



HAL
open science

Eléments de réponse au besoin des animateurs TRANSFORM en terme de sémantique

Rachel Boutrou

► **To cite this version:**

Rachel Boutrou. Eléments de réponse au besoin des animateurs TRANSFORM en terme de sémantique. 2021. hal-04674203

HAL Id: hal-04674203

<https://hal.inrae.fr/hal-04674203>

Preprint submitted on 21 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Éléments de réponse aux besoins des animateurs TRANSFORM en « sémantique » (R. Boutrou- Juin 2021)

La **sémantique** est une branche de la linguistique qui étudie les signifiés, ce dont on parle, ce que l'on veut transmettre par un énoncé.

Bibliographie

Une bibliographie est une liste structurée d'ouvrages ou d'autres documents, notamment d'articles, ayant des caractéristiques communes.

La bibliographie thématique regroupe des documents sur un même sujet ou un même thème. Elle donne la liste des documents utilisés pour réaliser le travail de recherche : faciliter la recherche intellectuelle, justifier et discuter le travail de recherche.

Bibliométrie

La bibliométrie est une méthode d'analyse quantitative utilisant les publications scientifiques et différents indicateurs pour mesurer la performance de la recherche, particulièrement 2 composantes importantes de la performance de la recherche:

- la production scientifique, soit la quantité de résultats scientifiques produits;
- l'impact scientifique, soit l'influence des résultats scientifiques sur la progression subséquente de la science.

Une définition plus contemporaine, associant bibliométrie, scientométrie d'outputs et infométrie, serait en termes d'« analyse quantitative de l'activité et des réseaux scientifiques ». Cette discipline, qui s'appuie sur l'analyse statistique des données et des réseaux, a un volet cognitif, en interaction avec les champs se donnant les sciences et les communautés scientifiques comme objet (économie de la connaissance, sociologie des sciences, épistémologie, histoire des sciences, etc.) et avec les sciences de l'information. Elle a un volet opérationnel, en liaison avec l'évaluation, le positionnement des acteurs et le management scientifique. Dans ce dernier cas, les études bibliométriques permettent de quantifier la communication d'un individu ou d'un groupe, non seulement en termes de volume, mais également de visibilité, d'influence, de partenariats, d'insertion dans les réseaux.

<https://fr.wikipedia.org/wiki/Bibliométrie>

Ainsi, la bibliométrie fournit aux chercheurs et aux institutions de l'information stratégique cruciale pour la mise en valeur de leurs résultats de recherche auprès de la communauté scientifique.

Les données bibliométriques sont utilisées pour :

- Mesurer et comparer la production scientifique de chercheurs, de groupes de recherche, d'institutions, de régions ou de pays par l'entremise d'indicateurs
- Identifier les revues les plus importantes ou influentes d'un domaine donné ;
- Suivre l'évolution dans le temps d'une discipline ou d'un sujet de recherche.

Terminologie

La terminologie est une discipline qui a pour objet l'étude théorique des dénominations des objets ou des concepts utilisés par tel ou tel domaine du savoir, le fonctionnement dans la langue des unités terminologiques, ainsi que les problèmes de traduction, de classement et de documentation qui se

posent à leur sujet (Dictionnaire Larousse en ligne). On appelle également terminologie l'ensemble des termes, rigoureusement définis, qui sont spécifiques d'une science, d'une technique, d'un domaine particulier de l'activité humaine. <https://fr.wikipedia.org/wiki/Terminologie>

La terminologie consiste en l'étude du choix et de l'usage des termes faisant partie des vocabulaires de spécialité, qu'on peut trouver dans tous les domaines de connaissance : informatique, grammaire, linguistique, mathématique, philosophie, médecine, musique, etc. Par exemple, le terme table relève à la fois de la terminologie de l'ameublement et de la terminologie de l'informatique.

Définition

D'après le CNRTL Centre National de Ressources Texturelles et Lexicales ;
<https://www.cnrtl.fr/definition/définition>

DÉFINITION, subst. fém.

A.- L'idée dominante est celle d'une limite ou d'un ensemble de traits qui circonscrivent un objet.

1. Détermination des limites d'un objet; résultat de cette action. Synon. délimitation, contour.
2. Détermination précise et concrète des caractères distinctifs d'un être.

Les définitions de mots, chez les lexicographes, ont pour but de faire connaître le sens d'un mot à ceux qui ont déjà une notion plus ou moins claire ou obscure, plus ou moins superficielle ou approfondie, de la chose que ce mot désigne.

« Mon métier consiste dans les mots, dans la connaissance du poids, du volume, de la densité des mots, leur maniement, leur usage, leur exacte application. (...) Mettre tout son effort en de justes et précises définitions oblige l'esprit à concevoir bien ». Bloch, Destin du Siècle, 1931, p. 20.

B.- LOG., usuel. Opération mentale qui consiste à déterminer les limites et le contenu d'un concept; proposition qui met en équivalence un être à définir, avec un ensemble d'attributs qui déterminent ses caractères essentiels :

« Au seuil de toute doctrine philosophique se rencontre une question inévitable : c'est la définition même de la philosophie. La définir, c'est déterminer la place qu'elle occupe dans la hiérarchie de nos connaissances, les rapports qu'elle soutient avec celles qui semblent les plus voisines, les parties dont elle se compose, la méthode qu'elle suit ». Ozanam, Essai sur la philos. de Dante, 1838, p. 91.

