

LA FABRIQUE DE L'INTELLIGIBILITÉ DES DONNÉES DANS LES ORGANISATIONS

Appel à articles numéro 64 – Décembre 2023

Data Intelligibility Manufacturing in Organizations

Résumé :

L'enjeu de ce numéro est de s'interroger sur le sens attribué aux données qui prolifèrent dans les organisations. Il s'agit, dans un contexte où le travail conjoint avec les informaticiens et les statisticiens est une réalité organisationnelle prégnante, d'envisager les choix épistémologiques, méthodologiques et professionnels qui donnent du sens aux données organisationnelles issues du monde digital. L'objectif consistera à comprendre comment s'y prennent les organisations pour passer d'un entrepôt de données accumulées à une plateforme de ressources mobilisables et exploitables pour leur conférer du sens.

Coordination :

Marie-Hélène HERMAND, Maître de conférences en Sciences de l'Information et de la Communication, Université Bordeaux Montaigne, MICA – marie-helene.hermand@u-bordeaux-montaigne.fr

Jean-Michel CAMIN, Docteur en Sciences de l'Information et de la Communication, Ingénieur de recherche, Orange – jean-michel.camin@u-bordeaux-montaigne.fr

Argumentaire :

Selon la formule « *Data is the new oil* » du mathématicien Humby (2006 ; World Economic Forum, 2011), les données sont assimilables à un actif économique au même titre que le pétrole. Dans ce contexte, les organisations sont confrontées à une problématique essentielle qui consiste à faire acquérir une valeur ajoutée aux données qu'elles produisent afin de les rendre intelligibles. C'est à cette question du sens des données organisationnelles, articulée à celle de la pratique professionnelle des outils numériques, que nous proposons de consacrer ce numéro.

Pour garantir l'intelligibilité de leurs données, les organisations doivent tenir compte de contraintes diverses liées au cycle de vie des données : les choix d'infrastructure et de stockage, les conditions de lisibilité, les coûts et risques organisationnels induits, le circuit économique, les modalités de délivrance des données par des spécialistes ou des sous-traitants, les multiples usages situés des données organisationnelles ou encore l'accompagnement des équipes professionnelles et des publics dans l'usage des données. Le travail avec les informaticiens et les statisticiens étant une réalité dans bon nombre d'organisations, une réflexion s'impose sur les choix épistémologiques, méthodologiques et professionnels qui donnent du sens aux données organisationnelles issues du monde digital. Par données, nous entendons les textes, les nombres, les calculs, les images et les sons, aucune de ces données n'étant brute et toutes transitant par des outils et des pratiques pour proliférer dans les organisations. Par monde digital, nous entendons la convergence entre les secteurs (informatique, téléphonie, médias), les équipements (serveurs, terminaux, bases de données, connexions, systèmes d'information) et la diversité des outils mobilisés dans la sphère professionnelle (smartphones, tablettes, applications, robots, automates, etc.) (Benedetto-Meyer et Boboc, 2021).

Le travail interdisciplinaire constitue un enjeu majeur de ce numéro qui propose de réunir des contributions issues de la communication des organisations et, de manière non exhaustive, des sciences du langage, de la sociologie, des sciences de gestion, de l'informatique, de l'IA, de la techno-anthropologie, de la philosophie ou du droit. L'objectif consistera à comprendre comment s'y prennent les organisations pour passer d'un entrepôt de données accumulées à une plateforme de ressources mobilisables et exploitables pour leur conférer du sens.

Les données dans les organisations : une interrogation large et récurrente en SIC

Depuis une dizaine d'années, l'importance des données dans les organisations est traitée selon divers angles dans *Communication & Organisation* : Moinet (2009) s'est penché sur le concept d'information utile pour le relier à ceux de la connaissance stratégique et de l'intelligence économique ; Lloveria (2014) a observé comment la data-visualisation laisse place au *data-design*, puis au *data-storytelling* ; Le Deuff (2015) s'est intéressé aux utopies documentaires qui redonnent du pouvoir à l'individu ; Lehmans (2017) s'est interrogée sur l'ouverture des données pour redéfinir la culture de l'information organisationnelle.

