

# L'information<sup>1</sup>

Louis-Claude Paquin

Professeur au Département des communications

Université du Québec à Montréal

*“La plus grande crise qui menace la civilisation moderne sera la façon de transformer l'information en connaissance structurée.”*

*Carlos Fuentes*

## **La surcharge informationnelle**

Depuis les cinquante dernières années, l'essor des technologies de diffusion a fait croître de façon exponentielle la quantité d'information disponible. En effet, nous sommes littéralement bombardés d'informations par les journaux et revues, la radio et la télévision et, depuis peu, par les réseaux informatiques.

L'information a toujours été pour nous une denrée vitale en nous permettant de maîtriser notre environnement. La privation d'information induit des états hallucinatoires, alors que la surcharge provoque de l'anxiété. Au niveau des organisations, l'acquisition et la circulation de l'information revêtent de plus en plus une importance capitale. L'information est venue à être considérée comme un produit dont la valeur est proportionnelle aux avantages qu'elle procure à ceux qui l'obtiennent.

Or, ce n'est pas parce que plus d'informations sont disponibles que nous sommes mieux informés, que notre compréhension du monde est meilleure et que notre vision est plus large. Bien au contraire, l'accroissement

---

<sup>1</sup> À paraître dans les *Actes du Séminaire du Mont Orford: Les Sciences du texte juridique : l'interprétation du droit : le sens, l'interprète, la machine*. Établissements Émile Bruylant (Bruxelles).

du volume d'information peut entraîner une réduction du focus, un effet de “tunnel”. L'information souhaitée est souvent difficile à discerner et à obtenir dans la masse des informations disponibles qui sont variées et disparates.

Nos capacités de traiter et d'assimiler l'information. sont limitées, ce qui entraîne un phénomène d'exformation, soit une accumulation d'informations disponibles qui ne sont pas traitées faute de temps et de compétence. Dans l'état actuel des programmes disponibles, l'ordinateur ne nous est pas d'un grand secours, il participe plutôt à la croissance du problème.

En plus de générer de l'anxiété, la surabondance entraîne une confusion entre les données brutes et l'information, entre les faits et la connaissance.

### **Les deux types informations**

Dans la foulée de l'avènement de la téléphonie<sup>2</sup>, le terme “information” a exercé une séduction croissante sur les intellectuels et les scientifiques, car il connotait le dynamisme, la modernité, la scientificité et la technicité. Alors que ce avait toujours été associé à la communication et à la connaissance, son acception a connu une restriction majeure pour désigner tout ce qui est envoyé par un canal mécanique ou électrique et, par extension, biologique que cela soit signifiant ou pas pour le récepteur.

La très grande popularité de cette acception qui s'est répandue dans la plupart des domaines de la science et de la technologie est source de confusion entre la capacité de transmettre des signaux ou des données avec la capacité de fabriquer des messages signifiants.

---

<sup>2</sup> Shannon et Weaver, *The Mathematical Theory of Communication*, 1949. Le traitement optimal de l'information qui est visé consiste à augmenter le débit du signal qui est appelé “information”. Cette théorie n'aide en rien à décider ce que l'on doit en faire ou encore à y attribuer une signification, mais prescrit comment transmettre une information de façon inaltérée en présence de bruit et autres interférences avec le signal.

De sorte que la plupart de ce qui est pris actuellement pour de l'information est en fait des données brutes. C'est ainsi que les patrons de "bits" qui circulent à l'intérieur des ordinateurs sont appelés information. Ces patrons ne seront jamais que des symboles dont la manipulation est régie par des règles syntaxiques très strictes : les langages de programmation<sup>3</sup>. L'ordinateur ne comprend pas la signification des symboles qu'il manipule!

Le signal ou la donnée ne peut devenir information que sous certaines conditions : ils doivent s'inscrire dans une activité relationnelle, être dotés de caractéristiques formelles qui les rendent recevables par le récepteur et qu'ils soient signifiants. Au delà d'une réponse béhavioriste, ils doivent s'intégrer aux structures cognitives de ce dernier. La donnée brute ou le signal doivent être mis en forme, mis en contexte, analysés et interprétés pour devenir de l'information.

