



**HAL**  
open science

## **ETHCOMOD : une nouvelle méthode de modélisation d'ontologie métier**

Charlotte Darricades, Christian Sallaberry, Sébastien Laborie, Eric Kergosien,  
Patrice de La Broise

### ► To cite this version:

Charlotte Darricades, Christian Sallaberry, Sébastien Laborie, Eric Kergosien, Patrice de La Broise. ETHCOMOD : une nouvelle méthode de modélisation d'ontologie métier. Congrès INFORSID 2026, Université Toulouse Capitole; Institut de Recherche en Informatique à Toulouse, May 2026, Toulouse, France. ⟨hal-05610842⟩

**HAL Id: hal-05610842**

**<https://hal.science/hal-05610842v1>**

Submitted on 4 May 2026

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 CC BY-ND 4.0 - Attribution - No Derivative Works - International License

---

# ETHCOMOD : une nouvelle méthode de modélisation d'ontologie métier

## Application dans le secteur de l'énergie

**Charlotte Darricades<sup>1,2,3</sup>, Christian Sallaberry<sup>2</sup>, Sébastien Laborie<sup>2</sup>, Eric Kergosien<sup>1</sup>, Patrice De La Broise<sup>1</sup>**

1. Université de Lille, Gériico

charlotte.darricades@univ-lille.fr, eric.kergosien@univ-lille.fr, patrice.de-la-broise@univ-lille.fr

2. Université de Pau et des Pays de l'Adour, E2S UPPA, LIUPPA

christian.sallaberry@univ-pau.fr, sebastien.laborie@univ-pau.fr

3. Pôle d'Études et de Recherche de Lacq, TotalEnergies SE, BP 47, 64170 Lacq

charlotte.darricades@totalenergies.com

---

*RESUME. La méthode ETHCOMOD étend les approches existantes en intégrant étroitement chercheurs, communicants et ressources sémantiques externes. Elle repose sur une démarche itérative combinant enquêtes de terrain, analyses de corpus scientifiques et traitements semi-automatisés du langage. Elle a été expérimentée en entreprise, dans un contexte de R&D multi-énergies. Les résultats montrent que l'ontologie produite facilite la désambiguïsation terminologique et la construction d'un langage commun entre experts et non-experts. Le prototype CommTools, développé comme preuve de concept, met en évidence le rôle médiateur de l'ontologie dans la génération de contenus vulgarisés fiables.*

*ABSTRACT. The ETHCOMOD method extends existing ontology engineering approaches by tightly integrating researchers, communication professionals, and external semantic resources. It relies on an iterative process combining field studies, scientific corpus analysis, and semi-automated natural language processing. The method was validated in an industrial multi-energy R&D context. The experiment shows that the resulting ontology supports terminological disambiguation and enables the construction of a shared language between experts and non-experts. The CommTools proof-of-concept highlights the mediating role of ontology-based knowledge structures in generating reliable popularized scientific content.*

*Mots-clés : Modélisation d'ontologie, Représentation de connaissances, Partage de connaissances, Communication*

*KEYWORDS: Ontology engineering, Knowledge representation, Knowledge sharing, Communication*

---

## 1. Introduction

Dans un environnement marqué par l'accélération des transitions énergétiques, numériques et sociétales, le secteur de la Recherche et Développement (R&D) ne se limite plus à inventer et développer des solutions techniques (Lanciano-Morandat et al., 2019 ; Darbellay, 2012) : il contribue activement à la construction des récits organisationnels, à l'identité des entreprises et à leur inscription dans un horizon collectif de transformation. Autrement dit, la transformation de la R&D en opérateur de récits organisationnels (Giroux & Marroquin, 2005) et de légitimation sociale ne peut se réaliser qu'à travers des dispositifs de communication interne et externe capables d'articuler la production de savoirs avec leur mise en circulation. La question de la communication devient alors centrale : comment ces savoirs et ces récits produits par la R&D circulent-ils dans l'organisation et au-delà de ses frontières ?

Dans ce prolongement, les travaux d'El Mezouari, Lotfi et Bouthir (2013) soulignent que la compétitivité des entreprises industrielles dans un contexte mondialisé ne peut être dissociée d'un système de communication interne et externe efficace. Les chercheurs, traditionnellement perçus comme des producteurs de connaissances techniques, sont aujourd'hui de plus en plus sollicités pour devenir des acteurs de communication. Ils ne se limitent plus à publier des résultats scientifiques dans des revues spécialisées : ils participent à des événements internes, coconstruisent des supports de vulgarisation, interviennent dans les médias et contribuent ainsi à l'humanisation et à la mise en visibilité de l'innovation. Face à eux, les communicants occupent une place spécifique mais complémentaire. Leur rôle ne se réduit pas à « traduire » les savoirs en langage accessible : ils cadrent les messages, organisent les dispositifs éditoriaux et participent à la scénarisation stratégique des projets. Ils opèrent une médiation constante entre les exigences de lisibilité et de visibilité d'une part, et les impératifs de précision et de crédibilité scientifique d'autre part. La R&D apparaît donc non seulement comme un espace de production scientifique, mais aussi comme un nœud communicationnel structurant les interactions entre chercheurs, communicants et directions.

Ainsi, la transmission et la vulgarisation de connaissances scientifiques constituent des enjeux stratégiques majeurs. Pourtant, au sein de grandes organisations en pleine mutation face au défi de la transition énergétique, comme, par exemple, TotalEnergies, une rupture sémantique persiste souvent entre les chercheurs, qui produisent un savoir hautement spécialisé, et les communicants, chargés de valoriser ces travaux auprès de publics non experts. Si de nombreuses méthodologies de structuration de la connaissance existent, telles que Methontology (Fernandez-Lopez et al., 1997), NeOn (Suàrez-Figueroa et al., 2012) ou SAMOD (Peroni, 2016), elles se concentrent principalement sur des aspects informatiques ou de gestion, délaissant souvent la dimension médiatrice de la communication. L'originalité de notre approche se distingue par une modélisation qui consiste plutôt à engager tous les acteurs qui participent à la gestion des connaissances d'une organisation, et cela tout au long du cycle de vie, c'est-à-dire de la création de ces connaissances jusqu'à leur valorisation.