RB. On peut remplacer « philosophie » par « fonctionnalité »

Ontologie

[https://fr.wikipedia.org/wiki/Ontologie_\(informatique\)](https://fr.wikipedia.org/wiki/Ontologie_(informatique))

En informatique et en science de l'information, une ontologie est l'ensemble structuré des termes et concepts représentant le sens d'un champ d'informations, que ce soit par les métadonnées d'un espace de noms, ou les éléments d'un domaine de connaissances. L'ontologie constitue en soi un modèle de données représentatif d'un ensemble de concepts dans un domaine, ainsi que des relations entre ces concepts. Elle est employée pour raisonner à propos des objets du domaine concerné.

L'objectif premier d'une ontologie est de modéliser un ensemble de connaissances dans un domaine donné. Les ontologies décrivent généralement :

- individus : les objets de base ;
- classes : ensembles, collections, ou types d'objets1 ;

- attributs : propriétés, fonctionnalités, caractéristiques ou paramètres que les objets peuvent posséder et partager ;
- relations : les liens que les objets peuvent avoir entre eux ;
- événements : changements subis par des attributs ou des relations

Une autre définition, plus opérationnelle, peut être formulée ainsi :

« Une ontologie est un [réseau sémantique](#) qui regroupe un ensemble de concepts décrivant complètement un domaine. Ces concepts sont liés les uns aux autres par des relations [taxinomiques](#) (hiérarchisation des concepts) d'une part, et [sémantiques](#) d'autre part. »

Une ontologie est une représentation partagée et consensuelle entre les collaborateurs qui a pour but de se mettre d'accord sur un sujet particulier avec un objectif commun. Le but est de définir un ensemble de connaissances dans un domaine donné.

<http://www.journaldunet.com/developpeur/tutoriel/theo/070403-ontologie.shtml>.

Au niveau du Département TRANSFORM, travail mené par Magalie Weber : Construire une ontologie TRANSFORM sur les procédés alimentaires et non alimentaires

Les objectifs opérationnels du projet sont d'élaborer une méthode et de sélectionner les outils implémentant cette méthode qui permettront de i) harmoniser les 6 ontologies de domaine déjà développées en utilisant PO2 (dairy products, meatylab et biorefinery) et @Web (matter transfer, milk microfiltration et biorefinery), ii) étendre le vocabulaire sur le périmètre d'intérêt pour TRANSFORM à l'aide de différents cas d'étude (aliments et bioproduits), iii) déterminer les référentiels internationaux de référence pour TRANSFORM, iv) aligner les vocabulaires du périmètre TRANSFORM avec ces référentiels internationaux et v) tester la capacité de la méthode à gérer de manière efficace la mise à jour ultérieure du référentiel TRANSFORM.

Patrice : Avoir un vocabulaire standardisé > un référentiel pour une communauté donnée.

Définir une méthode pour enrichir le thésaurus avec des termes utiles à TRANSFORM

1. Un thésaurus constitué de μthésaurus, les structurer ; remettre ensemble l'existant
2. Faire valider les définitions (RB. rôle des animateurs ici ?!)

Utilité : annoter des tableaux de données pour les chercheursTRANSFORM

Rq : Caroline Pénicaud est impliquée dans ce travail

Revue biblio+ (Scoping)

Exemple de l'EMC SagA.

But : Réaliser un bilan sur l'utilisation des méthodes d'évaluation multicritère dans les travaux menés au sein de l'Institut.

Un groupe d'experts INRAE a mené une réflexion transdisciplinaire. Par un travail important d'annotation et d'analyse statistique d'un corpus de 954 articles publiés entre 2007 et 2017 par des scientifiques de l'Institut, le groupe a identifié les potentialités des méthodes d'évaluation multicritère.

Méthode d'élaboration du corpus:

- 4920 articles annotés MCDA (954) ou pas MCDA (3966)
- 954 articles MCDA annotés en 8 catégories de caractéristiques

Data are accessible in a public repository (<https://data.inra.fr/>; <https://doi.org/10.15454/WB51LL>)

L'obtention du corpus est décrit dans un data paper : <https://doi.org/10.1016/j.dib.2019.104204>

L'analyse du corpus de 954 articles -> 4 clusters dont les particularités sont décrites dans un article de recherche. <http://dx.doi.org/10.1007/s13593-020-00650-3>

Revue systématiques de la littérature et Méta-analyse

Une revue systématique de la littérature (RSL ou SLR in English) est une synthèse de connaissances déjà publiées dont la réalisation repose sur des principes scientifiques favorisant l'objectivité :

- répliquabilité
- transparence, archivage
- prise en compte des biais (extrinsèques, intrinsèques)

Elle a pour but l'objectivité des conclusions (minimiser les biais). Pour cela un encadrement exigeant (mais bienveillant!)

- démarche systématique
- protocole + revue sont publiés/peer-revue

RSL. Une procédure codifiée dont les étapes doivent impérativement figurer dans un PROTOCOLE, lui-même déposé sur un site dédié :

- Cochrane, <https://www.cochrane.org>
- Prospero, <https://www.crd.york.ac.uk/prospéro/>
- PRISMA <http://www.prisma-statement.org>

Impératif : Evaluation du biais et sources de biais, voire sur le processus d'évaluation.

	Narrative	Scoping	SLR
Protocol	No	Yes	Yes
Protocol deposit (i.e. Prospero)	No	No	Yes
Search strategy	No	Yes	Yes
Standardized data extraction form	No	Yes	Yes
Bias analysis (i.e. Rob assessment)	No	No	Yes
Synthesis of finding « summary of finding »	No	No	Yes

Une méta-analyse est une RSL dont on analyse mathématiquement, les résultats et surtout les biais.

Def : the analysis of analyses

Il y a un gradient de MA. Quand il y a des différences, c'est montré dans une MA standard. Le plus c'est de montrer ce qui explique cette différence, via méta-régression