D'autres revues en SIC se sont intéressées aux données en contextes organisationnels. Dans *COSSI*, Marcon (2018) s'est penché sur la manière dont les PME s'approprient le discours sur le *big data*. *Les Cahiers du numérique* ont par exemple traité du *big data* dans le secteur médical (2016, vol. 12/1-2), de la visualisation des données en institutions éducatives (2016, vol. 12/4), de l'archivage des données dans les institutions éducatives et culturelles (2015, vol. 11/3) ou encore des usages de données dans les services publics (2013, vol.9/1). La revue *Études digitales* s'est quant à elle intéressée au gouvernement des données en étudiant la rationalité algorithmique comme un mode inédit de gouvernementalité (2016,

n°2). La revue *Études de communication* s'est penchée sur la transformation des données en connaissances (2016, n°43). La *Revue des interactions humaines médiatisées* s'intéresse quant à elle à la recherche d'information, à la fouille de données, à la structuration des archives, à la sémantisation des contenus et à l'exploitation des traces qui transforment les relations avec les usagers. La question des données organisationnelles est aussi régulièrement traitée sous l'angle de l'intelligence artificielle (IA), par exemple en contexte d'innovation sociale (trois numéros de *Technologie et développement* en 2022, 2021, 2020) ou du point de vue des débats normatifs qu'elle suscite (*Réseaux* n°232-233 en 2022).

La question de l'intelligibilité : une problématique traitée par de nombreuses disciplines

Les sciences cognitives et la linguistique se sont saisies de la question de l'intelligibilité en l'associant à quatre principaux critères : le critère de lisibilité – d'ailleurs critiqué pour son inaptitude à rendre compte de la complexité des contenus (Beudet, 2001) –, de cohésion (Maingueneau, 1991 ; Bronckart, 1996), de cohérence (Charolles, 1993) et de pertinence (Sperber & Wilson, 1986).

Pour appréhender la complexité de l'élaboration du sens, les théories de l'argumentation et de l'énonciation ont pour leur part mis en évidence les stratégies linguistiques (Ducrot, 1984), cognitives (Vignaux, 1988) et énonciatives (Culioli, 1990 ; Kerbrat-Orrechioni, 1980) ainsi que la logique des arguments (Perelman et Olbrechts-Tyteca, 1970 ; Plantin, 1990 ; Breton, 1996) et le rôle des présupposés et des stéréotypes (Angenot, 1980 ; Amossy 1997).

Saillante en contexte numérique depuis le début des années 2000, la question de l'intelligibilité a été traitée par le collectif Pédauque (2006) à partir de la notion de document numérique. Représentant un large spectre disciplinaire (informatique, linguistique, ingénierie des connaissances, SIC, télécoms, gestion, sciences sociales), Pédauque a lié le processus d'intelligibilité à quatre propriétés du document : la mémorisation, l'organisation des idées, la créativité et la transmissibilité du document.

Sensible au « fossé d'intelligibilité » causé par la perte de mémoire documentaire, Bachimont (2010) considère que l'intelligibilité d'une archive décontextualisée doit être restaurée. Le chercheur propose d'élaborer une « rhétorique de l'intelligibilité » (Bachimont, 2020) en interrogeant le récit des étapes aveugles de manipulation et de calcul opérées sur les données. Dans la revue *Intelligibilité du numérique*, il souligne cette question du sens et invite à dépasser les frontières disciplinaires pour la traiter.

L'intelligibilité des données : une problématique légitime pour la communication des organisations

Ces approches non ancrées en communication des organisations rencontrent certaines de ses préoccupations, montrant par-là que la communication des organisations ne peut être laissée à l'écart de la question qui touche au sens des données. Quelques chiffres permettent de situer l'ampleur du phénomène :

- les données fournies par l'indice 2021 de l'économie et de la société numériques (DESI, Commission européenne) montrent deux tendances : les grandes entreprises adoptent davantage les (nouvelles) technologies que les PME ; les grandes entreprises utilisent, par ordre décroissant, un progiciel de gestion intégré (à 80%), les médias sociaux (55%), la e-facturation (54%), le cloud (48%), la vente en ligne (38%) et le big data (35%) ;
- il existe aussi un important écart d'appropriation de l'IA entre PME et grandes entreprises : ce sont surtout ces dernières qui utilisent le *machine* ou *deep learning* (13% d'entre elles), l'analyse du langage écrit (10%), le traitement des images (10%), le traitement du langage parlé (8%), la gestion des

mouvements (6%) et la génération de langage parlé ou écrit (5%) ;

- selon une étude publiée par Dell et l'Institut pour le futur, 85% des emplois de 2030 n'existent pas encore : l'intelligence artificielle et la robotique, qui transforment en profondeur les métiers existants, vont en créer de nouveaux ;

- en 2015 en France, 62 % des entreprises utilisatrices de *big data* traitent des données de géolocalisation, 32 % exploitent des données issues des médias sociaux et 29 % des informations relevées par des objets intelligents, connectés ou des capteurs.