Cet article présente ETHCOMOD, une nouvelle méthode de modélisation d'ontologie métier conçue pour mieux intégrer les experts dans la boucle de l'identification, de la structuration et du partage des connaissances au sein d'une organisation. En s'appuyant sur une approche itérative et collaborative, ETHCOMOD vise à transformer les productions scientifiques en concepts sémantiques structurés pouvant être, par la suite, exploités à des fins de médiation, entre chercheurs et communicants, par exemple. Ainsi, nous détaillerons d'abord, en section 3, les cinq étapes de cette méthode, avant d'exposer un exemple de mise en œuvre concrète au sein du Pôle d'Études et de Recherche de Lacq (PERL) de TotalEnergies, en section 4. Enfin, dans la section 5, nous présenterons et expérimenterons le prototype CommTools qui exploite cette base de connaissance au service des communicants de TotalEnergies, pour un meilleur accès et une valorisation des productions scientifiques des chercheurs.

## 2. État de l'art

Un nombre important de travaux propose une méthodologie pour construire une ontologie. Parmi ceux-ci, nous pouvons notamment citer les méthodes Toronto Virtual Enterprise (TOVE) (Grüniger et Fox, 1995), Methontology (Fernandez-Lopez et al., 1997), Sensus (Swartout et al., 1997), OnTo-Knowledge (OTK) (Staab and al., 2001), Terminae (Aussenac-Gilles et al., 2008), Networked Ontologies (NeOn) (Suàrez-Figueroa et al., 2012), ou encore Simplified Agile Methodology for Ontology Development (SAMOD) (Peroni, 2016).

Toutes ces méthodes commencent par une phase d'acquisition de connaissances du domaine ou du métier, de rédaction de spécifications fonctionnelles ou de questions de compétence. Methontology et OTK sont très similaires. Les deux approches commencent par l'acquisition de connaissances et la rédaction de spécifications. Elles se poursuivent en modélisant le domaine d'abord d'une manière informelle puis dans un langage formel. Enfin, toutes deux proposent une évaluation de l'ontologie produite. NeOn et SAMOD édictent des consignes pour créer des ontologies modulaires suivant une approche par composition (Despres, 2014). Certaines méthodologies, notamment OTK, NeOn et SAMOD proposent un développement modulaire de l'ontologie construite pas-à-pas, soit en ajoutant, à chaque itération, la modélisation d'une partie supplémentaire du métier/domaine, soit en modélisant toutes les parties du métier/domaine d'abord, puis en les fusionnant. NeOn se distingue des autres méthodologies présentées car elle fournit de nombreuses approches pour élaborer une ontologie ou un réseau ontologique. Elle demande aux ontologues de réaliser préalablement une analyse approfondie du projet afin de pouvoir choisir la bonne combinaison des processus et activités proposés. Les trois méthodes OTK, NeOn et SAMOD intègrent également une phase d'évaluation de l'ontologie produite, et cela lors de l'étape finale. SAMOD semble ressortir du lot selon nos critères car elle propose une première étape permettant de créer un premier squelette d'ontologie à partir des besoins exprimés par les utilisateurs cibles. La méthode préconise d'impliquer fortement les experts concernés afin de préciser et d'étendre les besoins, et le modèle ontologie produit,

de façon itérative (Rawsthorne et al., 2022). L'aspect itératif impliquant les experts est primordial dans un secteur spécifique tel que celui de la R&D. Enfin, SAMOD intègre des phases de tests à différentes étapes du processus.

Le tableau suivant offre une synthèse de ces approches en mettant en avant leurs forces ainsi que leurs limites.

*Tableau 1. Comparaison des principales approches de création d'ontologies*

<b>Approches</b>	<b>Objectifs</b>	<b>Forces</b>	<b>Limites</b>
TOVE (1995)	Formaliser les connaissances organisationnelles dans les entreprises industrielles	Une approche rigoureuse et très structurée avec un ancrage fort dans la modélisation des processus d'entreprise	Complexe à exploiter car nécessite une expertise en logique formelle (peu adaptée à la collaboration non technique)
Methontology (1997)	Construire une ontologie complète à partir de la connaissance experte d'un domaine	Une méthode détaillée et systématique avec un guide d'évaluation clair	Interactions faibles entre experts et concepteurs
Sensus (1997)	Étendre et relier des ontologies existantes à partir de ressources linguistiques et sémantiques	Favorise la réutilisation de connaissances et l'interopérabilité entre ontologies	Aucune démarche participative et dépendance forte à des connaissances préexistantes
OnTo-Knowledge (2001)	Développer des ontologies pour améliorer la gestion des connaissances dans les entreprises	Une approche orientée entreprise qui prend en compte la maintenance et l'évaluation continue	Nécessite des ressources informatiques et humaines conséquentes ; ce qui accroît sa complexité
Terminae (2008)	Construire une ontologie à partir de corpus textuels et d'analyses linguistiques	Un fort ancrage sémantique qui est utile pour modéliser un vocabulaire métier	Nécessite des outils linguistiques et un corpus bien délimité (peu d'interactions humaines)
NeOn (2012)	Favoriser la création collaborative et modulaire d'ontologies ou de réseaux d'ontologies	Une approche modulaire et flexible qui permet la co-construction entre plusieurs acteurs	Nécessite un haut niveau d'expertise et une planification lourde
SAMOD (2016)	Concevoir une ontologie de manière agile et itérative en collaboration avec les experts du domaine	Une méthode itérative, participative, flexible et centrée sur les besoins réels des utilisateurs	Demande une implication constante des experts du domaine

L'analyse comparative des méthodologies existantes souligne ainsi la diversité des approches possibles, chacune répondant à des finalités spécifiques techniques, organisationnelles ou linguistiques, mais rarement à des enjeux communicationnels. Si ces méthodes offrent des cadres rigoureux pour la formalisation et la structuration des connaissances, elles demeurent majoritairement ancrées dans une logique informatique ou de gestion de projet, laissant en marge les dimensions symboliques, relationnelles et médiationnelles propres aux Sciences de l'Information et de la Communication. C'est pourquoi dans la section suivante, nous présentons notre approche ETHCOMOD pour répondre à ce besoin.

