

ENTRETIEN

« La connaissance n'est pas réductible à l'information »

Cette année 2016 marque le centenaire de la naissance de l'Américain Claude Shannon, l'un des pères de la théorie de l'information : l'occasion de revenir sur ce terme abondamment employé... et sur ses limites.



Jean-Louis ERMINE est professeur émérite à l'institut Mines-Télécom et expert en gestion des connaissances.

POUR LA SCIENCE

Pouvez-vous rappeler comment la théorie de Shannon définit l'information ?

JEAN-LOUIS ERMINE : La notion forgée par Shannon dans le cadre de sa théorie des télécommunications est assez éloignée de la signification courante du mot information. On considère une source qui émet un message, codé à l'aide de signaux élémentaires, à travers un canal de communication, et un récepteur qui doit décoder ce message. Supposons que le récepteur ait N décodages possibles du message, chacun ayant une certaine probabilité. La quantité d'information ou entropie de Shannon du message est alors, par définition, la quantité $-(p_1 \log p_1 + \dots + p_N \log p_N)$, où p_i est la probabilité associée à la i -ème possibilité et où le logarithme est, par convention, celui de base 2.

PLS

Quel est le rapport avec les bits d'information ?

J.-L. E. : Le bit est l'unité de mesure de l'information au sens ci-dessus. Or les techniques informatiques utilisent des signaux binaires représentés par 0 ou 1 et qu'on appelle aussi bits. Si les signaux 0 et 1 sont de même probabilité (égale à $1/2$), alors l'information correspondant à un bit informatique

est égale, selon la définition de Shannon, à $-(1/2) \log(1/2) - (1/2) \log(1/2)$, ce qui fait 1 bit d'information (car $\log(1/2) = -1$).

PLS

L'information à la Shannon est-elle un concept satisfaisant ?

J.-L. E. : En informatique et dans les technologies de la communication, oui. Mais l'information à la Shannon, ce n'est pas l'information du 20 heures à la télévision ! Elle ne mesure pas du tout le contenu sémantique des messages, c'est-à-dire leur signification. Elle ne rend donc pas compte des nombreuses acceptions qu'a, dans le langage courant, le mot information. Beaucoup de chercheurs ont essayé d'élaborer des définitions de l'information au sens sémantique qui soient précises et si possible quantitatives. Après des décennies de recherches et de polémiques, on arrive à peu près à un consensus appelé GDI (*General Definition of Information*), que l'on peut résumer en disant qu'une information est un ensemble structuré de données dotées d'un sens pour les acteurs du domaine concerné.

PLS

Pouvez-vous préciser la différence entre données et informations ?

J.-L. E. : Une donnée désigne un fait brut, une observation, une perception, un signal de

capteur... Par exemple, on peut avoir pour donnée le nombre 454 ; ce nombre devient une information lorsqu'on dit qu'il est la valeur de l'indice boursier CAC 40 à telle date. En d'autres termes, pour que des données deviennent des informations, il faut les traiter pour leur donner du sens. Les bases de données sont ainsi plutôt, à proprement parler, des bases d'informations : les données ne sont pas seulement rassemblées, elles sont aussi classées d'une certaine façon, complétées par des précisions, ce qui fait d'elles des informations, élémentaires certes, mais des informations quand même.

PLS

Vous êtes un spécialiste de gestion des connaissances. Comment s'articulent les notions d'information et de connaissance ?

J.-L. E. : Une connaissance est une information qui prend du sens dans un contexte donné. Elle n'est pas réductible à l'information. Voici deux exemples simples pour mieux comprendre. Les bases de données permettent parfois de conférer du sens à leurs contenus, mais elles ne fournissent pratiquement jamais de contexte d'interprétation ; elles produisent des informations (rudimentaires), mais pas de connaissances. C'est un peu différent avec le moteur de recherche de Google. Lorsqu'on lui soumet

une requête formée de plusieurs mots, il fait un lien sémantique entre ces mots et livre des résultats qui ont un sens (une liste de sites en rapport sémantique avec la requête). Mais il ne fait pas que cela : les sites indiqués sont ceux ayant les meilleurs scores, c'est-à-dire les plus consultés, ce qui correspond à un contexte d'usage. De ce point de vue, le moteur de Google constitue un outil de recherche de connaissances.

