



Emmanuelle Larroque :
le numérique a besoin
des femmes

p. 08



Emilie Berthod :
documentaliste musicale
en quête d'harmonie

p. 47



Pascal Boniface :
l'intelligence artificielle est
la rivalité des puissances

p. 48

archimag.com

[STRATÉGIES & RESSOURCES DE LA MÉMOIRE & DU SAVOIR]

n° 342

19 euros - mars 2021

serda

une publication du groupe

[dossier]

vrac numérique : comment mettre de l'ordre



alerte au vrac numérique !

Le marketing a bien fait son œuvre, adossé il est vrai à de réels arguments de confort fonctionnel : le cloud computing, ou cloud, est aujourd'hui un succès. Ceci quel que soit son modèle : infrastructure as a service (IaaS), platform as a service (PaaS) et software as a service (SaaS). Tous les observateurs, Gartner, Forrester et autres, constatent la croissance de ce marché. Ainsi, créer et stocker de l'information numérique est simple comme bonjour. Un abonnement auprès d'un prestataire cloud, un espace illimité ou facile à augmenter et ce n'est plus un souci. Avec un serveur d'entreprise la situation n'est guère différente : le coût d'un stockage supplémentaire est dérisoire. Mais la médaille a forcément son revers. Les utilisateurs, dans leurs applications métier et leur drive, s'en sont donné à cœur joie. Résultat : les volumes explosent. Surtout, l'information devient difficile à trouver, des documents d'activité importants ou ayant valeur de preuve perdent en visibilité, les actifs patrimoniaux sont dilués, à un tel point qu'opérer un véritable archivage électronique ou



Michel Remize

DR

un versement, dans le secteur public, relève du casse-tête. Parole d'archiviste : « Avec le passage au zéro papier, comment identifier des archives que je ne vois pas ? » Cette pandémie porte un nom : le vrac numérique.

Que faire ? Grave question qui interroge la gouvernance de l'information. On peut éduquer les utilisateurs avec des règles et des conseils — après tout, ils en seront les premiers bénéficiaires. Tout professionnel de l'information en a la compétence.

On peut mettre le nez dans le système d'information et utiliser à fond son moteur de recherche, son logiciel de gestion de contenu ou d'archivage. Tout records manager ou e-archiviste connaît. On peut attaquer le problème non par les documents, mais par leurs data. Metadata manager, tu deviendras ! ■

Michel Remize
[Rédacteur en chef]

Archimag soutient les étudiants

Archimag offre 6 mois d'abonnement numérique à tous les étudiants. Nous souhaitons ainsi témoigner notre soutien aux jeunes qui subissent de plein fouet les effets de la crise sanitaire du Covid-19 dans leurs études et leur quotidien. ■

nous faisons Archimag

Serda édition-IDP
24, rue de Milan, F-75009 Paris
Tél. : +33 (0)1 55 31 92 30
Fax : +33 (0)1 44 53 45 01
infos@archimag.com
www.archimag.com

contacts e-mail
prenom.nom@archimag.com
rédaction
rédacteur en chef
Michel Remize
directrice de la rédaction
Louise Guerre
directeur de la publication
Pierre Fuzeau
l'équipe de rédacteurs
Fabien Carré, Clémence Jost,
Eric Le Ven, Bruno Texier

nos correspondants
Arbido (Genève),
Ikram Bouzakar (Tanger),
Florian Delabie (Bruxelles),
Philippe Laurent (Bruxelles),
Élisabeth Laviguer (Montréal),
ont collaboré à ce numéro
Emmanuelle Asselin, Solenn Bihan, Jérôme Bondu, Laurence Crohem, Patrick Devos, Marc Maisonneuve, Michel Thomas
conception graphique
Julio Arias-Arnanz, Amcoat
maquette
Artislyck (Aline Paumard)
dessinateur
Yves Barros
publicité
directrice de la publicité
Cathy Potel
01 44 53 45 14

chef de publicité
Imane Erraoui
01 44 53 45 06
responsable marketing et commercial
Alexandre Corbier
01 44 53 45 00
vente au numéro
service abonnement
Zamila Nguyen
zamila.nguyen@archimag.com
réclamations infos :
zamila.nguyen@archimag.com
BP 95-92244 Malakoff Cedex
tarifs et conditions d'abonnement
valables jusqu'au 31-12-2021
France : 1 an, 137 euros
France : 2 ans, 237 euros
Tarif étudiant : 1 an, 30 euros
Tarif demandeur d'emploi : 1 an, 65 euros
Vente au numéro : 19 euros