### 3. La méthode ETHCOMOD

ETHCOMOD (*cf.* Figure 1) étend la méthode SAMOD (Peroni, 2016), illustrée plus haut, de la façon suivante : (i) ajout d'une première étape qui vise la cartographie de l'organisation en général et du métier ciblé en particulier ; (ii) intégration de différentes communautés dans la collecte et la validation des données ; (iii) préconisation de l'usage d'outils de TALN pour l'extraction d'information (labels et concepts) ; (iv) ajout d'une étape d'enrichissement des connaissances à partir de ressources externes ; (v) intégration des *modlets* (i.e., sous graphes ontologiques, agrégés en fin de processus) pour constituer l'ontologie finale qui sera expérimentée par les communautés du domaine dans leurs activités. Un modlet (Peroni, 2016) permet de modéliser une partie (domaine/sous-domaine) du métier visé sous la forme d'un sous-graphe ontologique.

ETHCOMOD (représentation simplifiée, *cf.* Figure 1) se décline en 5 étapes que nous allons présenter succinctement dans les sections 3.1 à 3.5. Par la suite, un exemple dans le secteur de l'énergie illustrera chaque étape en section 4.

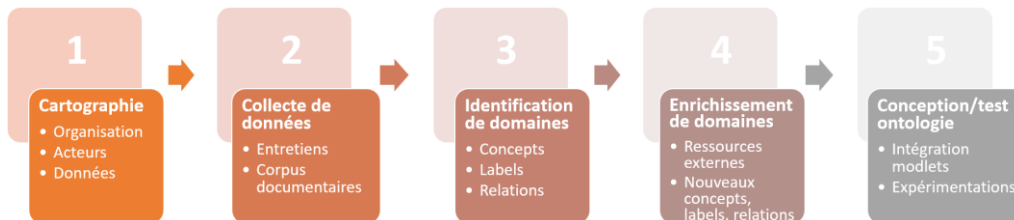


Figure 1 : Les 5 étapes de la méthode ETHCOMOD

#### 3.1. Cartographie de l'organisation

Cette étape consiste à créer une cartographie qui fait état des caractéristiques de l'organisation : directions, acteurs, groupes socio-professionnels, connaissances, informations, données. Plusieurs communautés peuvent émerger (e.g., ingénieurs et chercheurs qui publient des articles scientifiques relatifs à leurs travaux, communicants qui publient également selon différents canaux en interne et en externe pour valoriser des choix/orientations de l'organisation). Les ressources de données créées et/ou gérées par ces communautés sont également répertoriées.

### **3.2. Collecte de données du domaine (entretiens et corpus textuels)**

Il s'agit ici de récolter des données relatives au domaine ciblé. Les approches ethnographiques préconisées (enquêtes, entretiens semi-directifs et observations), menées auprès des communautés en interaction au cœur du métier à modéliser (les vocabulaires peuvent être très différents allant de celui de l'expert jusqu'à celui du néophyte), visent l'identification de corpus documentaires (confidentiels vs non-confidentiels) traitant du métier.

### **3.3. Identification de domaines**

Cette étape est consacrée à l'analyse des transcriptions d'interviews et des documents des corpus pour déterminer des sous-domaines et les concepts ainsi que les relations qui les caractérisent. Le regroupement de concepts et des labels associés permet ensuite l'organisation de ces concepts via des liens hiérarchiques. Le traitement de documents semi-automatisé est supporté par des outils de traitement automatique du langage. La validation/correction itérative des sous-domaines représentés par des *modlets* implique des représentants des différentes communautés d'acteurs concernées.

### **3.4. Enrichissement de domaines**

ETHCOMOD fait une hypothèse forte que des ressources externes expertes (bases de connaissances en lien avec un sous-domaine) peuvent enrichir efficacement des *modlets* en termes de concepts, labels et relations. Ici, comme précédemment, des traitements semi-automatisés d'interrogation et de parcours de bases de connaissances sont préconisés.

### **3.5. Conception et tests de l'ontologie (intégration des modlets, expérimentations)**

Les *modlets* sont intégrés et l'ontologie métier est expérimentée/utilisée par les communautés du domaine dans leurs activités. La validation/tests auprès d'experts est réalisée à l'aide d'un éditeur d'ontologie. Des applications informatiques exploitant l'ontologie sont également développées et mises à disposition des utilisateurs.

## **4. Application de la méthode ETHCOMOD aux métiers de l'énergie**

Nous avons choisi le contexte des métiers de l'énergie et plus particulièrement la collecte, la structuration et la vulgarisation des connaissances au sein du PERL (Pôle d'Etudes et de Recherche de Lacq), une entité du groupe TotalEnergies. Il s'agit, notamment, de permettre à deux communautés distinctes de l'organisation de partager des connaissances et des données relatives au domaine de l'énergie, dans leurs activités au service de l'organisation. Ainsi, une communauté d'ingénieurs et de chercheurs mène des activités de R&D dont elle fait la promotion par des

publications scientifiques auprès de pairs dans des conférences et revues nationales et internationales. Elle peut avoir des difficultés à vulgariser ces activités en interne auprès de non spécialistes lors de présentations ou autres comptes-rendus. De même, une communauté de communicants mène des activités de promotion des activités du groupe en interne et en externe. Elle s'appuie notamment sur les résultats scientifiques obtenus par ces ingénieurs et chercheurs. Elle rencontre des difficultés à interroger des documents scientifiques dont le vocabulaire spécialisé peut être difficile d'accès et éloigné du vocabulaire du public ciblé.

