



Cartographier le discours sur la littérature : l'ontologie comme modèle exploratoire

Motasem Alrahabi et Marine Riguet

Labex OBVIL - Sorbonne Université

Plan



- Les ontologies
- Analyse du discours critique sur la littérature
- Applications
- Perspectives

Qu'est-ce qu'une ontologie?



- Héritée initialement d'une tradition philosophique, une ontologie en informatique et en science de l'information est une description formelle des concepts dans un domaine de connaissance ou de discours [Gruber, 2009]:
 - Les concepts (classes) ont des attributs (propriétés) et sont instanciés par des objets (individus);
 - Ils sont liés les uns aux autres par des relations taxonomiques (hiérarchisation) et sémantiques.
- Difficultés: hiérarchisation des concepts, définir et arranger les classes et sous-classe, définir les attributs et les valeurs...
- Types: ontologies globales, - de domaine, - d'application [Guarino, 1998]
- Langages de spécification: OWL, RDF..., basés sur la logique du premier ordre, et représentent les connaissances sous forme d'assertion (sujet, prédicat, objet).

Analyse du discours critique sur la littérature



- L'analyse du discours sur la littérature permet de cerner les enjeux singuliers, à une époque donnée, de l'idée de littérature et des valeurs qui lui sont attachées.
- La critique littéraire se définit selon deux axes distincts (Macherey, 2004):
 - La critique comme appréciation (l'école du goût)
 - La critique comme savoir (la science de la production littéraire)
- Nous entendons par jugement critique sa première acception, à savoir l'expression d'une critique évaluative et normative de la littérature.

Principes d'une ontologie du discours critique



- Cerner l'expression du jugement de la critique littéraire en s'appuyant sur le schème linguistique suivant :

QUI (autorités) → DIT (modalités) → QUOI (objet) → COMBIEN (intensité) → COMMENT (valeurs)

→ **autorités critiques:** personnes, institutions, journaux...

→ **types d'énonciation:** affectif, évaluatif axiologique ou non axiologique, aléthique, déontique..

→ **champs de valeur convoqués:** esthétique, économique, politique, épistémique, idéologique..

→ **objets du discours:** œuvre, auteur, mouvement, savant...

Finalités et usages



- Développer l'ontologie dans un environnement numérique sous forme d'un outil de recherche, de découverte et d'exploration des textes littéraires
 - Parcourir les types de jugements émis sur un auteur, une oeuvre, un courant
 - Cibler un type de jugement particulier sur un objet (ex: quels éloges du théâtre, quelles condamnations du romantisme)
 - Comparer le jugement critique porté sur différents objets (ex: quel jugement porté selon le genre, entre les écrivaines et les écrivains)
 - Définir la singularité d'un jugement critique (ex: qu'est-ce qui distingue la critique de Sainte-Beuve de celle de Barbey d'Aurevilly ?)

Choix méthodologiques



- Corpus de travail :
 - Corpus critique constitué numériquement par le Labex OBVIL ([lien](#)), et comptant plus de 400 ouvrages (essais et recueils d'articles) de critique littéraire française publiés entre la fin du XVIIIe siècle et le début du XXe siècle.
 - Ce corpus transéculaire a l'avantage d'observer une cohérence formelle (nature des documents homogène, orthographe modernisée) tout en couvrant une variété d'auteurs (239) et une période historique au cours de laquelle la critique littéraire débat sur ses conditions d'existence.
- Granularité :
 - Plutôt que d'annoter l'ensemble du discours de critique littéraire, qui contient sa part de digressions, de définitions, de descriptions méthodologiques, nous centrons notre analyse sur les phrases porteuses d'une énonciation subjective, de l'ordre de l'évaluatif affectif ou axiologique. Les énoncés considérés comme neutres (observations, définitions, anecdotes, etc.) sont écartés.

Ressources linguistiques



La constitution empirique des ressources lexicales s'appuie sur:

- une base linguistique générale déjà constituée [Motasem Alrahabi, 2010]
 - une fouille du corpus critique par le biais de l'analyseur ([Alix](#)) et de repérages manuels (enrichissement par des adjectifs et adverbes)
- Les autorités / les sources du jugement
 - substantifs : l'université, la presse, l'opinion, le public...
 - entités nommées (personnes et organisations) : Sainte-Beuve, Revue des Deux Mondes...
 - Les objets du jugement
 - substantifs : récit, nouvelle, pièce, poète, style, génie...
 - entités nommées (personnes et organisations) : Balzac, Comédie française...
 - L'expression du jugement
 - verbes : aimer, louer, rejeter, insulter...
 - adjectifs : beau, intelligent, cacophonique, ridicule...
 - adverbes : passionnément, détestablement, sottement...

