche se fait de façon manuelle par l'utilisateur. Il a une démarche de recherche active, à l'instant qu'il choisit, sur les sites documentaires qu'il choisit. Mais, il doit se tenir à une certaine régularité dans ses consultations pour éviter de passer à travers un document contenant des signaux faibles en émergence. Dans le « pull », le veilleur documentaire peut modifier ses filtres de recherche. Il peut faire évoluer,

modifier, changer ses équations de recherche, ses champs d'intérêts, ses mots clefs, ses sélections de revues au fur et à mesure de la navigation et à chaque recherche. Dans chaque site documentaire consulté, le veilleur trie et analyse les documents en même temps qu'il les collecte. Vu que la recherche s'effectue source par source, il ne se retrouve pas submergé par des documents sans valeur ajouté les uns par rapport aux autres provenant de plusieurs sources différentes, comme cela peut être le cas avec le « push ».

4.2 Outils de type « push »

Par définition, le « push » est une méthode propre à l'Internet qui consiste à apporter à l'utilisateur ou à pousser vers l'utilisateur, de manière directe et automatique, dans sa boîte aux lettres électronique, l'information documentaire qu'il a programmée en fonction des critères qu'il a choisis.

Notamment, le choix a été fait par l'Unité de veille sanitaire de recevoir les alertes des sommaires des revues du Nature Publishing Group (Nature Reviews Microbiology, Nature Medicine, Nature...) pour la correspondance des thématiques de ces revues aux thématiques d'intérêt de la veille sanitaire de défense en terme de recherche et développements sur les agents du risque agressif.

Le choix d'être abonné par mail à des lettres d'informations a également été fait par l'Unité de veille sanitaire. Ce choix s'est porté sur les *Newsletters* de l'Institut Pasteur. Ces lettres informent mensuellement le veilleur sur les dernières recherches en infectiologie, bactériologie ou virologie et sur les conférences à venir dans ces domaines.

L'Unité de veille sanitaire est, également, abonnée, par mail, aux Bulletin Epidémiologique Hebdomadaire (BEH) publié par l'Institut national de Veille Sanitaire (InVS) qui fait un point hebdomadaire de la situation épidémique mondiale.

Dans la même optique, le veilleur documentaire reçoit, par courriel, Morbidity and Mortality Weekly Report (MMWR) et Emerging Infectious Diseases (EID), deux revues en texte intégral, publiés par le Center for Diseases Control and Prevention (CDC) d'Atlanta.

Il est alerté des nouvelles publications des revues en libre accès (11), notamment celles de BioMedCentral (BMC) : BMC Infectious Diseases, BMC Microbiology...

Le veilleur reçoit les nouvelles publications correspondant aux thématiques d'intérêts provenant des bases documentaires médicales et scientifiques en texte intégral, ScienceDirect®, de l'éditeur hollandais Elsevier, via son service Alerts. Cette base contient les références bibliographiques et le texte intégral, depuis 1995, des revues d'Elsevier, Academic Press, Pergamon, American Institut of Chemical Engineers, North Holland et Gauthier Villard. L'intérêt de cette base vient du fait qu'elle répertorie des revues dans les domaines de la médecine et des sciences et techniques qui ne sont pas indexées dans d'autres bases, comme Medline, Pascal et Francis. Sa mise à jour est quotidienne. En outre, l'utilisateur peut télécharger le résumé et le texte intégral des publications de plus de 1700 revues. Les articles pré-publiés sont également consultables. La diffusion sélective des documents s'effectue selon une procédure automatisée. Dans un premier temps, l'utilisateur sauvegarde son équation de recherche ou le sommaire d'un périodique d'intérêt. Il donne son adresse électronique et détermine la fréquence de réception de ses alertes. Dès qu'une nouvelle publication en rapport avec cette équation ou chaque fois qu'un nouveau numéro de la revue d'intérêt est indexé sur le site, l'utilisateur est alerté par mail. Le lien Save as Citation Alert permet également d'alerter le veilleur documentaire chaque fois qu'une publication est citée dans les références d'une autre publication.