Cette expérimentation de la méthode ETHCOMOD a un double objectif : traduire les discours scientifiques en concepts sémantiques structurés, afin d'en permettre la réutilisation et la valorisation par les communicants. Ainsi, il s'agit de créer un langage commun aux chercheurs et communicants, permettant de relier les terminologies expertes à des catégories plus génériques, compréhensibles par le grand public. Notons que, en interne, TotalEnergies dispose d'un thesaurus géologique complet, utilisé pour les métiers du sous-sol, mais aucun référentiel spécifique aux énergies renouvelables ou à la R&D multi-énergies n'existe. De même, en externe, plusieurs ontologies ou thesaurus (VoInrae, UNBIS, PubChem), détaillés dans (Darricades, 2026) couvrent certains domaines (chimie, biologie, environnement, respectivement), mais aucune ne relie ces champs dans une structure interdisciplinaire.

C'est dans ce contexte que nous avons mis en application chaque étape de la méthode ETHCOMOD en collaboration avec les chercheurs et communicants du PERL, chez TotalEnergies.

#### ***4.1. Cartographie de l'organisation***

Une étude itérative, reliant observations, entretiens, analyses documentaires et interactions de terrain a permis de tracer les bases de la cartographie présentée sur une structure visuelle (voir Darricades, 2026). Les premières investigations ont permis de dégager les lignes hiérarchiques et les pôles décisionnels directement impliqués dans la R&D. En parallèle, une analyse des groupes socio-professionnels a permis de distinguer les chercheurs, les chercheurs-communicants, les communicants et les communicants-technophiles. Les circuits de communication formels et informels, les collaborations inter-équipes et les points de passage entre chercheurs et communicants ont pu être affinés. Un second travail a consisté à recenser, trier et classer l'ensemble des documents produits au sein du centre, en distinguant deux grands ensembles : d'une part les documents confidentiels (cahiers de laboratoire, notes de synthèse, brevets et demandes de brevets, rapport annuels, notes mensuelles de faits marquants de la R&D, les rapports de projets de R&D) et d'autre part, les documents non confidentiels, à savoir les publications scientifiques et communications validées pour diffusion dans un outil interne nommé ROPA (Relation avec les Organismes Professionnels Amont) dédié à la capitalisation des productions scientifiques.

Cette cartographie constitue le premier livrable de la méthode ETHCOMOD, matérialisant les connexions entre directions, individus et groupes professionnels et pointant les ressources informationnelles de chacun.

#### **4.2. Collecte de données du domaine (entretiens et corpus textuels)**

La deuxième étape est consacrée à la collecte et à la structuration des données de la R&D et des communicants. Pour notre expérimentation, cette phase s'est concentrée sur la constitution du corpus des publications scientifiques, seules sources publiques, validées institutionnellement et diffusables. Au total, 56 publications produites par les chercheurs du PERL ont été sélectionnées dans le cadre du travail de recensement (sélection représentant les différents domaines de recherche du PERL) via l'outil interne ROPA, avant d'être analysées, classées par thématique et sélectionnées pour la modélisation sémantique. En parallèle, treize entretiens semi-directifs (via grille d'entretien d'une durée moyenne de 1h30) ont été réalisés auprès de sept chercheurs et six communicants du PERL, complétés par des réunions internes et un lien direct avec le terrain à travers quatre stages d'observation. Ces immersions relèvent d'une observation participante visant à saisir les logiques communicationnelles à l'œuvre dans la production et la valorisation scientifique (Latour et Biezunski, 1989). Enfin, un troisième travail intègre une dimension comparative et participative, à travers la diffusion d'un questionnaire (voir Darricades, 2026) : un premier lot, adressé aux communicants R&D internes à TotalEnergies (12 répondants, 9 basés en France, et 3 en Belgique, en Inde et au Brésil) et un second, destiné à des communicants R&D externes à TotalEnergies dans des domaines d'activités stratégiques différenciés (5 répondants français issues des domaines de la construction, pharmaceutique, agroalimentaire et fabrication pneumatiques).

#### **4.3. Identification de modlets (processus par sous-domaine, itérations)**

Il s'agit désormais de soumettre les données scientifiques collectées au processus d'extraction terminologique et conceptuel d'ETHCOMOD. Les retranscriptions d'entretiens ont été analysées manuellement par des experts communicants R&D de TotalEnergies. Les 56 publications répertoriées ont, quant à elles, été analysées via une chaîne de traitement semi-automatisée. Deux choix fondamentaux ont été opérés : l'utilisation de TextRazor (Hollink et al., 2016) comme outil d'extraction afin d'identifier automatiquement les entités nommées et l'adoption de Wikidata (Vrandečić, 2012) comme référentiel sémantique garantissant la cohérence, la traçabilité et la hiérarchisation des concepts scientifiques (via les identifiants Wikidata QID). Ces choix, parmi de nombreux outils d'extraction d'information et ressources externes, ont été faits à l'issue d'expérimentations sur le corpus (Darricades, 2026).

Ainsi, dans le cadre de cette étude, l'extraction d'information et l'identification de *modlets* à l'étape 3 d'ETHCOMOD se décline en 4 phases :

- 1) Extraction de termes à forte densité sémantique dans une publication avec TextRazor. L'analyse porte sur les *titre, résumé, introduction et mots-clés*

d'une publication. Ainsi, TextRazor détermine le contexte sémantique approprié dans Wikidata. Ici, le terme *adsorption* est identifié dans une des parties ciblées du texte relatif à une publication du corpus (cf. Figure 2) ;



Figure 2 : Extraction de termes

- 2) Étiquetage conceptuel de termes avec Wikidata. Le terme *adsorption* est rattaché à un concept Wikidata dont le QID est Q180254 (cf. Figure 3) ;