De multiples entraves organisationnelles doivent en outre être prises en compte : la dépendance au terrain lors du processus de recueil, la qualité variable et l'intelligibilité plus ou moins nette des données, les problèmes terminologiques, le point de vue défendu par les modélisations adoptées, les choix graphiques de représentations formelles ou encore la manière dont les outils digitaux utilisés pour manipuler les données cadrent notre perception des organisations.

Les propositions de contributions

Un premier axe donnera l'occasion de proposer et de discuter des cadres théoriques permettant d'appréhender la fabrique de l'intelligibilité des données organisationnelles. Sans ambition d'exhaustivité, nous citons quelques cadres repérés et des exemples de leur mobilisation. Le courant de la médiologie initié par Debray dans les années 1990 a facilité l'étude du transport du sens dans l'espace et dans le temps (Merzeau, 1998). La théorie de l'acteur-réseau et la sociologie de la traduction (Akrich, Callon, & Latour, 2006) ont été mobilisées dans la détection de l'incertitude des données (Camin, 2014). La sémantique interprétative et les principes fondateurs de la textométrie (Pincemin, 2011) ont permis d'observer le processus d'attribution de sens à des organisations territoriales expérimentales (Hermand, 2020). L'approche sémiotique construite autour de la sémantique différentielle (Bachimont *et al.*, 2011) a permis d'étudier la restauration des conditions d'intelligibilité d'archives institutionnelles patrimoniales (Leyoudec, 2015). Le cadre des affordances (Treem & Leonardi 2013) a amené Lambotte *et al.* (2018) à interroger l'intelligibilité des données massives issues de l'activité siconumérique des entreprises. La pragmatique de la communication numérique a permis de se pencher sur l'influence de la structuration de données sur l'intelligibilité mutuelle (Marcoccia, 2016). La technosémiotique vise pour sa part à faciliter l'étude de l'influence des dispositifs techno-éditoriaux sur les effets de sens (Julliard, 2015). La sémiotique doxastique (Stockinger, 2017a) ouvre quant à elle la voie à la description, à l'analyse et à l'interprétation critique d'« univers de sens » véhiculés dans les écosystèmes digitaux élaborés par les organisations (Hermand, à paraître).

Un deuxième axe permettra d'interroger les cadres méthodologiques et les outils de manipulation des données et du sens. Pour ce qui concerne les méthodes de restitution et de représentation formelle, peuvent être explorés les héritages et les réactualisations des cadres conceptuels et cognitifs (Minsky, 1975), des scripts ou des scénarios (Schank & Abelson, 1977), des réseaux contractuels (Smith, 1980) ou des graphes conceptuels (Sowa, 1984). Par exemple, la conception de méthodes d'extraction du sens et de thématisation de données a fait l'objet de travaux consacrés à la médiatisation d'archives digitales institutionnelles issues du monde de la recherche (Stockinger, 2017b) ou encore à la médiatisation du patrimoine des bassins miniers (Beloued *et al.*, 2015). Un autre exemple montre comment s'intéresser aux nuages de tags pour visualiser des corpus structurés de liens hypertextes dans des sites web des universités (Pinède & Lespinet-Najib, 2013). Un dernier exemple montre le recours à l'iconographie des corrélations et l'utilisation de méthodes statistiques en vue d'étudier les proximités sémantiques dans le discours (Camin, 2018).

Un troisième axe sera consacré aux expertises et aux métiers. Les professionnels de la donnée (analystes de données, architectes de données, statisticiens, géomaticiens, etc.) ne sont pas les seuls à « faire parler » les données et à travailler leur valeur cognitive. D'autres professionnels sont amenés à se pencher sur la qualité et la valorisation des données (juristes, financiers, traducteurs, rédacteurs, communicants, sociologues, ergonomes) en fonction de leur expertise ou de la spécialisation thématique de leur organisation. Parmi les initiatives remarquables qu'il serait utile d'étudier, on identifie notamment :

- le travail de *start up* spécialisées dans l'exploitation et l'enrichissement de données ;
- les opérations de valorisation des organisations par les données ;
- la mutualisation de données par des organisations territoriales ou à différents échelons territoriaux ;
- les laboratoires (*data lab*) et les observatoires engagés dans la détection de l'implicite des données ou d'actes de langage remarquables ;
- les démarches participatives et coopératives visant à augmenter le volume de données organisationnelles.