Approche pour l'annotation des textes



- Analyse de surface (outil à base de règles)
- Stratégie d'annotation:
 - Un premier repérage des expressions des modalités (approbation, désapprobation, éloge, polémique, adhésion, répulsion, indignation...).
 - Puis, autour de ces modalités, repérage des adjectifs et adverbes modaux qui sont, le plus souvent, les termes qui obéissent à un champ de valeur [Gérard, 2011]
 - Les modalités appréciative et dépréciative, en particulier, sont ainsi déclinées selon qu'on y attache une valeur esthétique (« beau » / « laid »), éthique (« moral » / « immoral »), informative (« nouveau » / « ancien ») ou psychoaffective (« bouleversant » / « rebutant »).
 - La valeur référentielle (« vrai » / « faux ») proposée par Ch. Gérard est incluse dans la modalité de jugement propre à la notion de vérité (« correct » / « incorrect »).
 - Enfin, dans ces phrases modalisées, une identification des autorités et des objets du jugement critique, réalisée à la fois par une analyse des substantifs et des entités nommées.

La plateforme Ariane



- <http://obvil.huma-num.fr/ariane>
- Plateforme de fouille sémantique, permettant l'exploration des corpus de l'OBVIL préalablement annotés selon l'ontologie, en croisant des recherches par:
 - métadonnées ;
 - mots-clés (au sein des titres, des textes, des phrases annotées) ;
 - catégories sémantiques (= modalités) ;
 - entités nommées (travail en cours) ;
 - classes d'objets
- Guide de la plateforme : <https://frama.link/Ariane>

La plateforme Ariane: présentation

The screenshot displays the Ariane platform interface. At the top, the logo 'Lectura' is visible on the left, and the title 'Goethe et Diderot' is centered. On the right, the author 'Barbey d'Aurevilly, Jules' and the date '1842 (1841)' are shown. Below the title, there is a section for 'Lecture synthétique' with '48 catégories' listed in a sidebar. The main content area shows a text excerpt from Barbey d'Aurevilly's 'Le roman expérimental', discussing Goethe's reputation and the French perception of him. The text is annotated with various tags and icons, indicating the platform's analytical capabilities.

Lectura
Goethe et Diderot
Barbey d'Aurevilly, Jules
1842 (1841)

Lecture synthétique
48 catégories

Appréhension
Dépense
Compréhension
Inférence
Mots
Observation
Appréhension Précoce
Accès
Évaluation
Lect
Plan
Cohés
Structuration
Appréhension Critique

Goethe, qui n'est pas seulement une gloire allemande, mais une gloire de l'esprit humain (un fervent de Barbey d'Aurevilly avait coupé, il ne sait malheureusement plus dans quel journal, cette curieuse note : « Au mois d'octobre 1802, Goethe écrit à Schiller : « Je vous attends à Westphal : nous discuterons littérature et « nous parlerons de la France. J'ai rêvé qu'un jour cette « nation, qui a tant aidé à ma réputation, briser le piédestal qu'elle m'a élevé. Vous m'expliquez « ce songe. « Mais jusqu'à présent, cette lettre a été cherchée en vain dans la Correspondance de Goethe. (Note de l'éd.). La France, seule, a fait plus pour cette gloire que l'Angleterre, par exemple, dont le sang saxon touché au sang german dans ses veines, et qu'onque son Walter Scott en traduit Goethe et que son grand Byron lui ait dédié respectueusement son Manfred. La France a fait plus que l'Allemagne elle-même. Ouf ! la France, la séduisante France, qui s'prend de toute chose et de toute pensée étrangère, a eu également la gloire de Goethe. Sans elle, il serait encore dans le fossé de l'Allemagne. Mais elle a mis au service de sa gloire sa langue, qu'on entend et qu'on parle partout. Chose inouïe ! Napoléon lui-même, ce grand choqueur d'opinion publique, qui convenait, à Santa-Hélène, d'avoir été renversé pour l'avoir choquée, — le même Napoléon qui devait trouver que le livre de madame de Staël sur l'Allemagne n'était pas assez français et qui le frappa de l'interdit de sa police, — eut sa minute de madame de Staël. « Vous êtes un homme, vous ! « Oh ! il n'y a un jour à Goethe. Qu'en savait-il ? L'arrêt il su ? publie Goethe n'a jamais agi !