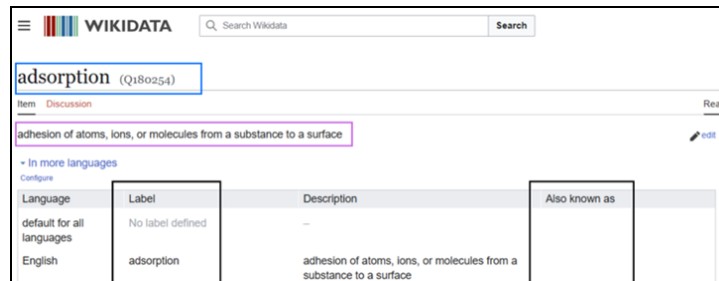


Figure 3 : Étiquetage conceptuel

- 3) Identification et association de concepts dans des *modlets*. Le terme *adsorption* va devenir à la fois label et concept d'une première version de *modlet* (cf. Figure 4, le rond orange représente un concept tandis que les ronds clairs représentent les labels associés à ce concept) ;

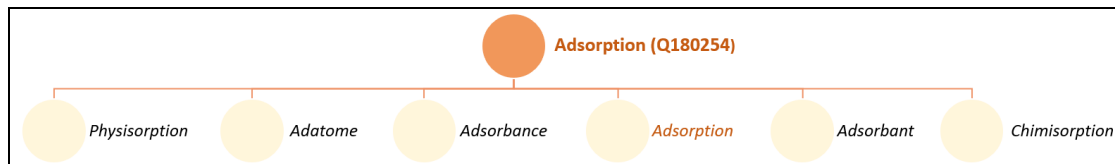


Figure 4 : Identification et association de concepts

- 4) Validation/correction des *modlets* avec les chercheurs. Ce *modlet*, construit à partir d'une publication, est validé ou corrigé par les chercheurs du PERL.

#### 4.4. Enrichissement de *modlets* (tests, intégration de concepts, labels, relations, issus de ressources externes)

Les *modlets* construits à l'étape précédente sont désormais enrichis à l'aide de la ressource externe Wikidata. Des labels sont rajoutés (dans différentes langues) ainsi que des concepts (cf. Figure 5).

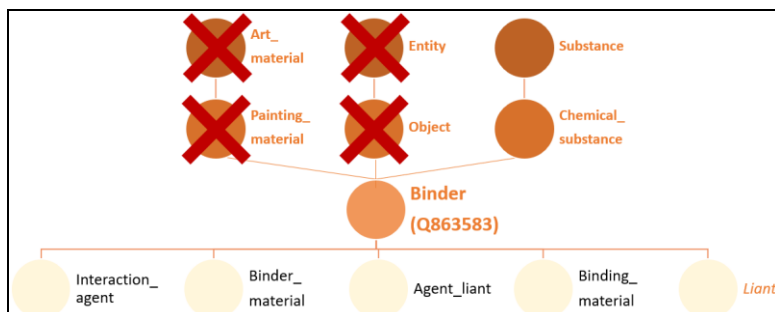


Figure 5 : Enrichissement de modlets

Dans ce second exemple, il faut noter que le terme *Liant* a été extrait d'une publication différente de celle qui a mis en exergue *Adsorption*. Ici, conformément au contexte, *Liant* est devenu un label rattaché au concept *Binder*. Les labels de couleur orange (ronds clairs) sont des termes chercheurs issus de publications scientifiques du PERL. Les labels en noir (ronds clairs) sont proposés par Wikidata. Les six concepts parents de *Binder* (ronds orange) sont également rajoutés sur la base de la ressource Wikidata. Ici, les chercheurs ont validé *Chemical\_substance* et *Substance* mais ont supprimé quatre autres concepts sémantiquement éloignés du domaine.

#### 4.5. Conception et tests de l'ontologie (intégration des modlets, expérimentations)

Les *modlets* sont intégrés pour constituer l'ontologie. À titre d'exemple (cf. Figure 6), nous prenons deux *modlets* *Adsorption* et *Electrophoresis* qui, dans la hiérarchie Wikidata, ont un même concept parent *Seperation\_Process*. Ainsi, ceci nous permet de bâtir l'ontologie finale de manière itérative en assemblant les modlets deux à deux.

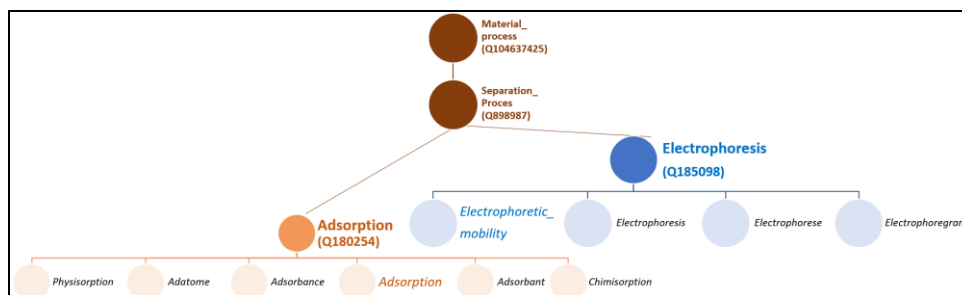


Figure 6 : Intégration de modlets

L'analyse terminologique issue du corpus scientifique du PERL met en évidence un total de 748 labels extraits des publications contenant effectivement 2335 occurrences. Le français représente 42% des labels, l'anglais domine avec 58%, conformément aux normes de production scientifique internationale et aux pratiques rédactionnelles des chercheurs du PERL. L'ontologie produite contient 350 concepts distincts (89 % des étiquettes provenant de Wikidata et 11 % des chercheurs du

PERL). La moyenne de 5,44 labels par concept constitue un indicateur déterminant : elle révèle l'existence d'une pluralité de formulations pour désigner un même objet scientifique, ce qui renforce la nécessité d'un travail de désambiguïsation. Les corrections, opérées par les cinq chercheurs mobilisés, ont permis de stabiliser les hiérarchies sémantiques et d'ajuster la base de connaissances en fonction des usages effectifs observés dans les publications du PERL.