Le coût de l'abonnement à cette base, consultable sur Internet, est très élevé. Le Service de santé des armées a dû négocier avec l'éditeur pendant près d'un an pour qu'il revoie ses tarifs à la baisse et augmente la couverture des revues. Ceci n'a été possible que grâce à l'adhésion du SSA au

Consortium Couperin. Ce consortium, devenu le deuxième en Europe, regroupe différentes structures de recherches et universitaires. Il oeuvre pour la mutualisation des ressources documentaires électroniques. Il intervient pour l'évaluation, la négociation et l'achat au meilleur prix des produits documentaires électroniques. Il intervient dans la mutualisation des compétences, ressources humaines et financements documentaires. Il participe à la coopération internationale et notamment dans l'espace européen. Il construit un réseau national de compétences et d'échanges en matière de systèmes d'information documentaire, notamment en ce qui concerne l'accès intégré aux ressources électroniques, l'archivage pérenne, la signalisation des ressources alternatives et les publications en ligne de ses membres. Enfin, il favorise la communication des membres sous la forme de listes de diffusion et d'un site web.

L'Unité de veille sanitaire a choisi d'être alerté par mail à chaque nouvelle publication d'un ouvrage d'épidémiologie, de maladies infectieuses ou encore de santé publique édité par les Editions Médicales INTERNationales (EMINTER), Vigot-Maloine et Masson, grâce à leur service d'alertes par mail. Ces éditions publient, régulièrement des ouvrages en langue française, très synthétique et toujours à la pointe, dans les domaines de recherche de la veille sanitaire de défense. Ils constituent une littérature complémentaire aux publications des revues.

Dans le « push », le veilleur documentaire automatise sa stratégie de recherche. Au préalable, il établit un plan de concept permettant de traduire ses sujets d'intérêts en mots-clés, trouver des synonymes, des mots connexes, des expressions ou des termes équivalents en plusieurs langues. Ce plan lui permet d'élaborer un profil de recherche qui contient les équations de recherches, les filtres, une sélection des champs d'intérêts, une sélection de revues. Le veilleur doit ensuite enregistrer et sauvegarder son profil de recherche sur le site documentaire, inscrire ses données personnelles, donner son adresse électronique, ce qui n'est pas sans poser de problèmes pour garder confidentiel les thématiques d'intérêts des armées. Pour cette raison, le veilleur a choisi de multiplier les adresses électroniques de réception des courriers. La relance de la recherche s'effectue ensuite automatiquement dans la boîte aux lettres du veilleur en fonction d'une périodicité qu'il aura établi. Mais, le profil de recherche, lui, n'évolue pas automatiquement. Le veilleur doit donc se rendre sur le site s'il veut modifier ses équations de recherche et ses filtres. Etant donné que la veille s'effectue sur plusieurs sites où l'information est parfois redondante, le veilleur risque d'être très vite encombré et de voir sa boîte aux lettres submergée par des documents redondants sans valeur ajoutée les uns par rapport aux autres. Pour certains outils, l'alerte peut être intéressante mais ne permet pas le lien vers le texte intégral. L'utilisateur devra donc dans un deuxième temps se rendre sur le site de l'éditeur et faire du « pull ».

4.3 Outils de syndication des contenus : les fils RSS

Les fils RSS (Really Simple Syndication) sont des flux de contenus gratuits en provenance de sites Internet. La lecture de ces fils se fait via un logiciel de lecture des flux RSS. Ces fils incluent les titres des articles, des résumés et des liens vers les articles intégraux et sont exploitables dynamiquement par d'autres sites comme les agrégateurs de News (syndication de contenu).

L'Unité de veille sanitaire utilise le service RSS feed de PubMed qui permet d'attacher un fil à ses équations de recherche. Le veilleur est informé à chaque fois qu'une nouvelle publication en rapport avec cet équation est indexé sur leur site.

Le veilleur documentaire met également à profit le fil RSS que l'éditeur Ingenta attache au sommaire de ses revues électroniques pour se tenir informé des nouvelles publications des revues qu'il a sélectionnées, comme Military Medicine.

Le fil RSS semble la solution d'avenir pour faire de la veille documentaire au bon moment sans être contraint de se rendre directement sur les sources documentaires identifiées pour en tirer les informations les plus récentes dans les domaines d'intérêt. Le veilleur documentaire est informé des

nouveautés directement via son fil sur son poste de travail et il détermine lui même la fréquence de consultation de ses fils. Le fil RSS présente un autre avantage par rapport au « push » qui est son anonymat. En effet, à aucun moment l'adresse électronique n'est dévoilée d'où une relative confidentialité (6). Cela évite le débordement des boîtes aux lettres électronique comme c'est souvent le cas avec le « push » (5).