imprimeur
Inore Groupe Impression
4 rue Thomas Edison
58640 Varennes Vauzelles
éditeur
IDP Sarl, au capital de 40 000 euros
Information, documentation, presse
Numéro de commission paritaire : 1221 T 85484
ISSN : 2260-166X
Dépôt légal à parution du numéro
crédits photos
Couverture :
© Alengo/Stock
Page 50 haut : © Guillaume Bourdages/Unsplash
Intérieures : droits réservés, sauf mentions différentes

annonceurs
Ammareal : 2
Cotranet : 11
Crescendo : 41
IDP : 2^e de couverture, 9, 13, 19, 3^e et 4^e de couverture
GM Invent : 37
Numilog : 27
Serda Formation : 21
Xelians : 23, 33
Archimag sur les réseaux sociaux

Facebook
→ www.facebook.com/pages/archimag/102327699812643

Twitter
→ twitter.com/ArchimagRedac



Archimag est une publication du groupe Serda. Toute adaptation ou reproduction même partielle des informations parues dans Archimag est formellement interdite sauf accord écrit d'IDP SARL.



Ce document est imprimé sur papier certifié PEFC

Annoncez-vous sur Archimag et Archimag.com
Contactez Cathy Potel : 01 44 53 45 14, cathy.potel@archimag.com
et Imane Erraoui : 01 44 53 45 06, imane.erraoui@archimag.com
Abonnez-vous à Archimag : www.archimag.com/boutique



sommaire

[actualités]

- 04 quelle gouvernance des données pour la France ?
- 05 Conseil de l'Europe : archives à la BNU de Strasbourg
- 06 Centre Pompidou : 2023-2026, fermeture pour travaux
- 07 Paris, ville la plus numérique de France
- 08 3 questions à Emmanuelle Larroque
- 09 une direction à deux têtes pour le CNNum
- 10 UiPath : levée de fonds record
- 12 la start-up du mois : Yelda



12

DR

[dossier]

- 14 **vrac numérique : comment mettre de l'ordre**



L'entassement croissant et pas toujours cohérent de documents sur les serveurs est un phénomène qui touche les entreprises comme les administrations. Quelle est l'origine de ce vrac numérique ? Comment se manifeste-t-il ? Quelle est son ampleur ? Pour faire face, à quelle théorie archivistique peut-on se rattacher ?

Sommaire p.15

[outils]

- 24 les nouveaux horizons de la bibliométrie
- 26 livres numériques en bibliothèque : PNB est-il trop éloigné des réalités du terrain ?
- 29 la renaissance de la Librairie des ducs de Bourgogne
- 30 ils ont archivé le confinement
- 32 numérisation : sous quel format conserver les documents ?
- 34 logiciels pour bibliothèques : petite progression
- 42 opendata et datavisualisation : que faire des données qui dorment dans des tableurs ?
- 44 mettre en place une dynamique de veille sous pression
- 46 favoris : des moteurs de recherche écolos
- 47 Emilie Berthod : documentaliste musicale en quête d'harmonie



29

KBR

[perspectives]

- 48 Pascal Boniface : « l'intelligence artificielle est la nouvelle frontière de la rivalité des puissances »
- 50 archimag store : applis, cadeaux, beaux livres...
- 52 dans les archives d'Archimag, mars 2001



50

plus d'infos tous les jours



sur Archimag.com

Conseil de l'Europe : archives à la BNU de Strasbourg

La Bibliothèque nationale et universitaire de Strasbourg (BNU) va prochainement accueillir les archives du Conseil de l'Europe. Ce corpus documentaire couvre la période allant de 1949 à nos jours et comprend les publications produites par le Conseil de l'Europe, mais aussi celles commandées à des institutions extérieures. La volumétrie de ce fonds d'archives est estimée à environ 150 mètres linéaires.

Ces archives viendront compléter les collections de l'ancienne bibliothèque du Conseil de l'Europe qui ont déjà été déposées à la BNU, près de 45 000 monographies et plus de 900 périodiques, soit environ 2 kilomètres de collections.

recrutement d'un catalogueur

« Cet ensemble documentaire, unique et cohérent, contribue significativement à enrichir les collections d'excellence de la BNU sur les questions européennes et renforce encore davantage le partenariat avec le Conseil de l'Europe », souligne Geoffrey Girost en charge du service sciences humaines et sociales au sein de la BNU.