La prépondérance des labels issus de Wikidata (89%) atteste par ailleurs de l'intégration massive de référentiels externes dans le processus de normalisation terminologique. L'ancrage dans Wikidata permet de stabiliser les identifiants conceptuels, de réduire les ambiguïtés lexicales et d'assurer l'interopérabilité des connaissances produites. À l'inverse, les labels spécifiquement issus du PERL (11%) témoignent de la présence d'un lexique local, situé, reflétant les pratiques discursives des chercheurs. Leur intégration dans l'ontologie garantit que le modèle ne se limite pas à un simple alignement sur un vocabulaire global, mais incorpore la dimension de la connaissance propre au terrain. Dans cette perspective, l'ontologie produite ne constitue pas uniquement un outil de formalisation, mais un dispositif de médiation visant à faire converger des ressources lexicales multiples vers une structure commune, mobilisable autant par les chercheurs que par les communicants.

Cette opération, bien que construite avec les experts du domaine, soulève cependant une question centrale : dans quelle mesure une base de connaissances patiemment construite, par un travail minutieux de tri, de confrontation et de validation experte, peut-elle pleinement déployer son potentiel tant qu'elle n'a pas été mise à l'épreuve d'usages réels ?

## 5. Preuve de concept

### 5.1. Le prototype *CommTools*

L'outil développé intitulé « *CommTools* » constitue la preuve de concept de notre méthode ETHCOMOD. Ce prototype met à l'épreuve la solidité du modèle en le confrontant à un usage réel dans le contexte de TotalEnergies. L'application permet de vérifier la robustesse sémantique de l'ontologie (cohérence, traçabilité, interopérabilité) tout en évaluant sa valeur médiatrice, c'est-à-dire sa capacité à favoriser le passage des savoirs scientifiques vers des formats communicationnels fiables.

L'architecture de *CommTools* a été pensée pour répondre à trois finalités principales :

- (1) **Pour les chercheurs de l'entreprise**, elle constitue un espace de valorisation et de visibilité. Les publications indexées y apparaissent associées à des annotations conceptuelles issues de l'ontologie produite par ETHCOMOD, garantissant la fidélité des termes et la cohérence terminologique ;
- (2) **Pour les communicants**, l'application représente un levier d'appropriation et de médiation. Elle leur donne accès à des contenus validés, contextualisés et non confidentiels, tout en offrant un niveau de lecture adapté à leurs

besoins de communication. Elle leur permet ainsi de construire des narratifs alignés sur la rigueur scientifique, tout en restant accessibles à des publics variés ;

- (3) **Pour l'entreprise**, *CommTools* renforce la crédibilité des discours institutionnels autour de la R&D et participe à la construction d'une image d'innovation responsable et transparente.



Figure 7 – Interface principale de l'outil *CommTools*

Cette première interface, visible dans la Figure 7, constitue le point d'entrée dans notre outil *CommTools*. Elle permet de rechercher un concept, un auteur ou un terme, grâce à un système d'*auto-complétion* connecté à l'index stocké dans une base GraphDB<sup>1</sup>. Cette base de connaissances RDF/OWL contient les annotations du corpus scientifique en lien avec l'ontologie métier construite par notre approche ETHCOMOD. Le moteur de recherche repose sur des requêtes SPARQL exécutées via une API Flask, qui interroge la base et retourne les résultats au format JSON. Les documents correspondants au résultat de la requête sont présentés sous forme de cartes enrichies (lorsque l'utilisateur sélectionne une carte celle-ci est entourée en rouge) : chaque carte affiche le titre, les auteurs, les équipes associées, et un résumé vulgarisé correspondant à des articles scientifiques du corpus répondant au critère de recherche de l'utilisateur de *CommTools*. Dans l'exemple illustré dans la Figure 7, l'utilisateur cherche des publications relatives au concept « *adsorption* » et 12 références de documents lui sont restitués. Évidemment, notre outil reconnaît automatiquement toutes les occurrences du concept « *adsorption* » dans le corpus, y compris lorsque le terme exact n'est pas utilisé (ex., si l'article emploie les expressions : *adsorbent*, *adsorption capacity* ou *surface interaction*).

Lorsque l'utilisateur souhaite disposer de plus de détails sur une publication renvoyée lors d'une recherche, une nouvelle interface permet de présenter plus

<sup>1</sup> <https://graphdb.ontotext.com>

d'informations sur l'article en lui-même comme on peut le voir dans la Figure 8. Sur la gauche on retrouve l'article en PDF, sur la droite un résumé vulgarisé et en bas un graphe interactif qui permet de naviguer au sein de l'index pour lequel on retrouve notamment les concepts, les auteurs ainsi que des relations associées. En sélectionnant par exemple le nœud « *adsorbants* » dans ce graphe, toutes les occurrences associées au concept sont mises en évidence dans le résumé ainsi que dans le PDF de l'article, les labels liés (*synonymes, termes équivalents, dérivés scientifiques*) s'affichent également, et la définition issue de Wikidata<sup>2</sup> ou de notre ontologie métier issue de TotalEnergies apparaît dans la zone d'information. Cette navigation au sein de l'index permet à l'utilisateur de visualiser graphiquement comment un terme ou un concept s'insère dans la logique scientifique de l'article (*concepts parents, sous-concepts, relations méthodologiques ou phénoménologiques*).

The screenshot displays the TotalEnergies interface for a document. On the left, there is a PDF viewer showing the article "Gas porosimetry as an efficient tool for the assessment of the shaping effect in commercial zeolites". The article title is highlighted in blue. On the right, there is a summary section titled "Les zeolites : des petites briques pour des applications géantes". The summary includes a "Résumé" (Summary) and an "Introduction" section. The summary text describes zeolites as porous materials used for gas capture and storage. The introduction explains that zeolites are used in various applications, such as in the production of carbon black. Below the summary, there is a section titled "2. Comment mesurer l'impact du binder ?" which discusses the impact of binder on the adsorption of CO<sub>2</sub> and CH<sub>4</sub>. At the bottom of the interface, there is an interactive ontology graph showing various concepts and their relationships. The graph is a network of nodes connected by lines, with nodes representing concepts like "adsorbants", "zeolites", "adsorption", etc. A sidebar on the right of the graph shows the "Dernier nœud sélectionné" (Last selected node) as "adsorbants" and lists its labels and definition.