La solution choisie pour la lecture des fils par l'Unité de veille sanitaire est le téléchargement du lecteur de fil, RSS Reader, car il est gratuit et son paramétrage est facile. Mais, il faut savoir que des agrégateurs de contenus au format RSS proposent la consultation de leur site quand le veilleur documentaire souhaite prendre connaissance des fils d'informations sélectionnés (ex : My Yahoo®, My Desktop®...). Des navigateurs intègrent également une fonction lecture de fil RSS (ex : Firefox®).

5 Description des outils de l'analyse

L'Unité de veille sanitaire expérimente actuellement les outils d'analyse bibliométriques et scientométriques qui permettent de tirer de l'information à valeur ajoutée d'un corpus de documents ou qui permettent de quantifier les résultats d'une collecte ou d'appuyer l'analyse du veilleur.

A l'heure actuelle, l'Unité de veille sanitaire applique les outils d'analyse textuelle par approche linguistique, notamment l'outil de contraction de texte Copernic Summarizer® et l'outil de traduction automatique Systran Professional®.

Le choix du logiciel Copernic Summarizer® a été fait car il permet de faire des résumés concis de documents ou de pages Web de 5% du texte jusqu'à 1000 mots. Ce logiciel est fondé sur des algorithmes basés sur des calculs statistiques et des données linguistiques. Il permet en outre d'identifier les concepts clés d'un texte et en extrait les mots clefs et les phrases les plus marquantes. Le veilleur documentaire utilise cet outil notamment pour extraire les concepts clefs de rapports publiés par l'OMS qui font souvent plus de 100 pages et pour en faire un résumé qui permet d'avoir une vision globale des concepts.

Le logiciel Systran Professional® permet de traduire directement des textes au format Word®, Excel®, Pdf ou PowerPoint® et les pages web en temps réel. Ce logiciel contient en outre un dictionnaire médical français/ anglais et anglais/ français intégré avec plus de 2000 mots et permet dix traductions possibles. Cet outil est intéressant pour le veilleur documentaire car la grande majorité des documents collectés le sont en langue anglaise et sont d'origine médicale ou scientifique. L'emploi de ce traducteur permet d'avoir une idée globale sur le sens d'un texte scientifique mais ne permet pas d'avoir une vision très détaillée, ni une traduction parfaite. C'est d'ailleurs le cas de la plupart des logiciels de traduction automatique en vente actuellement.

Ces différents outils appuient l'analyse face à un flux de documents souvent de qualité et de quantité variables. Mais, l'analyse humaine faite à froid est indispensable. Le veilleur doit effectuer une sélection et une validation des documents, fonction des événements traités, et par comparaison aux événements antérieurs et connus. Il doit analyser le caractère innovant de l'information apportée par chaque document, son intérêt, le facteur d'impact de la publication (9), la rigueur du contenu, la crédibilité des auteurs, et la correspondance aux thématiques définies. Il doit également analyser le fond et la forme du document, les objectifs et les hypothèses de recherche et l'approche mise en œuvre. De plus, il doit synthétiser et mettre en forme les documents collectés et les transformer en véritable information utile. Ce traitement des documents bruts prend deux aspects : des fiches de synthèses brèves et des documents de synthèses plus élaborés.

6 Matériel informatique

La veille documentaire est un processus chronophage. Elle nécessite donc des matériels qui doivent permettre de gagner du temps sur toutes les étapes du traitement de l'information documentaire.

Des écrans ergonomiques, c'est à dire grands et doublés, pour permettre l'utilisation simultanée de plusieurs sources ou programmes, sont donc utilisés. Des postes avec des capacités de stockage de données conséquentes de l'ordre du Téraoctet sont également indispensables.

Des postes connectés à Internet à haut débit de téléchargement avec au moins 2 Mégabits par seconde, sinon plus, en différenciant les lignes pour chaque acteur de cette veille et en multipliant les fournisseurs pour pallier à d'éventuelles défaillances sont primordiaux.