« Michel sante ce mot qui sante Goethe, et qui met quinze barbes, sous forme iconique et brisée, de l'ennemi de madame de Staël. Ce jour-là tout aussi femme gloire !

Il était, comme elle, victime d'une gloire à laquelle il n'avait lui-même en y croyant :

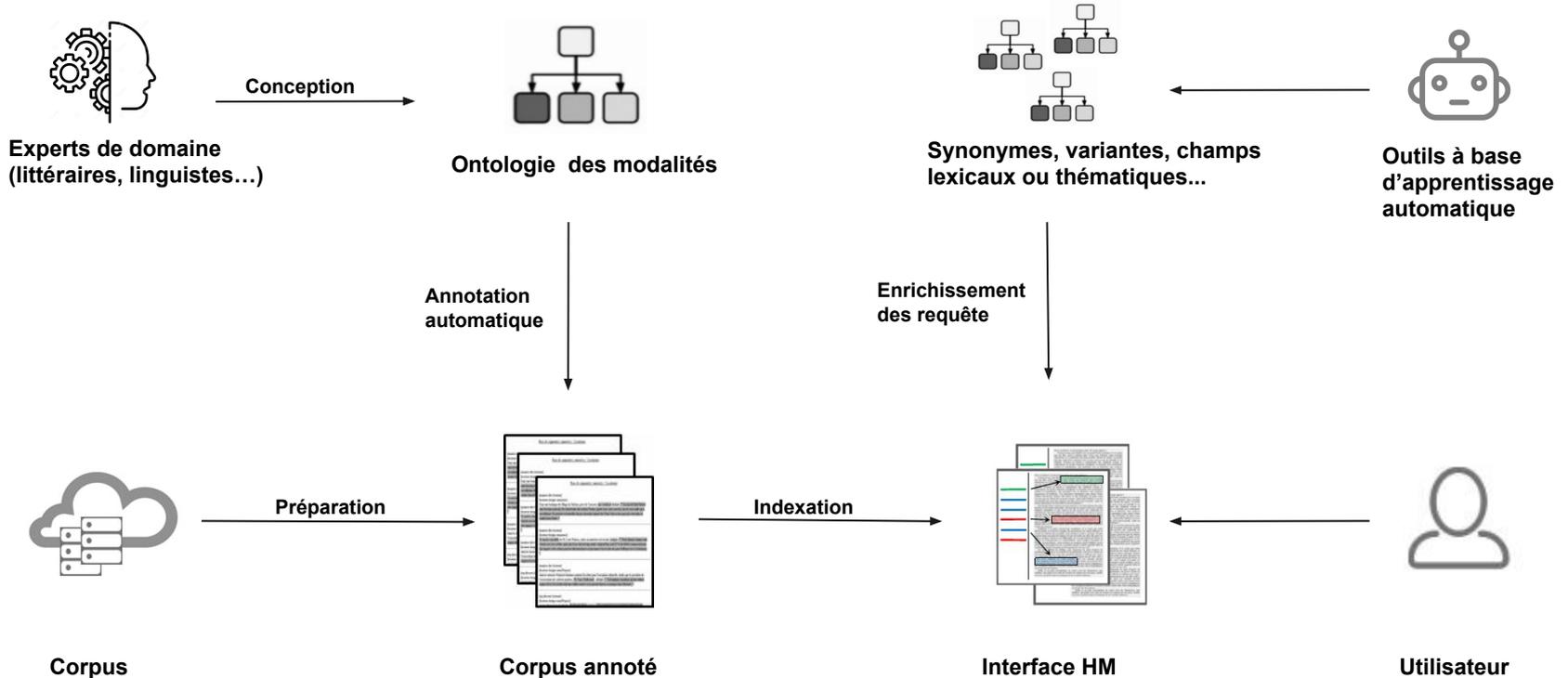
« Et pourquoi y croyait-il ? Pourquoi la France y croyait-elle ?

« Qu'y avait-il de plus corinaire à l'esprit surajouté et à la netteté transcendante de l'esprit de Napoléon que cette grande redieuse de l'esprit de Goethe, qui, du reste, n'embarassait pas beaucoup ses adorateurs jusqu'à la comparer à la Voie Lactée et en fait un tourbillonnement d'indiscernables étoiles ?