### **Modèle graduel de l'information et de sa valeur**

- les données brutes dont le volume est souvent très grand, qui sont disparates, éparses et qui sont destinées à d'autres fins que la communication et la connaissance, soit à répertorier des faits ;
- les données organisées qui sont le résultat d'une extraction et d'une disposition en fonction d'une focalisation particulière et dont la finalité est informationnelle c'est-à-dire qu'elles sont destinées à l'analyse et à l'interprétation de certains faits ;
- l'information qui est le résultat d'une analyse et d'une interprétation de données organisées dans un contexte défini et en fonction d'un modèle préconstruit ;
- les indications de l'information qui sont ajoutées à l'information sous la forme, soit d'un marquage de la

---

<sup>3</sup> J. Haugeland, *L'esprit dans la machine*, 1985 trad. fr. 1989, p. 107 et sq.

structure (dans le cas d'un document, la liste des sous-titres), soit d'une indexation des objets et concepts impliqués, soit de liens avec d'autres informations ;

- l'information secondaire, c'est-à-dire une condensation de l'information qui en indique la valeur : cette condensation peut être sommaire, le clip genre "fiche de lecture" offrant les mots-clés et les concepts principaux ainsi que des indications signalétiques ; la condensation peut être plus complète, le schéma montre les relations entre les concepts et le résumé rend compte plus intégralement de l'argumentation ;
- l'information élaborée ou synthèse qui est littéralement de l'information sur les informations en ce que différentes informations sur un thème précis sont colligées, présentées de façon analytique et intégrées à un domaine de connaissance; cette information est dite stratégique en ce qu'elle facilite la prise de décision.

Il est à noter que le statut peut varier selon les individus : pour certains des données organisées constitueront de l'information, alors que pour d'autres il ne s'agira que de données brutes.

Tout comme la valeur de l'information, les coûts de médiatisation et les délais de diffusion vont croissant. Ajouter des indications à l'information requiert une main d'oeuvre qualifiée, alors que la production d'informations élaborées (synthèses) ne peut qu'être le fait d'experts.

### **Le traitement de l'information**

Le traitement de l'information par les humains implique une série d'activités mentales aussi complexes les unes que les autres : la perception, l'analyse et l'interprétation.

La perception est l'activité qui internalise l'information. Elle implique trois opérations : le déchiffrement, la segmentation, et le filtrage. Déchiffrer c'est distinguer ce qui fait partie de l'information de ce qui n'en fait pas

partie : les marges du texte, les taches des images, le son du bruit, etc. Segmenter c'est découper ce qui est déchiffré en unités signifiantes, symboles et énoncés. Filtrer c'est opérer une sélection des énoncés pour ne retenir que ceux qui sont pertinents en fonction d'un projet ou de préoccupations. Percevoir c'est aussi prédire dans la mesure où, en présence d'ambiguïtés ou de bruit, le récepteur peut quand même discerner l'information; dans le cadre des théories cognitives, on dirait que ce dernier applique des "prototypes" qu'il a en tête<sup>4</sup>.

L'analyse consiste à mettre à jour la structure des énoncés retenus et, par-delà, l'organisation de l'information. Les opérations de l'analyse sont le décodage, la catégorisation et le classement. Décoder c'est associer un référent aux symboles et identifier les relations qu'ils entretiennent, par rapport à un langage donné qui délimite l'ensemble des symboles admissibles et leur référence et qui en dicte les règles d'assemblage en des énoncés. Catégoriser c'est rendre explicites, mentalement ou en ajoutant des étiquettes, les multiples niveaux des structures formelles et signifiantes. Classer c'est regrouper les éléments constitutifs de l'information en fonction de critères provenant à la fois du besoin d'information et des connaissances. Le contexte et le cadre de référence sont importants, car ils permettent de suppléer à l'absence ou l'équivocité, tant de la référence d'un symbole que de relation explicite entre ceux-ci.

L'interprétation est la moins connue des activités mentales. Il s'agit d'une intégration de l'information traitée à des schèmes existants cognitifs et culturels pré-existants. L'information est alors comprise et peut être utilisée dans un raisonnement, pour accomplir une action, pour solutionner un problème ou encore plus

---

<sup>4</sup> Pour une vue d'ensemble de l'application de la notion de "prototype" aux sciences du langage, voir : J. R. Taylor, "Linguistic Categorization", *Prototypes in Linguistic Theory*, Oxford, 1989.

simplement pour produire des informations nouvelles. Les schèmes permettant l'interprétation peuvent aussi se trouver modifiés par l'intégration de nouvelles informations ce qui permet l'adaptation des individus à leur environnement.