Un scanner de très haute qualité, pouvant numériser jusqu'à 15 ppm, des documents de 50 pages, sans surveillance grâce à un chargeur automatique, s'est révélé très utile pour l'Unité de veille sanitaire. Il permet de sauvegarder sous format numérique des documents imprimés comme les rapports des armées et les enquêtes de surveillance épidémiologique.

Le stylo scanner a montré également tout son intérêt. Il permet, en effet, d'encoder instantanément sur Word®, les parties intéressantes de textes imprimés.

7 Conclusion

Pour collecter les documents scientifiques permettant la mise à jour des bases documentaires qu'elle met à disposition de son autorité de tutelle, l'Unité de veille sanitaire de l'IMTSSA applique les outils de la veille documentaire de type « pull », « push » et « fil RSS ».

Ces trois types d'outils présentent des avantages et inconvénients mais sont complémentaires dans la mission de veille documentaire confiée. Le fil RSS connaît aujourd'hui un véritable engouement et constitue la solution d'avenir pour l'unité car il évite à l'utilisateur le pollu-postage de sa boîte aux lettres électronique comme c'est souvent le cas avec le « push » et car il permet à l'utilisateur de consulter les nouveautés directement sur son lecteur de fil, au moment qu'il choisit.

Les outils de traduction automatique et de contraction de texte appliqués soutiennent le veilleur dans sa fonction mais ils ne remplacent pas l'analyse qu'il doit faire à froid.

Ces outils de la veille documentaire ne remplacent pas non plus le réseau de personnes ressources et le réseau documentaire du veilleur qui lui apportent la littérature grise souvent indispensable et inaccessible par ces outils.

L'activité de veille documentaire de l'Unité de veille sanitaire de l'IMTSSA doit évoluer avec les technologies de l'information et de la communication. Elle doit suivre la montée en puissance des web-logs et des fils RSS. Elle doit apprendre à tirer parti des outils de « text mining » pour appuyer son analyse des documents, tout ceci au gré des futures opérations extérieures et pour une meilleure prise en charge sanitaire et médicale des personnels.

8 Bibliographie

- [1] BOUTIN JP., RIBIERE O., VAN CUYCK H., MALOSSE D., *Pour une veille sanitaire de défense*, Médecine et armées, 2004, 32, 4, p 366-372
- [2] BOYNTON J., GLANVILLE J., MCDAID D., LEFEBVRE C., *Identifying systematic reviews in MEDLINE: developing an objective approach to search strategy design*, Journal of Information Science 1998, 24, p 137-157
- [3] BRAND-DE HEER DL., *A comparison of the coverage of clinical medicine provided by Pascal BIOMED and MEDLINE*, Health Information and Libraries Journal, 2001, 18, p 110-116
- [4] BUSSIERES JF., LEBEL D., *Utilisation pratique et efficace de PubMed*, Québec pharmacie, 2003, 50, 1, p 20-22
- [5] EVEILLARD P., *Comment prévenir le spam et le fishing*, Revue du Praticien Médecine Générale, 712/713, 05 déc 2005, p 1385
- [6] EVEILLARD P., *Les fils et les posts prennent la relève du push*, Revue du Praticien Médecine Générale, 700/701, 12 Sept 2005, p 913
- [7] EVEILLARD P., *Veille documentaire : premiers pas, premières stratégies*, Revue du Praticien Médecine Générale, 17 nov 2003, 631, p 1515
- [8] FEGHERASSI-PAGEL H., FOHANNO V., MOREL-PAIR C., *L' INIST/CNRS, du document scientifique à la veille, un accès privilégié à l' information biomédicale internationale. Communication médicale et scientifique* Revue médicale de l'assurance maladie, 2001, 32, 2, p 167-174
- [9] HECHT F., HECHT BK., SANDBERG AA., *The Journal "Impact Factor": a misnamed, misleading, misused measure*, Cancer Genetics and Cytogenetics, 1998, 104, p 77-81
- [10] HUPIN C., TANTI M., FERCHICHI S., MIGLIANI R., BOUTIN JP., *Veille sanitaire de défense : objectifs, définition et mise en place au sein du Service de santé des armées*, Journées de veille sanitaire, 2005
- [11] SALAUN JM., *Libre accès aux ressources scientifiques et place des bibliothèques*, Bulletin des Bibliothèques de France, 2004, 49, 6, p 20-30