



La data literacy

-

un état de l'art des discours scientifiques et professionnels

04 avril 2024 – MSH - Journée des doctorants
« De la pluridisciplinarité à l'interdisciplinarité en Sciences de
l'Information et de la Communication ».



Contextualisation

- Évolutions numériques, technologiques où les données prennent une nouvelle place au sein des pratiques professionnelles résumées sous le vocable de « révolution des données ».
- Nécessité d'une nouvelle compréhension et maîtrise des pratiques par le développement de littératies des données depuis le début des années 2000. Le *data literate* est cet individu formé attendu sur le marché du travail.
- La *data literacy* est une littératie, à savoir l'étude de la lecture et de l'écriture sur les sociétés telle qu'elle a pu être exposée par Jack Goody, Levy-Strauss, David Olson, Brian Street, etc.

Contextualisation

- La *data literacy* est la capacité d'accéder, d'interpréter, d'évaluer, de gérer, de manipuler et d'utiliser de manière critique et éthique les données grâce à une connaissance des caractéristiques de la donnée, de son cycle de vie et des différents impacts engendrés par son usage, notamment en termes de sécurité et de protection de la vie privée.
- Elle se présente à la fois comme un ensemble de compétences, de connaissances et de postures critiques. En corollaire, en tant que littératie, elle possède un versant pédagogique et requiert donc un apprentissage.
- Son étude est au carrefour des sciences de l'information, de l'éducation, de l'informatique, de l'anthropologie, de la sociologie, et bien d'autres disciplines.

Recherches complémentaires

Titre	Auteurs	Année	Revue
La <i>data literacy</i> distribuée : périmètres définitionnels, origines documentaires, perspectives réticulaires	Verdi, Ugo & Le Deuff, Olivier	2020	Les cahiers du numérique
Quelle(s) réponse(s) à l'enjeu d'acculturation aux données ? Un état de l'art des caractéristiques de la <i>data literacy</i>	Verdi, Ugo	2023	Revue française des sciences de l'information et de la communication
L'éthique des données dans les chartes éthiques des collectivités territoriales	Verdi, Ugo	2023	Communication & Organisation

Problématique et hypothèses

- **Problématique générale**
- Qu'est-ce que la littératie appliquée aux données ?
- **Questionnements**
 - Qu'est-ce qu'acculturer aux données ?
 - Quels sont les liens avec les autres concepts des données et littératies voisines ?
 - Quelles évolutions dans le temps ?
 - Selon quels acteurs et quelles thématiques ?
 - Existe-t-il une *data literacy* typiquement française ?

Problématique et hypothèses

- **Hypothèses**
- La *data literacy* est un **objet frontière**, à savoir « la capacité d'une idée, d'une théorie ou d'une pratique à se traduire au-delà de frontières culturellement établies » (Fox, 2011),
- Au même titre que David Olson qui parlait de cultures de l'écrit « avec des conséquences et des implications fort différentes les unes des autres », on doit parler de ***data literacies*** et non de « la » ***data literacy***,
- Les définitions et enjeux diffèrent selon le **domaine de rattachement** des acteurs (ex : universitaire, entrepreneur, institutionnel, etc.)

Méthodologie de thèse

- **Recherche** : Constitution d'un état de l'art scientifique et professionnel lors de la première année amenant à déterminer le produit de recherche final et le(s) type(s) de recherche de terrain à réaliser en conséquence.
- **Possibilité** : état de l'art des cadres de compétences et cursus sur les données afin de proposer un modèle à appliquer en contexte pédagogique avec une application et des retours d'expérience.
- **Choix effectué** : centrer les recherches sur le recueil et l'analyse de discours, sur la base des travaux du linguiste Dominique Maingueneau (de la deuxième à la troisième année).