Bien qu'elles ont été présentées séquentiellement, ces activités sont interdépendantes : pour percevoir et sélectionner une information, il faut déjà un peu l'interpréter. De plus, ces activités mentales sont intriquées à des états psychologiques qui en conditionne la réalisation. Parmi les états psychologiques mentionnons les émotions, les croyances, les passions, les intentions, les espoirs, les craintes, le plaisir, etc.

Selon les états psychologiques particuliers des individus, leurs connaissances et leur expérience, c'est-à-dire les événements vécus, les mêmes informations entraîneront parfois des réponses et des comportements différents. Elles seront pertinentes pour certains, alors que pour d'autres elles ne le seront pas.

### **La production de l'information**

Les phases de la production de l'information sont la conception, la médiatisation, la diffusion et la gestion.

Concevoir l'information c'est effectuer une représentation de la pensée. Cette représentation s'effectue en deux temps : l'abstraction et la mise en symboles. L'abstraction est un processus de simplification, de condensation et de réduction de la pensée pour focaliser sur ce que l'on veut communiquer. Le niveau de généralisation est fonction du niveau des aspects particuliers retenus.

Médiatiser c'est encoder et présenter ce qui a été abstrait de la pensée dans un système symbolique. Les symboles sont des entités communicationnelles qui tiennent lieu des choses représentées tout en permettant leur organisation en des énoncés selon un code. Ces symboles sont réalisés dans un système de signes : l'écriture (mots et chiffres), l'iconique (images

fixes ou animées, graphiques, schémas), le sonore ou encore un mélange des trois : le multimédia. Ces symboles sont inscrits sur un support à l'aide d'une technique, ce qui leur confère la permanence, mais qui entraîne une déperdition de la genèse, l'intention et la pensée de l'émetteur.

Diffuser c'est transmettre l'information au récepteur; cela peut se faire par le dialogue, l'inscription, les mass-média et depuis peu par les réseaux informatiques.

Le dialogue permet une diffusion optimale de l'information : le récepteur peut ajuster le débit, réclamer des clarifications, etc. Cette diffusion est toutefois limitée dans le temps et l'espace.

L'inscription permet une diffusion beaucoup plus large d'un volume beaucoup plus grand d'information et à beaucoup plus de personnes, éloignées dans le temps et l'espace, surtout depuis l'avènement de l'imprimerie. Cependant, la qualité de cette diffusion souffre de ce que le récepteur ne peut en valider sa compréhension.

Les mass-media ont augmenté de façon substantielle la quantité et le débit tout en réduisant les délais de diffusion. L'information a été diffusée à des gens qui, jusque là, n'y avaient pas accès. Par contre, l'accès est contrôlé par les producteurs et la maximisation des profits prime souvent sur d'autres considérations.

Les nouvelles technologies issues du couplage des ordinateurs et des télécommunications (les autoroutes électroniques ou optiques) pour constituer des réseaux de bases de données à l'échelle quasi-planétaire sont entrain de transformer la diffusion de l'information. L'utilisateur peut à sa guise choisir son interlocuteur ou son service dans le réseau.

Gérer l'information c'est accumuler les archives qui la renferment, s'assurer de leur conservation et de leur intégrité. Il s'agit là de la fonction des bibliothèques, des cinémathèques, des audiothèques, des vidéothèques.

La production de l'information touche donc des aspects économiques, technologiques, médiatiques et culturels.

### **L'accès à l'information**

L'accès à l'information le plus simple est le déroulement linéaire d'une source : écouter une émission de télévision, feuilleter un imprimé, faire défiler les pages-écran, etc. Il s'agit là d'un accès très peu efficace sinon aléatoire face aux besoins spécifiques du récepteur. Ce dernier doit donc pouvoir exercer un contrôle qui peut aller de la sélection de la source d'information jusqu'à la formulation d'une requête spécifique.

Pour qu'une requête puisse être formulée et adressée à un ensemble d'informations, des indications doivent y avoir été ajoutées. Ces indications sont de trois types : des indications signalétiques, des indices de classification et des descripteurs.