Figure 8 – Interface qui présente les informations au sujet d'un document

<sup>2</sup> <https://www.wikidata.org>

## 5.2. Évaluation avec les chercheurs et les communicants

Une expérimentation a été conduite pour évaluer les résultats générés par les interfaces de notre application *CommTools*, et en particulier la production par l’outil de textes vulgarisés et contextualisés via différents types de prompts prédéfinis. Cette démarche d’évaluation s’est appuyée sur les travaux de Kassogué et al. (2019) qui proposent une grille analytique structurée pour l’examen critique de productions scientifiques. L’enjeu de cette évaluation ne porte donc pas sur les prompts eux-mêmes en tant qu’instructions d’intelligence artificielle, mais plutôt sur la qualité, la fidélité et la lisibilité des textes générés, considérés comme objets de médiation entre la sphère scientifique et la sphère communicationnelle. Autrement dit, il s’agissait d’observer dans quelle mesure la génération assistée par ontologie, matérialisée dans *CommTools*, favorise une traduction intelligible, rigoureuse et légitime des savoirs scientifiques, apte à être mobilisée par les communicants dans leurs actions de valorisation de la R&D.

Pour chaque publication, trois versions de résumés vulgarisés sont soumises à évaluation (voir Darricades, 2026) : la première issue d’une génération automatique sans intégration ontologique (noté prompt A) ; la deuxième produite à partir des concepts proches de premier niveau de l’ontologie et leurs labels (noté prompt B) ; et la troisième (notée prompt C) qui exploite l’ensemble des concepts généralisants de l’ontologie et de leurs labels, offrant un niveau de contextualisation supérieur. Ces trois versions de résumés vulgarisés constituent des versions expérimentales destinées à comparer la qualité des prompts et à identifier celui produisant la traduction la plus satisfaisante. Le type de prompt retenu à l’issue de cette phase d’évaluation sera ensuite intégré de manière pérenne dans les interfaces concernées (cf., Figure 7 et Figure 8), afin d’optimiser la génération automatique de textes vulgarisés et contextualisés.

Chaque participant a évalué un ensemble de productions automatiques à l’aide d’une grille d’analyse construite sur la base de mesures quantitatives et d’observations qualitatives qui porte sur la fidélité scientifique (*présence et exactitude des notions centrales, rigueur du propos*), la clarté et la vulgarisation (*lisibilité, accessibilité du vocabulaire, compréhension pour un non-spécialiste*), la cohérence terminologique (*usage stable et pertinent des concepts issus de l’ontologie*), la capacité de médiation (*facilitation du dialogue chercheurs-communicants et communicants-grand public*), ainsi que la perception identitaire (*reconnaissance de soi dans le texte pour les chercheurs, exploitabilité pour les communicants*). Des espaces de commentaires libres permettent aux participants de motiver leurs choix, d’exprimer leurs perceptions et de proposer des ajustements.

Pour les chercheurs, l’évaluation du prompt A, généré sans ontologie, met en évidence une faiblesse notable dans la restitution des notions centrales : les chercheurs le jugent « dilué », imprécis et trop généraliste. La rigueur moyenne n’atteint que 3,2/5, et six chercheurs sur sept relèvent des erreurs ou omissions majeures. Ce résultat confirme que, sans structure sémantique sous-jacente, la génération automatique peine à maintenir la cohérence interne du discours scientifique et à restituer la hiérarchie des concepts. Le prompt B, intégrant des

concepts et labels proches issus de l'ontologie, marque une nette amélioration. La rigueur monte à 4/5, la fidélité scientifique est jugée satisfaisante par la majorité des chercheurs, et les commentaires soulignent une meilleure correspondance entre le texte généré et la logique méthodologique de la publication. L'ontologie joue donc ici un rôle de cadre de stabilisation du sens : elle aligne le vocabulaire sur la logique disciplinaire des chercheurs, tout en permettant une vulgarisation plus contrôlée. Le prompt C, enrichi par tous les concepts généralisants, obtient une évaluation globalement positive (rigueur 4/5, clarté 4,1/5), mais suscite des retours plus ambivalents : s'il est jugé plus fluide et accessible, il est aussi perçu comme moins précis scientifiquement. Certains chercheurs y voient un risque de simplification. Ces résultats montrent que l'intégration ontologique renforce la confiance des chercheurs envers le dispositif, tout en soulignant que la vulgarisation doit rester encadrée pour éviter la perte de profondeur scientifique.

Du côté des communicants, l'évaluation du prompt A, bien que jugé clair (4,4/5), est perçu comme descriptif et peu utile en communication stratégique, c'est-à-dire compréhensible mais sans profondeur. Le prompt B, plus détaillé, est salué pour sa rigueur mais jugé « dense », ce qui le rend moins exploitable pour des contenus de vulgarisation. Le prompt C s'impose en revanche comme la version la plus performante : il obtient les meilleures moyennes sur l'ensemble des indicateurs, 5/5 pour l'accessibilité, 4,5/5 pour la dynamique, 5/5 pour l'attractivité, et fait l'objet d'un consensus sur son intérêt communicationnel. Les communicants soulignent que le prompt C facilite l'appropriation rapide du sujet, notamment grâce à une contextualisation fluide et une hiérarchisation claire des idées, rendues possibles par l'appui ontologique. Sur le plan de l'usage, *CommTools* est perçu comme un outil de médiation assistée : il permet aux communicants de comprendre plus aisément des sujets techniques et d'en extraire les éléments clés pour la valorisation interne ou externe. Les communicants insistent également sur la nécessité d'une supervision humaine, rappelant que la médiation ne peut être automatisée sans une validation épistémique. Ce positionnement confirme leur compréhension du dispositif non comme une substitution, mais comme un amplificateur de la compréhension et de la circulation des savoirs.