Ce déménagement sera suivi par le recrutement d'un catalogueur en vue de l'intégration de ces nouveaux fonds.

Fondé en 1949, le Conseil de l'Europe est une organisation internationale composée de 47 États membres rassemblant environ 830 millions de personnes. ■

Donald Trump, destructeur compulsif d'archives ?

Que restera-t-il du mandat de Donald Trump à la Maison-Blanche ? Peu d'archives en tout cas si l'on en croit l'agence de presse AP. L'ancien président avait en effet pour habitude de déchirer et jeter des documents liés à sa présidence sans autre forme de procès. Selon Solomon Lartey, ancien analyste des archives de la Maison-Blanche, des responsables de la National archives and records administration (Nara) ont pourtant rappelé à Donald Trump qu'il n'était pas autorisé à détruire des documents officiels : « Ils lui ont dit d'arrêter de le faire. Il ne voulait pas s'arrêter ». Mis devant le fait accompli, l'ancien archiviste explique : « Mon directeur est venu me voir et m'a dit : "vous devez scotcher les morceaux ensemble" ».

La documentation produite au format numérique n'a guère été mieux traitée. Selon un ancien conseiller, certains agents de la Maison-Blanche avaient pris l'habitude d'utiliser leur messagerie personnelle au détriment des outils de communication officiels.

Des historiens redoutent de découvrir un « trou béant » lorsqu'il s'agira d'accéder aux archives de la période 2017-2021. Certains d'entre eux ont annoncé des poursuites contre l'ancien président. ■

les tweets du mois

Dr.Pierre-Al.Chalze @docteurbagarre
En réponse à @PAMM_JU
La dématérialisation est trop vue comme un outil de contrôle des coûts, alors que c'est avant tout un démultiplicateur de force.

Connaissancesarts @Cdearts · 17 fév.
Leonard de Vinci : le chateau du Clos Lucé va proposer une immersion numérique dans l'œuvre du génie de la Renaissance
#leonard #leonarddevincent #leonard #leonarddevincent #leonarddevincent
#leonarddevincent #leonarddevincent #leonarddevincent

La France au Vietnam @FranceAuVietnam · 20 fév.
Là ! La collection Patrimoines Partagés de @hallof s'enrichit d'une magnifique bibliothèque numérique 🇫🇷 en coopération avec la Bibliothèque nationale du Vietnam !
Plus de 2000 documents des deux bibliothèques sont en ligne : manuscrits, cartes...
#hallof #france-vietnam

Arthur Sarazin @arsarazin
Et si vous voulez re-faire un #ChallengeData, toute la documentation est en libre accès sous licence CC-BY-SA : méthodes, tutoriels vidéos, templates de livrables, tout est réutilisable !
Ici : opendatacanvas.org/challenge-data

Clémence René @clémencerene
Pourquoi aller vers le #Digital ?
L'usage d'internet en 2021 dans le monde c'est :
◆ 5,22 milliards d'utilisateurs mobiles uniques
◆ 4,66 milliards d'internautes
◆ 4,20 milliards d'utilisateurs actifs des réseaux sociaux
#TransfoNum #MBADM8

INDICATEUR	VALEUR	POURCENTAGE
UTILISATEURS MOBILES UNIQUES	5,22 milliards	66,4%
INTERNET	4,66 milliards	59,5%
UTILISATEURS ACTIFS DES RÉSEAUX SOCIAUX	4,20 milliards	53,6%

■ ■ ■ ■ suivez-nous sur Twitter :
→ twitter.com/ArchimagRedac



- 17 traiter un vrac numérique, mode d'emploi
- 20 outils : cerner le vrac numérique
- 22 vrac numérique et situation de crise au ministère des Armées

vrac numérique : comment mettre de l'ordre

L'entassement croissant et pas toujours cohérent de documents sur les serveurs est un phénomène qui touche les entreprises comme les administrations. Quelle est l'origine de ce vrac numérique ? Comment se manifeste-t-il ? Quelle est son ampleur ? Pour faire face, à quelle théorie archivistique peut-on se rattacher ? En pratique, une méthode rigoureuse permet de le résorber. De plus, différentes sortes d'outils logiciels sont à disposition pour cette gestion. Retour d'expérience au ministère des Armées.