Les indications signalétiques décrivent l'archive qui comporte l'information parmi celles-ci on retrouve : l'auteur, le titre, la date, le bas de vignette, la provenance, le nom du fichier, etc. Si ces indications permettent de retrouver les archives, elles sont peu utiles pour la recherche de l'information qu'ils contiennent.

Les indices de classification organisent l'information en fonction d'un critère discriminant. Les critères sont limités : l'alphabet (les dictionnaires), le temps (les chronologies), le lieu (les atlas), un continuum (des meilleures aux pires créances) ou une division d'un champ du savoir ou de la science en "domaines" (droit pénal, civil de la famille, etc.). Le critère facilite l'accès à l'information mais en le restreignant puisque en regard d'autres critères, l'information peut se retrouver éparpillée.

Les descripteurs ou mots-clés ajoutés lors d'une indexation permettent un accès à l'information par les concepts à l'oeuvre dans une information. Pour permettre un accès constant et cohérent à l'information,



les descripteurs utilisés doivent être contrôlés, c'est-à-dire constitués en langage . Le thésaurus est ce langage prescrivant l'éventail des descripteurs admissibles et des relations (synonymie, généralité, spécificité, associativité) qu'ils entretiennent.

L'indexation est une opération longue et coûteuse car elle nécessite une analyse souvent fine de l'information. Il s'agit toutefois de la seule façon de s'assurer que toute l'information pertinente (exhaustivité), mais seulement celle-ci a été dépistée (bruit). De toute façon, cette analyse est incontournable, si ce n'est pas fait lors de la diffusion de l'information, c'est accompli individuellement par les récepteurs.

Par ailleurs, une barrière à la découverte de l'information pertinente est de ne pas savoir ce que l'on veut ou d'être incapable de le formuler en une requête. Dans ce cas la compréhension de la structure de l'information et de son insertion dans un ensemble permet de briser la linéarité de la réception pour "naviguer" d'une information à une autre. La plupart du temps, cette organisation est explicitée par une table des matières.

Pour faciliter la compréhension de la structure, il suffit lors de la médiatisation d'ajouter à l'information des marques dotées d'un statut particulier et d'une forme prescrite. Ces marques peuvent être typographiques : les sous-titres en gras, les idées secondaires en retrait, etc. Ces marques peuvent aussi prendre d'autres formes comme des icônes. Ces marques peuvent être réalisées directement ou de façon déclarative au moyen d'une norme comme SGML (Standard Generalized Markup Language)<sup>5</sup>

Pour faciliter la navigation dans informations, des liens conceptuels peuvent être établis entre divers constituants lors de la médiatisation. Un réseau conceptuel est ainsi constitué sur l'ensemble d'informations pour favoriser un accès associatif à

---

<sup>5</sup> adoptée par l'ISO (8879) en 1986.

l'information. Dans le cas des textes, ce type de réseau est appelé hypertexte.

## **Conclusions**

Il apparaît nécessaire de mieux comprendre les fonctions cognitives reliées au traitement de l'information et de développer des modèles mieux adaptés.

Il apparaît également nécessaire de développer et d'enseigner de nouveaux langages médiatiques en fonction des nouvelles technologies pour tirer le meilleur parti des capacités humaines étant données les limitations du médium.

Il apparaît nécessaire de développer des programmes informatiques afin d'augmenter la capacité de compréhension de l'information des récepteurs comme on l'a fait avec la gestion et la diffusion de l'information.

Le paradigme de la simulation qui est celui de l'intelligence artificielle s'est avéré inefficace à faciliter la compréhension de l'information car si des analyseurs syntaxiques performants ont pu être développés, on s'est buté au mur de la connaissance de sens commun.

Par ailleurs, il apparaît impossible de développer des "applications" universelles ou clés en main, car le traitement de l'information dépend de la culture locale et de l'utilisation qui est faite de l'information.

Un paradigme prometteur est celui de la coopération où les ordinateurs et les humains sont les composantes d'un même système. Chacune des composantes est utilisée de façon optimale selon ses capacités : l'humain pour ses capacités cognitives et l'ordinateur pour sa puissance de calcul appliquée au traitement formel du signal et des données.