



## Colloque Jeunes Chercheurs PRAXILING 2019

### « Constitution assistée d'un parcours de formation s'appuyant sur les outils du web sémantique et l'analyse de corpus textuels » *Présentation des travaux*

*Pierre-Bernard TOUBOL*  
Ingénieur de recherche  
Doctorant  
Humanités Numériques

Sous la direction de :

*Chrysta Peljssier*  
PRAXILING - Équipe 2  
et :  
*Mathieu Lafourcade*  
LIRMM - Équipe TEXTE

# Menu

A. État de l'art et objectifs de la recherche et substrats théoriques

B. Démarche, méthodes et premiers résultats

C. Participations et communications

D. Planning prévisionnel

# Positionnement de la recherche

- Massification des enseignements dans le supérieur.
- Besoin de personnalisation des apprentissages fonction des acteurs (apprenants et enseignants), des contextes et des objectifs (de formation et pédagogiques).
- Systèmes de gestion numérique d'apprentissage (EIAH) nécessitant un accompagnement des enseignants dans la création de scénarisations de leur enseignement.

# Objectif de la recherche

- Un objectif :
  - bâtir une méthode, et un modèle permettant de structurer les connaissances disciplinaires destinées à alimenter un outil d'aide à la production d'un parcours de formation.

La didactique en tant que modélisation d'une connaissance en vue de son apprentissage, est une démarche pouvant s'appuyer sur une logique *épistémologique* que l'on pourrait décrire et générer à l'aide d'une méthode d'analyse sémantiques automatisée.

# Épistémologie et langage scientifique



- Épistémologie en tant que concept participant à l'élaboration du langage scientifique par la reconnaissance d'unité des formes dans la connaissance dans leur rapport avec la science
- Les connaissances sont en perpétuelles croissance : c'est un problème central pour l'épistémologie...
  - nécessité de la prise en compte de leur évolution

# Choix du champ disciplinaire

## La thermodynamique

- Champs conceptuel et lexical très large faisant appel à des disciplines multiples : la chimie, la physique, les mathématiques et... l'histoire en raison de sa récente émergence en tant que science
- Il est d'usage de présenter sa didactique comme au carrefour de toutes ces champs disciplinaires qui ont participé à son développement.
- La problématique didactique conduit à remanier les outils conceptuels ou méthodologiques que la recherche leur emprunte (S. Rozier 1988)

# "Phylogénie" des concepts savants

*"Un concept est l'ensemble des rapports entre une forme lexicale stable qui se répète durant un intervalle de temps fini par un ou plusieurs locuteurs, des formes lexicales cooccurrentes qui accompagnent de manière variable cette forme stable, et des événements non lexicaux (sociaux, cérébraux, biologiques, etc.)."* (P. Charbonnat 2014)

- Charbonnat emprunte à la biologie la notion de **phylogénie** qui correspond à l'étude des liens existant entre concepts apparentés.
- Il fait apparaître au cœur de sa définition la notion de **contexte** que nous chercherons à modéliser par les relations internes aux corpus textuels
- Pour bâtir ces contextes, nous chercherons à exprimer formellement ces relations entre les différentes formes lexicales et par les relations entre "acteurs" de l'apprentissage  
Des analyses approchant cet objectif (mise en évidence des "obstacles épistémologiques") ont été développées par G. Brousseau en mathématiques ou L. Viennot en physique ou encore S. Rozier en thermodynamique.



# Objectif des enquêtes mises en œuvre

- Mettre en évidence et analyser des pratiques didactiques des enseignants du supérieur
  - entretiens semi-directifs
  - enquêtes par observation
- Faire émerger une "phylogénèse" des concepts savants de thermodynamique par une analyse de corpus conduite selon une méthode de "théorisation ancrée" (Glaser et al. 1968, Strauss et Corbin 1990, Lejeune 2016)



# Corpus didactiques : identification et capture des relations sémantiques



- Après sélection d'un ensemble de support de cours (en accès libre ou bénéficiant d'une autorisation explicite)
- Capture des relations sémantiques en faisant appel à des techniques de TALN.
- Alimentation de la base de connaissance structurée dans le réseau lexical JDM (Lafourcade 2008) : plus approprié pour modéliser une phylogénie ainsi qu'une ontogénèse (évolution dans le temps) de concepts qu'une simple ontologie.