Calendrier :

- Envoi des propositions sous forme de résumés : 20 janvier 2023
- Retour aux auteurs de la sélection des propositions : 27 janvier 2023
- Remise de l'article intégral : 27 avril 2023
- Retour aux auteurs de l'évaluation par le comité de lecture : 5 juin 2023
- Retour des articles définitifs (revisés après évaluation) : 11 septembre 2023
- Publication du numéro : 4 décembre 2023

Consignes de rédaction des propositions (résumés) :

- 6000 caractères, espaces compris
- Bibliographie non comptabilisée dans le nombre de caractères
- Sur une page de garde : titre de la proposition, prénom et nom de l'auteur, université, laboratoire, adresse électronique, cinq mots-clés
- Le résumé doit permettre de bien identifier la problématisation, le cadre théorique et conceptuel, la méthode, les analyses et la discussion.

Les propositions seront envoyées aux adresses suivantes :

- marie-helene.hermand@u-bordeaux-montaigne.fr
- jean-michel.camin@u-bordeaux-montaigne.fr

Consignes de rédaction des articles définitifs :

- 35 000 caractères, espaces compris, bibliographie comprise
- Les normes de mise en page des articles définitifs sont accessibles en ligne sur le site de *Communication & organisation* <https://journals.openedition.org/communicationorganisation/5909>
- La mise en forme finale selon les normes fournies conditionnera l'acceptation définitive de l'article.
- L'évaluation des articles complets sera faite en double aveugle par le comité de lecture de la revue. La liste des membres du comité de lecture est accessible à l'adresse suivante :

À propos des coordinateurs du numéro :

Marie-Hélène Hermand est maître de conférences en sciences de l'information et de la communication à l'Université Bordeaux Montaigne et chercheuse au laboratoire MICA-Médiations Informations Communication Arts (MICA – UR 4426). Qualifiée également en sciences du langage, elle a notamment publié l'ouvrage *Eurorégions. L'éclosion de la communication transfrontalière* (2020). Ses recherches actuelles portent sur la modélisation sémantique de la communication des organisations émergentes à l'échelle transfrontalière.

Jean-Michel Camin occupe un poste d'ingénieur de recherche chez Orange. Il est ingénieur en télécommunication, docteur en sciences de l'information et de la communication de l'Université Bordeaux Montaigne et chercheur associé au laboratoire MICA. Ses travaux de recherche portent sur l'incertitude et la communication, la confiance et l'identité dans les environnements numériques.

Références :

Akrich, M., Callon, M., Latour, B. (2006), *Sociologie de la traduction : textes fondateurs*. Presses des Mines.

Amossy, R. (1997). *Stéréotypes et clichés*. Nathan.

Angenot, M. (1982). *La parole pamphlétaire*. Payot.

Bachimont, B. (2020). Rhétorique de l'intelligibilité, vers des humanités du numérique. Communication à la MSH Sud de Montpellier, 28 février 2020, consultable sur : <https://projet.numerev.com/lab/65-rhetorique-de-l-intelligibilite-vers-des-humanites-du-numerique>

Bachimont, B. (2020). La présence de l'archive : réinventer et justifier. *Intellectica*, 53-54, 281-309.

Bachimont, B., Gandon, F., Poupeau, G., et al. (2011). Enjeux et technologies : des données au sens. *Documentaliste-Sciences de l'Information*, 4/48, 24-41.

Beaudet, C. (2001). Clarté, lisibilité, intelligibilité des textes : un état de la question et une proposition pédagogique, *Recherches en rédaction professionnelle*, 1/1, 1-19.

Beloued, A., Lalande, S., Stockinger, P. (2015). Modélisation et formalisation RDFS/OWL d'une ontologie de description audiovisuelle, *Les Cahiers du numérique*, 11/3, 39-70.

Benedetto-Meyer, M., Boboc, A. (2021). *Sociologie du numérique au travail*. Armand Colin.

Breton, P. (1996). *L'argumentation dans la communication*. La Découverte.

Bronckart, J.-P. (1996). *Activité langagière, textes et discours*. Delachaux et Niestlé.

Camin, J.-M. (2018). *Trust, risk and vulnerability in the context of identity management. A culturonomics*

perspective. Poster présenté au colloque *SRA Annual Meeting, Society of Risk Analysis*. New-Orleans (Louisiana, USA), 2-6 décembre 2018.

Camin, J.-M. (2014). *Approche communicationnelle de l'incertitude dans les projets innovants en phase de lancement*. [Thèse de doctorat en sciences de l'information et de la communication]. Université Bordeaux Montaigne.

Charolles, M. (1994). Cohésion, cohérence et pertinence du discours. *Revue internationale de linguistique française*, 29, 125-151.