Les discours

- En tant que « signes de richesse destinés à être évalués, appréciés, et des signes d'autorité, destinés à être crus et obéis » (Bourdieu, 2014, p.72), les discours construisent la représentation du réel que le locuteur souhaite faire partager par son allocataire (Seignour, 2011).
- Selon Maingueneau, ils permettent à la fois de désigner des objets d'analyse et de montrer que l'on adopte un certain point de vue. En effet, les discours sont des actes de communication par lequel « le locuteur autorise la révélation et l'objectivation partielles de son savoir et de sa volonté sous la forme d'un système de croyances contenu dans l'esprit de l'auditeur »

Les discours

- Reflètent des croyances et plus précisément des imaginaires, décrits par Charaudeau (2014) comme une « image de la réalité, mais en tant que cette image interprète la réalité, la fait entrer dans un univers de significations »

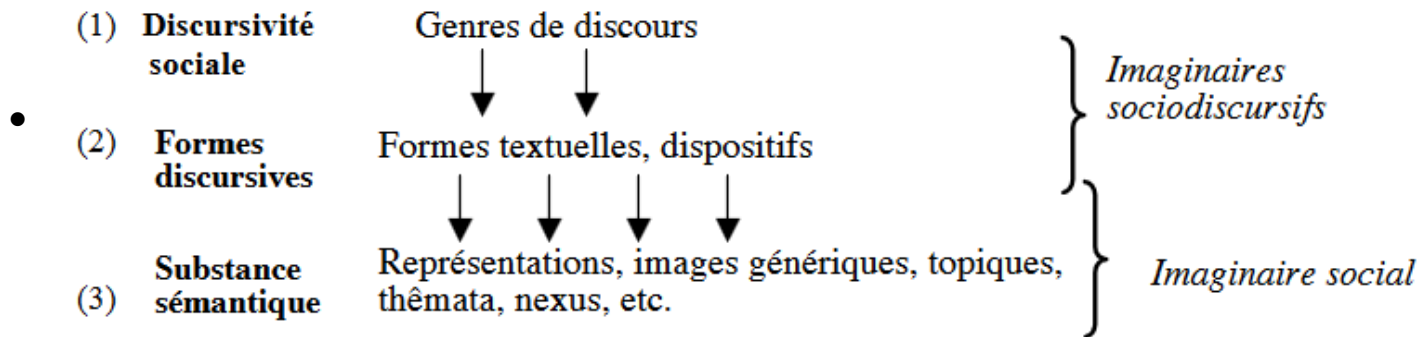
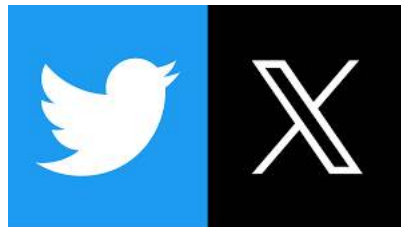


Figure 1 : *La sémiotique sociale ou les imaginaires du discours*

La constitution des corpus



Plusieurs centaines d'articles scientifiques, contenus professionnels français et internationaux (2021/2024)



44544 tweets intégrant le hashtag #dataliteracy (01/09/21 au 01/09/22)

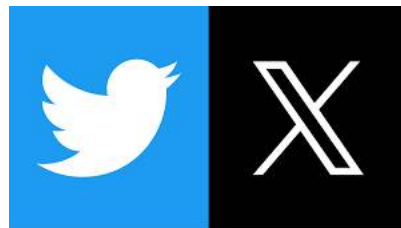


32 entretiens semi-qualitatifs auprès de 4 types d'acteurs français (2021 à début 2023)

La constitution des corpus



Plusieurs centaines d'articles scientifiques, contenus professionnels français et internationaux (2021/2024)



44544 tweets intégrant le hashtag #dataliteracy (01/09/21 au 01/09/22)



32 entretiens semi-qualitatifs auprès de 4 types d'acteurs français (2021 à début 2023)

Objectifs : déterminer la manière dont est discutée la *data literacy* et pouvoir ainsi repérer les éléments de langage, les acteurs qui portent ces discours, les sujets des données discutés, la spécificité des discours français, etc. Et plus globalement de comprendre l'écosystème des acteurs français.