# Réseau lexical vs ontologie de domaine

- Les ontologies de domaine issues de bases de connaissances spécifiques ont une valeur certaine mais sont limitées dans leur portée, spécialisé et dont l'agilité est limitée.
- JDM apporte des méta-informations
  - la force d'association (poids)
  - des annotations (fréquence et pertinence)
- réseau lexical généraliste "Jeux de mots" (Lafourcade 2008) intègre une méthode d'apprentissage sans fin.

# La Cible

## Sources textuelles :

- Entretiens (enquêtes)
- Support de cours
- Autres

Paramètres  
de contexte

extraction de triplets  
C-R-C

## Base de connaissances

JDM (Jeux de mots)

IA : interrogation  
de la base

mise en forme  
"DataViz"

**Parcours**  
d'apprentissage  
et  
champs conceptuels  
contextualisés

# Menu

A. État de l'art et objectifs de la recherche  
et substrats théoriques

**B. Démarche, méthodes et premiers  
résultats**

C. Participations et communications

D. Planning prévisionnel

# La démarche

- Analyse de discours issus d'entretiens.
  - objectifs :  
Recueillir et questionner les pratiques (*Pélissier*) et les expériences des acteurs de la formation concernant la création d'un parcours de formation pour une discipline donnée.



# Premiers résultats : l'analyse d'un processus "d'aide"

- Choix de privilégier une approche Système d'Information pour l'analyse du processus :

L'aide à la transmission de connaissances est vue comme un processus interne au Système d'Information (processus de support au sens de Porter (1986))

- le processus d'aide est décrit grâce à un formalisme de description et de documentation concernant :
  - les acteurs
  - les étapes
  - les produits
  - les indicateurs de qualité
- on définit le processus de l'aide par son périmètre

# Modèle de référence de travail

- Modéliser la connaissance dans la Théorie Anthropologique du Didactique (*Chaachoua 2010*) :
  - Toute activité humaine met en œuvre une organisation que Chevallard (1998) note [ T / τ / θ / Θ ] et qu'il nomme praxéologie, ou organisation praxéologique.
    - [ T / τ ] étant la pratique – ou encore le savoir- faire ;
    - [ θ / Θ ] le logos – ou encore le savoir.

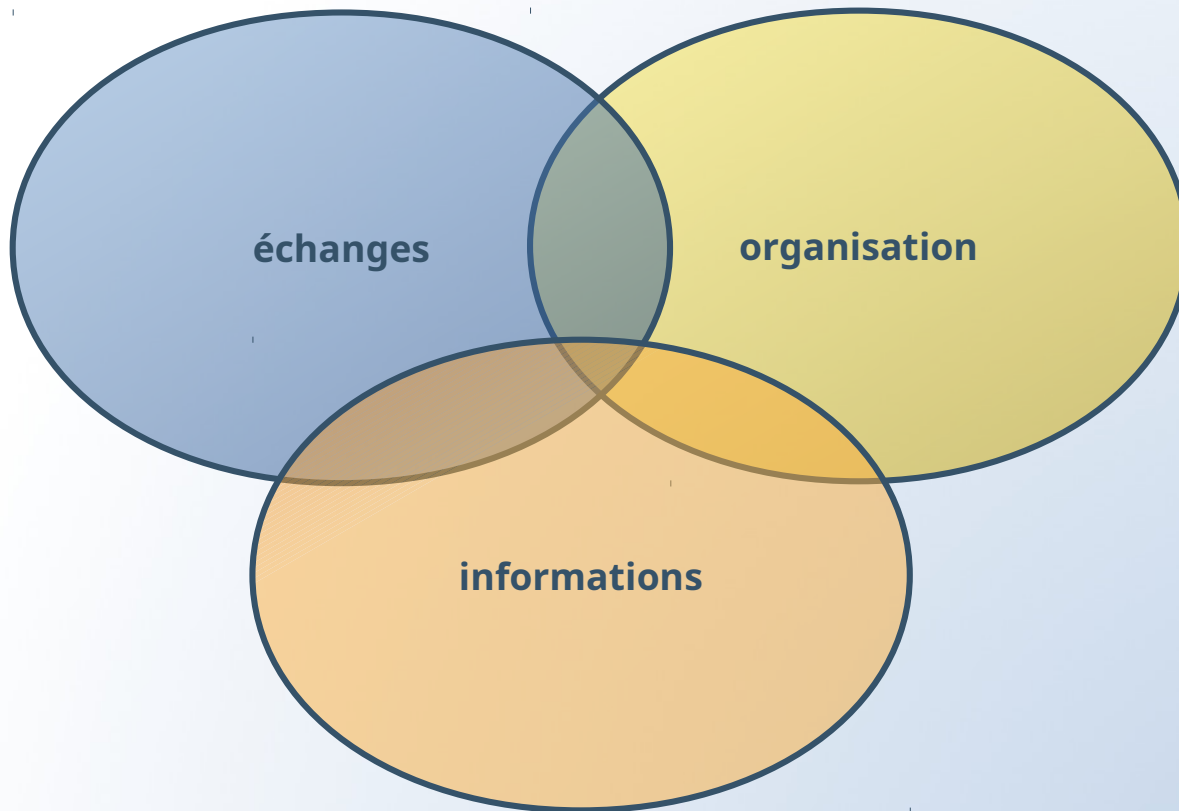
→ On s'appuiera sur la praxéologie didactique avec comme type de tâche T les tâches d'étude propre à la discipline choisie.