PLS

Pourquoi faut-il « gérer » les connaissances ?

J.-L. E. : La connaissance est une capacité humaine attachée à des compétences qui relèvent à la fois des individus et de la collectivité. Or quand des personnes ayant des connaissances ou savoir-faire particuliers quittent l'entreprise ou l'organisme où ils travaillent, ces connaissances peuvent être perdues, car une grande partie d'entre elles (on estime près de 80 %) sont tacites et non consignées quelque part. Ainsi, la Nasa n'a plus les connaissances qui lui ont permis d'envoyer des astronautes sur la Lune il y a quarante ans. Et il y a une multitude d'exemples d'entreprises qui ont eu ainsi des pertes de connaissances ou de savoir-faire ayant entraîné des pertes financières. Cela représente globalement des sommes énormes, et les entreprises en prennent de plus en plus conscience. Il faut donc gérer les connaissances, de façon à ce qu'elles soient bien partagées dans la structure en question et que leur transmission soit assurée dans le temps.

PLS

Dans quelle mesure cette problématique concerne-t-elle le métier de chercheur ?

J.-L. E. : Les chercheurs en science fondamentale sont moins exposés aux risques de pertes de connaissances en raison du système bien installé des publications dans des revues professionnelles, où les travaux et leurs résultats sont décrits (et préalablement évalués par des spécialistes). Mais ils ne sont pas complètement à l'abri, parce que dans ce domaine aussi, il y a du tacite, des savoir-faire et des façons

de raisonner qui n'apparaissent nulle part explicitement. Le problème se pose davantage pour la recherche appliquée ou industrielle, où les scientifiques ne sont pas maîtres de leurs objectifs et où les savoir-faire techniques occupent en général beaucoup plus de place.

PLS

Que proposez-vous pour bien gérer les connaissances ?

J.-L. E. : Il y a plus d'une vingtaine d'années, à l'époque où très peu de gens étaient conscients de cet enjeu, j'ai commencé à développer une méthodologie nommée MASK, qui s'implémente à l'aide de divers outils. Le CEA, en France, a été le premier à l'appliquer, en 1993. Depuis, elle a mûri et a fait l'objet, en France et dans le monde, d'interventions dans beaucoup d'organisations (ou entreprises) publiques et privées.

PLS

En quoi consiste-t-elle ?

J.-L. E. : Dans ses grandes lignes, la méthode MASK procède en quatre étapes. La première consiste à effectuer une analyse stratégique du patrimoine de connaissances de l'entreprise, c'est-à-dire à déterminer quelles sont les connaissances qui sont critiques par rapport aux objectifs et à la stratégie de cette entreprise. La deuxième étape est la création et l'organisation d'un « référentiel » de connaissances qui contient tout ce dont les collaborateurs dans l'entreprise ont besoin pour travailler. Cela implique d'organiser les systèmes d'information, d'interroger toutes les personnes détenant des connaissances tacites critiques, d'explicitier ces dernières, etc. La troisième phase est l'organisation du partage des connaissances au sein de la structure, à l'aide notamment d'outils collaboratifs (wikis, réseaux sociaux d'entreprise...). Dans la dernière, enfin, il s'agit de mettre en place des outils et des méthodes susceptibles de stimuler la créativité et de favoriser l'innovation. ■

Propos recueillis par Maurice MASHAAL

■ POUR INFO

Jean-Louis Ermine interviendra dans la table ronde « Extraire l'information, retrouver la connaissance » qui se tiendra de 16 heures à 18 heures le samedi 19 novembre 2016 à la Bibliothèque nationale de France, à Paris, dans le cadre de la 15^e rencontre *Physique et interrogations fondamentales*.

Cette rencontre, organisée par la Société française de physique et la BNF, a pour thème « La science, l'information, la connaissance ». Inscriptions gratuites (dans la limite des places disponibles) et programme sur <http://pif15.sfpnet.fr>

■ BIBLIOGRAPHIE

J.-L. Ermine (dir.), *Management et ingénierie des connaissances : modèles et méthodes*, Hermes-Lavoisier, 2008.

Association pour la gestion des connaissances dans la société et les organisations (AGECSO) : www.agecso.com

Club Gestion des connaissances : www.club-gc.asso.fr