La constitution des corpus



Veille

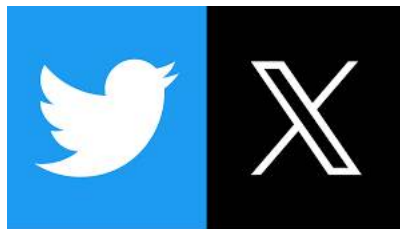
Google scholar,
Publish or Perish

Recherche

Bibliothèques, archives
et thèses similaires

Centralisation

Zotero
Zetlr



Absorption

API Twitter

Nettoyage et compilation

Excel et Open Refine

Analyse

langage R



Préparation

Guide d'entretien

Captation

Enregistrement visio

Transcription

Manuelle

Analyse

Grille méthodologique
Travaux de Maingueneau

La question de l'intégrité scientifique

- **Les questionnements**
 - Reproductibilité de la recherche : que mettre à disposition ? Sous quelle(s) forme(s) ?
 - Pratiques éthiques : droit d'auteur et autorisations à prendre en compte.
- **La méthodologie suivie**
 - Formulaire de consentement envoyé en amont aux interrogé.es,
 - Données stockées sur un serveur universitaire uniquement accessibles au doctorant,
 - Choix d'anonymisation générale,
 - Choix de publication des entretiens (avec autorisation) en annexe et lien pérenne sur *github*.

Résultats

- Les réseaux sociaux ne sont pas un lieu d'échanges et de conceptualisation mais de commercialisation où la *data literacy* sert avant tout de vitrine.
- La *data literacy* est effectivement un objet frontière au sens d'une appropriation par tous types d'acteurs (universitaires, entrepreneuriaux, associatifs, institutionnels). Il existe actuellement plus d'une trentaine de définitions, incluant des concepts spécifiques (état d'esprit critique, éthique, gouvernance, etc.) et des évolutions multiples catégorisées comme des littératies spécialisées.
- Les discours diffèrent selon les types d'acteurs non pas sur les caractéristiques générales de la *data literacy* mais sur les objectifs visés et sur les applications locales qui en découlent.

Limites et perspectives

- Sujet en évolution rapide dont le périmètre définitionnel devient de plus en plus nébuleux de par les interactions grandissantes avec les algorithmes et l'intelligence artificielle. Il en est de même pour l'écosystème des acteurs qui se modifie très rapidement.
- L'aspect didactique de conceptualisation de cursus, de mise en place pratique est abordé mais ne fait pas l'objet de la thèse et est un terrain d'études à explorer par la suite.
- Définir « le » système des données (donnés, IA, Algos...) va constituer le coeur de futures recherches.

Bibliographie

Littérature

- Goody, J. (1979). La raison graphique : La domestication de la pensée sauvage. Les Editions de Minuit.
- Goody, J., Privat, J.-M., & Maniez, C. (2007). Pouvoirs et savoirs de l'écrit. La Dispute.
- Olson, D. R. (2017). Goody et ses critiques. In É. Guichard (Éd.), Écritures : Sur les traces de Jack Goody (p. 30-38). Presses de l'enssib. <http://books.openedition.org/pressesenssib/1950>
- Street, B. (1985). Literacy in Theory and Practice. Cambridge University Press.

Analyse des discours

- Maingueneau, D. (1993). Analyse du discours et archive. Semen. Revue de sémio-linguistique des textes et discours, 8, Article 8. <https://doi.org/10.4000/semen.4069>
- Maingueneau, D. (2021). Discours et analyse du discours. Armand Colin. <https://www.cairn.info/discours-et-analyse-du-discours--9782200631710-p-3.htm>

Bibliographie

- **Objet frontière**
- Star, S. L. (2010). Ceci n'est pas un objet-frontière ! (M. B. Bah, Trad.). *Revue d'anthropologie des connaissances*, 4(1), Article 1. <https://doi.org/10.3917/rac.009.0018>
- **Data literacy**
- Verdi, U., & Le Deuff, O. (2020). La data literacy distribuée. Périmètres définitionnels, origines documentaires, perspectives réticulaires. *Les Cahiers du numérique*, 16(2-3-4), 137-173. Cairn.info. <https://doi.org/10.3166/LCN.2020.006>
- Verdi, U. (2023). Quelle(s) réponse(s) à l'enjeu d'acculturation aux données ? Un état de l'art des caractéristiques de la data literacy. *Revue française des sciences de l'information et de la communication*, 26, Article 26. <https://doi.org/10.4000/rfsic.14589>