doublons, triplons, fichiers périmés, nommage approximatif, versions redondantes... Ces fléaux sont aujourd'hui parfaitement identifiés par les professionnels de la gouvernance documentaire. Les causes, elles aussi, sont bien connues : des serveurs organisés en silos, des applications dispersées sur lesquelles les salariés produisent et partagent leurs documents de travail, le manque de cohérence, etc. Et au rythme actuel de la croissance des volumes d'information, toutes les organisations seront confrontées, à un moment ou à un autre, à un problème de

dispersion de leurs fichiers numériques. Selon l'«*Abrégé d'archivistique*» qui fait autorité en la matière (1), le vrac numérique désigne un ensemble de documents bureautiques très hétérogènes se caractérisant par une absence fréquente de règles de classement et de nommage, l'utilisation de caractères susceptibles de poser des problèmes, un stockage aléatoire et une arborescence construite sans logique prédéfinie. Lors de l'édition 2020 du salon Documation, la question du vrac numérique a fait l'objet de plusieurs conférences et confirmé une tendance : «*Ce déluge*

informationnel provoque de multiples désordres : perte de temps, risque de disparition de fichiers, redondance, conservation d'information inutile ou obsolète», constate Nouredine Lamriri (Everteam). Tout salarié peut en effet témoigner du temps perdu à rechercher une information située dans un document qu'il a du mal à localiser : dans sa messagerie ? sur le disque dur de son ordinateur ? sur les serveurs de l'entreprise ? dans les nuages ? Selon certaines études, un salarié passe environ 7 heures et 30 minutes par semaine



livres numériques en bibliothèque : PNB est-il trop éloigné des réalités du terrain ?

Le déploiement du dispositif de prêt numérique en bibliothèque (PNB) se poursuit à un rythme régulier dans les bibliothèques françaises. Mais le modèle économique est dénoncé par Réseau Carel qui pointe les conditions imposées par les éditeurs.

Près de sept ans après son lancement, PNB se porte plutôt bien. Le dispositif prêt numérique en bibliothèque est aujourd'hui installé dans 260 réseaux de bibliothèques et semble en passe de convaincre les bibliothécaires malgré des réticences que l'on continue d'observer ici et là. Le catalogue de PNB propose désormais plus de 70 000 titres contre 17 500 en 2015. Le nombre de prêts, quant à lui, a connu une très forte hausse passant de 67 493 en 2015 à 851 000 en 2020.

À l'occasion d'une journée d'étude de Réseau Carel qui s'est tenue au mois d'octobre dernier à la bibliothèque publique d'information du Centre Pompidou, plusieurs retours d'expérience ont été présentés aux participants. Notamment celui de la bibliothèque départementale de la Vienne qui a déployé PNB au mois de septembre 2017. Aujourd'hui, une soixan-



taine d'établissements partenaires sont en mesure d'intégrer cette bibliothèque numérique baptisée « Lire en Vienne », soit un bassin de 160 000 habitants. Fin 2020, le service comptait 28 000 inscrits.

Les statistiques d'emprunt confirment la curiosité des usagers pour le support numérique. En 2019, 7 424 prêts avaient été enregistrés; en 2020, ce chiffre atteignait 10 427. Dans le même temps, le nombre de lecteurs au format numérique passait de 669 à 1 002. « Nous avons enregistré un bond des prêts lors du premier confinement », explique Stéphane Joseph, responsable du pôle informatique de la bibliothèque départementale de la Vienne. Rappelons que, lors du premier confinement, les Français se sont tournés

vers le livre numérique faute de pouvoir se rendre dans les librairies. Selon le baromètre des usages du livre numérique réalisé par la Sofia, le SNE et la SGDL, 27 % des Français ont découvert le livre numérique pendant le confinement et la moitié des 15-34 ans sont aujourd'hui des lecteurs de livres numériques.

un quasi-doublement du budget d'acquisition dans la Vienne

L'offre éditoriale de la bibliothèque départementale de la Vienne est composée de romans pour adultes (85 %), de romans destinés à la jeunesse (10 %) et de documents divers (5 %). Lors du lancement

logiciels pour bibliothèques :

petite progression

Pour une année atypique comme l'a été 2020 marquée par la pandémie, les confinements et la fermeture des bibliothèques, la légère hausse du marché des logiciels est appréciable. En coulisses, les rachats de sociétés se poursuivent. La 27^e enquête de Tosca Consultants livre son analyse du marché.