La lecture croisée des deux panels (chercheurs d'un côté et communicants de l'autre) révèle une complémentarité forte mais différenciée entre les attentes des chercheurs et celles des communicants. Les chercheurs valorisent la fidélité scientifique et la cohérence terminologique, portées par le prompt B, tandis que les communicants privilégient la clarté, la fluidité et l'adaptabilité éditoriale, incarnées dans le prompt C. Ainsi, le prompt B est considéré par les chercheurs comme le plus fiable scientifiquement, car il restitue les concepts clés sans les simplifier à l'excès, tout en maintenant la structure logique des publications. De leur côté, les communicants identifient le prompt C comme le plus exploitable, en raison de sa capacité à rendre les thématiques scientifiques intelligibles, contextualisées et communicables, tout en conservant un niveau de rigueur satisfaisant grâce à l'appui ontologique.

## 6. Conclusion

La méthode ETHCOMOD démontre qu'une ontologie peut dépasser sa fonction technique de structuration de données pour devenir un véritable dispositif de médiation. En intégrant des ressources externes comme Wikidata et en impliquant activement les communautés métiers, cette approche permet de stabiliser un lexique commun tout en respectant les spécificités terminologiques locales. L'expérimentation menée avec le prototype *CommTools* confirme l'efficacité de cette démarche. L'évaluation croisée a révélé : que (1) le recours à l'ontologie garantit la fidélité scientifique indispensable aux chercheurs, et que (2) l'enrichissement sémantique (notamment via le prompt qui intègre des concepts généralisants) offre aux communicants la clarté et l'attractivité nécessaires à la valorisation des savoirs.

En conclusion, ETHCOMOD offre un cadre rigoureux pour assurer la circulation des connaissances dans un environnement industriel. Les perspectives de ce travail résident désormais dans la mise à l'épreuve de ces bases de connaissances au travers de multiples usages variés réels et quotidiens, afin de valider pleinement leur potentiel de transformation des pratiques communicationnelles en entreprise, et ainsi améliorer la communication interne et externe des organisations.

Ces résultats ouvrent alors une réflexion plus large sur la capacité des dispositifs ontologiques à reconfigurer durablement les modalités de circulation et de partage des savoirs au sein des organisations, interrogeant notamment la mesure dans laquelle l'intégration de référentiels mondiaux comme Wikidata dans des systèmes d'information locaux peut transformer durablement les pratiques de partage des savoirs, en faisant de l'ontologie non plus un simple dépôt de données, mais un véritable agent de confiance entre experts scientifiques et communicants.

## Bibliographie

- Aussenac-Gilles N., Despres S., Szulman S. (2008). The terminae method and platform for ontology engineering from texts. Paul Buitelaar and Philipp Cimiano, editors, *Bridging the Gap between Text and Knowledge*, pp 199-223
- Darbellay, F. (2012). La circulation des savoirs. *Interdisciplinariété, concepts nomades, analogies, métaphores*. Bern: Peter Lang.
- Darricades C. (2026). *Les contributions info-communicationnelles de la Recherche-Développement (R&D) à la traduction d'une mutation industrielle : le cas de TotalEnergies*. 2026, Thèse en Sciences de l'Information et de la Communication, Université de Lille.
- Despres, S. (2014, May). Construction d'une ontologie modulaire pour l'univers de la cuisine numérique. In *IC-25èmes Journées francophones d'Ingénierie des Connaissances* (pp. 27-38).
- El Mezouari, S., Lotfi, M., & Bouthir, Y. (2013). L'importance de la communication interne dans les entreprises: Cas d'une entreprise industrielle marocaine. *Revue Economie & Kapital*, (4).
- Fernández-López, M., Gómez-Pérez, A. & Juristo Juzgado N. (1997). METHONTOLOGY: From Ontological Art Towards Ontological Engineering. In *AAAI-97 Spring Symposium Series*, 24-26 March 1997, Stanford University, EEUU.

- Giroux, N., & Marroquin, L. (2005). L'approche narrative des organisations. *Revue française de gestion*, 159(6), 15-42.
- Grüninger, M., & Fox M.S., (1995). Methodology for the design and evaluation of ontologies. *In Proc. IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing*.
- Hollink, L., Bedjeti, A., Van Harmelen, M., & Elliott, D. (2016, May). A corpus of images and text in online news. *In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 1377-1382).
- Kassogué, B., Kassogué, P. T., & Dolo, S. (2019). Méthode de Recherche: Grille d'Analyse d'un Travail Scientifique. *International journal of Scientific and Research*, 10(1), 1399-1402.
- Lanciano-Morandat, C., Rolle, P., & Bouffartigue, P. S. (2019). Le travail de recherche. Production de savoirs et pratiques scientifiques et techniques. CNRS éditions.
- Latour, B., & Biezunski, M. (2005). La science en action : introduction à la sociologie des sciences.
- Peroni, S. (2016). SAMOD: an agile methodology for the development of ontologies. *In Proceedings of the 13th OWL: Experiences and Directions Workshop and 5th OWL reasoner evaluation workshop (OWLED-ORE 2016)* (pp. 1-14).
- Rawsthorne, H. M., Abadie, N., Kergosien, E., Duchêne, C., & Saux, E. (2022, June). ATLANTIS: Une ontologie pour représenter les Instructions nautiques. *In IC 2022, 33es journées francophones d'Ingénierie des connaissances* (pp. 154-163).
- Soegaard M. (2008). *Interaction Styles*, [http://www.interaction-design.org/encyclopedia/interaction\\_styles.html](http://www.interaction-design.org/encyclopedia/interaction_styles.html)
- Staab, S., Studer, R., Schnurr, H. P., & Sure, Y. (2001). Knowledge processes and ontologies. *IEEE Intelligent systems*, 16(1), 26-34.
- Suàrez-Figuero M.C., Gómez-Perez A., Motta E., Gangemi A. (2012). The NeOn Methodology for Ontology Engineering, Springer, pp 9-34
- Swartout, B., Patil, R., Knight, K., & Russ, T. (1996, November). Toward distributed use of large-scale ontologies. *In Proc. of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems* (Vol. 138, No. 148, p. 25).
- Vrandečić, D. (2012, April). Wikidata: A new platform for collaborative data collection. *In Proceedings of the 21st international conference on world wide web* (pp. 1063-1064).