« Et qu'y a-t-il de plus antipathique au génie clair, ouvert, rapide et absolu de la France, que ce qu'on appelle — et peut-être pour l'entendre ! — le génie de Goethe ? Je sais bien que l'esprit français, l'esprit de la rue, s'était lassé entendre bien avant l'avènement de Goethe par l'idée protestante et philosophique dont l'Allemagne moderne et Goethe sont sortis. Il Je sais bien que « Diderot », le précurseur « Diderot », avait bien présenté la gloire de Goethe avec la sienne :

« mais Voltaire, le seul homme du XVIII^e siècle chez qui l'ambicille philosophie n'avait pas enroulé l'esprit, resté français,

La plateforme Ariane: architecture générale



Perspectives



Cette ontologie s'avère être une méthode exploratoire pertinente pour mener des études sémantiques transversales, qui restent proches des textes. Elle a servi cette année à conduire plusieurs travaux littéraires, comme l'étude de la réception des femmes écrivaines par la critique littéraire du XIXe siècle (Riguet et Alrahabi, 2020) et l'analyse des polémiques dans les correspondances (Alrahabi et al., 2020).

Nous entendons poursuivre son développement par :

- l'enrichissement des ressources linguistiques (adjectifs, adverbes et substantifs).
- la prise en compte de l'intensité des modalités.
- l'ajout de la recherche par entités nommées sur la plateforme.
- l'évaluation des résultats à l'aide d'un Gold Standard et la comparaison avec d'autres approches.

Bibliographie



Adam, Jean-Michel, *La Linguistique textuelle. Introduction à l'analyse textuelle des discours*, Paris, Armand Colin, 2005.

Alrahabi, Motasem, *EXCOM-2 : plateforme d'annotation automatique de catégories sémantiques. Applications à la catégorisation des citations en français et en arabe*, thèse de doctorat en Traitement Automatique de Langue dirigée par J.-P. Desclés, Université Paris-Sorbonne, 2010.

Alrahabi, Motasem, Flepp, Pauline, Koskas, Camille, « Polémiques dans le rituel épistolaire : les cas de la correspondance Ponge et Paulhan », *Revue Épistolaire*, 2020. (article soumis).

Canvat, Karl, et Legros, Georges, *Les valeurs dans/de la littérature*, Namur, Presses Universitaires, « Diptyque », 2004.

Chartier Jean-François, Slimane Malika Ould, et Djoua Brahim, *Repérage et classification automatique des énoncés à modalité axiologique dans les textes*, *JADT 2012*, présenté à Actes des 11^e Journées internationales d'Analyse statistique des Données Textuelles, Université de Liège, 2012.

Chollier, Christine (dir.), *Qu'est-ce qui fait la valeur des textes ?*, Reims, EPURE, 2011.

Courtès, Joseph, *Analyse sémiotique du discours de l'énoncé à l'énonciation*, Paris, Hachette, 1991.

Desclés, Jean-Pierre, « Prise en charge, engagement et désengagement », *Langue française* 2, 2009, n° 162, p. 29-53.

Bibliographie



Desclés Jean-Pierre, et Le Priol Florence (dir.), *Annotations automatiques et recherche d'informations*, Paris, Hermès, 2009.

Dominicy, Marc, et Frédéric, Madeleine (dir.), *La Mise en scène des valeurs : la rhétorique de l'éloge et du blâme*, Lausanne, Delachaux et Niestlé, 2001.

Gruber, Thomas, *Ontolingua : A mechanism to support portable ontologies*, KSL Report, KSL-91-66, Stanford University, 1991.

Guarino, Nicola, (ed.), *Formal Ontology in Information Systems. Proceedings of FOIS'98*, Trento, Italy, 6-8 June 1998.

Kerbrat-Orecchioni, Catherine, *L'Énonciation. De la subjectivité dans le langage*, Paris, Armand Colin, 1980.

Laurendeau, Paul, « Modalité, opération de modalisation et mode médiatif », Delamotte-Legrand. R. dir. *Les médiations langagières*, Volume 1, *Des faits de langue aux discours*, Dyalang CNRS, Publications de l'Université de Rouen, Rouen, pp 83-95.

Mangueneau, Dominique, « L'énonciation et analyse du discours », Corela [En ligne], HS-19, 2016.

Noy, Nathalia F., et Mcguinness, Deborah L., "Ontology Development 101 : A Guide to Creating Your First Ontology". 2001.

Riguet, Marine et Alrahabi, Motasem, "Analyse automatique pour une étude du genre : quels jugements des écrivaines au XIX^e siècle ?", *Digital Humanities Conference*, Ottawa, 2020.

Smith, Barry, et Welty, Christopher (dir.), *Formal Ontology in Information Systems*, ACM Press, 2001.

Voisin, Patrick (dir.), *La Valeur de l'œuvre littéraire, entre pôle artistique et pôle esthétique*, Paris, Classiques Garnier, 2012.