1. un chiffre d'affaires en progression

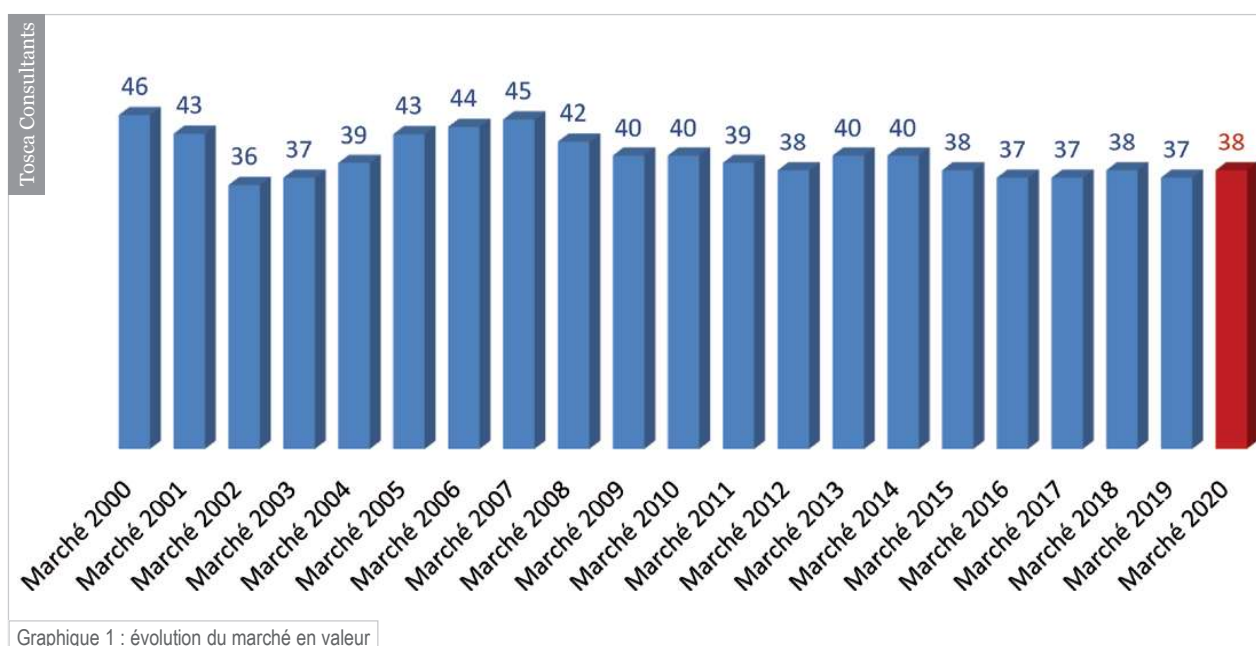
Le marché 2020 des logiciels pour bibliothèque représente un chiffre d'affaires de 38 millions d'euros contre 37 millions d'euros l'année précédente. Cette modeste progression est plutôt une bonne surprise, la pandémie avec ses périodes de confinement et son report des élections municipales laissant craindre une nouvelle baisse du marché. Les adaptations fournies par les éditeurs permettant de mettre en œuvre « *biblio drive* » ou service de « *click and collect* » ont peut-être permis de compenser la baisse du nombre de contrats.

Parmi les 45 fournisseurs recensés, 9 réalisent un chiffre d'affaires de plus d'un million d'euros dans le secteur des bibliothèques. Par ordre décroissant du chiffre

d'affaires réalisé auprès des bibliothèques françaises, il s'agit de C3RB Informatique, d'Archimed, de Decalog, d'Ex Libris France, du groupement AFI-BibLibre, de Réseau Canopé, de PMB Services, d'Infor France SAS et de BiblioMondo France. Ces 9 fournisseurs détiennent 76 % des parts de marché. Les 36 autres fournisseurs se partagent un chiffre d'affaires d'un peu plus de 9 millions d'euros, soit une moyenne de 254 000 euros par entreprise (voir graphique 1 ci-dessous).