Culioli, A. (1990). *Pour une linguistique de l'énonciation*. Ophrys.

Ducrot, O. (1984). *Le dire et le dit*. Éditions de Minuit.

Hermant, M.-H. (à paraître). Une approche modélisante de l'univers du discours transfrontalier : les discours constituants et les récits édifiants dans le portail web des Transfrontier Peace Parks. *Semen*, 52.

Hermant, M.-H. (2020). *Eurorégions. L'écllosion de la communication transfrontalière*. Éditions de l'Université de Bruxelles.

Julliard, V. (2015). Les apports de la techno-sémiotique à l'analyse des controverses sur Twitter. *Hermès*, 73/3, 191-200.

Kerbrat-Orecchioni, C. (1980). *L'énonciation*. Armand Colin.

Lamotte, F., De Meyère, D., Andry, T., Dumont de Chassart, P. (2018). L'intelligibilité des données massives de l'activités socionumérique : de la conception du dispositif à la construction de sens en contexte. *Communication et Organisation*, 54, 149-167.

Le Deuff, O. (2015). Utopies documentaires : de l'indexation des connaissances à l'indexation des existences. *Communication et Organisation*, 51, 93-106.

Lehmans, A. (2017). Données ouvertes et redéfinition de la culture de l'information dans les organisations. *Communication et Organisation*, 51, 93-106.

Leyoudec, L. (2015). Reconstruire les conditions d'intelligibilité du document numérique patrimonial : mobilisations documentaires et sémiotique des *Linked Open Data*. *Les Enjeux de l'information et de la communication*, 16/2, 99-112.

Lloveria, V. (2014). Data design-moi un mouton. *Communication et Organisation*, 46, 99-112.

Maingueneau, D. (1991). *L'analyse du discours*. Hachette.

Marcocchia, M. (2016). *Analyser la communication numérique écrite*. Armand Colin.

Marcon, C. (2018). Les PME croient-elles aux données massives (big data) ? ». *Revue COSSI*, 4,

consultable sur : <https://revue-cossi.info/numeros/n-4-2018-big-data-thick-data/704-1-2018-revue-marcon>

Merzeau, L. (1998). Ceci ne tuera pas cela. *Les Cahiers de médiologie*, 6/2, 27-39.

Minsky, M. (1975). A framework for representing knowledge, in P.H. Winston (dir.), *The Psychology of Computer Vision*. McGraw-Hill.

Moinet, N. (2009). De l'information utile à la connaissance stratégique : la dimension communicationnelle de l'intelligence économique. *Communication et Organisation*, 35, 215-225.

Pédaque, R. T. (2006). Document : forme, signe et médium, les re-formulations du numérique, consultable sur : https://archivesic.ccsd.cnrs.fr/sic_00001741/fr/

Perelman, C., Olbrechts-Tyteca, L. (1970). *Traité de l'argumentation*. Presses de l'Université de Bruxelles.

Pinède, N., Lespinet-Najob, L. (2013). Visualiser des corpus structurés de liens hypertextes sous forme de nuages de tags : atouts et limites, communication, consultable sur : <https://www.canal-u.tv/chaines/fmsh/contextes-langues-et-cultures-dans-l-organisation-des-connaissances-ixe-colloque-5>

Plantin, C. (1990). *Essais sur l'argumentation*. Kimé.

Pincemin, B. (2011). Sémantique interprétative et textométrie – Version abrégée. *Corpus*, 10, 259-269.

Schank, R., Abelson, R. (1977). *Scripts, Plans, Goals and Understanding. An Inquiry into Human Knowledge Structures*. Lawrence Erlbaum.

Sowa, J.F. (1984). *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley.

Sperber, D., Wilson, D. (1989). *La pertinence : communication et cognition*. Éditions de Minuit.

Stockinger, P. (2017a). « Chapitre 5. Éléments théoriques et méthodologiques pour une expertise des sites web » ». In Rouquette S. (dir.), *Site internet : audit et stratégie*. De Boek Supérieur, 133-158.

Stockinger, P. (avec Steffen Lalande and Abdelkrim Beloued) (2017b). « Studio Campus AAR: A Semantic Platform for Analyzing and Publishing Audiovisual Corpuse ». In Szoniecky, S., Bouhaï, N. (dir.), *Collective Intelligence and Digital Archives: Towards Knowledge Ecosystems*. John Wiley & Sons, 85 - 133.

Treem, J.W., Leonardi, P.M. (2013). Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association. *Annals of the International Communication Association*, 36/1, 143189.

Vignaux, G. (1988). *Le discours acteur du monde*. Ophrys.