2. baissent du volume des ventes

Hors Réseau Canopé qui comptabilise ses ventes en nombre d'abonnements annuels et Ebsco Information Services qui n'a pas communiqué le volume de ses ventes, les fournisseurs déclarent avoir vendu un total de 1887 produits en 2020



mettre en place une dynamique de veille sous pression

Nous avons tous vécu des moments où la pression pour mener un projet nous imposait une cadence inhabituelle. Ces moments sont en même temps stressants et intéressants car ils nous imposent de trouver des chemins de traverse, des nouvelles voies, pour arriver plus rapidement à destination. La veille n'est pas épargnée par ces « coups de bourre » professionnels. Il est donc important de voir comment l'on peut mettre en place une dynamique de veille avec quelques raccourcis salutaires.

1. qui peut être concerné ?

La mise en place d'une dynamique de veille rapidement et sous pression va intéresser différents acteurs. Non seulement les petites ou moyennes entreprises qui vont vouloir aller vite avec des outils gratuits, mais aussi les grandes entreprises. Soit qu'elles doivent mettre rapidement une veille en place, généralement parce que le haut management s'est emparé du sujet, soit parce qu'une entité doit pouvoir agir rapidement et avec agilité et ne veut pas passer par un service centralisé de veille qui ferait les choses certainement très bien, mais avec des délais allongés.

Partons du principe que le feu vert vient de tout en haut de la hiérarchie. Que votre supérieur hiérarchique direct semble jouer sa promotion sur la réussite de ce projet. Et que vous-même n'avez jamais vu une telle effervescence pour la veille ! Vous êtes décidé à jouer votre va-tout. Voyons donc comment faire vite et bien pour mettre en place une dynamique !

2. quelle méthodologie de mise en place ?

■ définition des besoins

Cette première phase passe normalement par des entretiens avec les collaborateurs finaux, clients de la veille. Ceci pour avoir la meilleure vision des besoins réels de l'organisation (qui a besoin de quoi, quand, sur quel support...) et de l'existant en matière de veille. Cette phase est souvent passionnante car elle vous fait voyager dans l'organisation et vous fait rencontrer un certain nombre de collaborateurs. Mais dans notre cas, il faudra se contenter d'une perception réduite. Pas ou peu d'entretiens donc dans cette première étape. De toute façon, les lignes directrices ont été données par les grands chefs. Elles sont floues — comme souvent ! —, mais l'on va devoir s'en contenter.

■ choix des outils

Cette seconde phase est plus technique, forcément. Les choses sérieuses commencent ! Trois cas de figure se présentent.

❶ Si vous n'avez pas d'outil et peu de budget, il sera possible de se rabattre sur le gratuit ou le quasi gratuit. Avec Inoreader, nous avons là un agrégateur de flux RSS de qualité. À 50 euros d'abonnement annuel, pas d'appel d'offres, pas de demande de budget. Tout va bien.

❷ Maintenant, si votre entreprise a la

chance d'avoir précédemment débloqué les fonds pour l'abonnement à une plateforme de veille professionnelle, et que vous avez été formé, il va falloir mettre les mains dans le cambouis. Ce n'est pas désagréable. Et je trouve personnellement que l'on conduit une voiture d'autant mieux que l'on comprend le rôle de l'embrayage, des pistons ou des bougies. C'est la même chose pour la conduite de la veille qui se fait d'autant mieux que l'on comprend ce que sont un flux RSS, des opérateurs de recherche ou des expressions régulières.

❸ Vous voulez en profiter pour réclamer un budget et acquérir une plateforme ! C'est bien joué... mais c'est aussi dangereux. Si vous avez les coudées franches pour choisir l'outil, pourquoi pas. Mais si vous devez passer par le service achats avec un process rigoureux, mais lent, vous risquez de rater le coche.

■ paramétrage : intégration des sources et création des règles

Durant cette phase, le temps sera votre ennemi. Il faudra faire vite. Selon la célèbre loi de Pareto, vous ferez 80 % du paramétrage en 20 % du temps. Et tant pis pour les igno-lages. Voici quelques idées pour booster la recherche des sources pertinentes.

■ On peut commencer par récupérer ses favoris pour se faire une base de liens utiles. On pourra aussi demander aux collègues concernés par la veille de faire de même.

■ On peut aussi parfois tomber sur des listes de sources. Par exemple, la requête [filetype:xls lemonde.fr/lesechos.fr liberation.fr RSS] vous donnera une jolie petite liste de flux de la PQN prêts à être intégrés à votre agrégateur. On peut tenter d'autres pistes, comme une interrogation Twitter avec la requête [filter:links +les mots clés