# Le processus

- C'est un ensemble d'activités reliées entre elles, réalisées par des acteurs et concourant à atteindre un objectif fixé. Il décrit l'enchaînement des activités nécessaires à la production des résultats.
- Qui fait quoi et comment :
  - Les acteurs qui interviennent ;
  - Les actes réalisés (transformation d'information) ;
  - Selon quel ordre, quelles séquences (mise en évidence des frontières entre acteurs) ?
- Comment le font-ils :
  - Description des actes incluant les outils de traitement

# 3 vues - 3 paradigmes



# Vue "échanges"

C'est une vue extérieure au processus.  
Elle décrit les échanges d'information  
entre:

- les interlocuteurs à qui ce processus destine les résultats de son action.
  - tous les acteurs dont il dépend pour produire ces résultats.
- C'est une vue axée sur les échanges



# Vue "organisationnelle/fonctionnelle"

C'est une vue interne, elle décrit :

- Les acteurs (ou groupe d'acteurs) physiques ou organisationnels qui collaborent à l'atteinte des résultats (objectifs). Ces acteurs sont illustrés par leurs compétences (domaine de responsabilité, pouvoir de décision).
- Elle met en évidence la façon dont l'organisation prend en charge ses missions



# Vue "informationnelle"

- L'ensemble des informations manipulées (les concepts et leurs inter-relations) qui constituent un micro-univers : la représentation restreinte du monde dont s'est dotée l'organisation pour accomplir ses missions.
- La démarche consiste à identifier, extraire, et modéliser l'information
- L'information n'est pas une information structurée : elle a donc besoin d'être modélisée, formalisée afin de simplifier son analyse.
- On cherche à en décrire son cycle de vie, depuis sa production jusqu'à sa délivrance.

# Synthèse et premiers livrables

Le recueil et l'identification des informations récoltées dans le cadre de ces trois paradigmes (les trois vues) et articulées les unes par rapport aux autres feront l'objet d'un guide de référence du processus considéré :

→ Le "référentiel d'aide"

# Choix des outils : les principes

Une chaîne de traitement numérique résolument basée sur des outils "open-source" et ouverts permettant des traitements multiples au travers

- d'outils dédiés (RQDA, Tropes, Iramuteq, TXM) ;
- de langages orientés vers le traitement de données et l'analyse statistique (R ou Python).



# RQDA - fonctions "avancées"

- Un ensemble d'items accessibles au travers de menus contextuels relatifs :
  - aux fichiers importés
  - aux codes créés
  - aux catégories
  - aux mots apparaissant dans le corpus
  - aux typologies de fichier
- Toutes les variables créées peuvent être documentées et/ou journalisées



- Chaque concept/catégorie extrait est projeté dans le réseau lexical
- Objectifs :
  - Construire sur un réseau lexical généraliste un réseau lexical dédié à la didactique des sciences physiques (Thermodynamique) et d'aide à l'apprentissage.
  - Ce réseau combinera poids et annotations sur des relations typées entre des termes et des concepts. Il sera à l'origine de référentiels devant aboutir à la proposition de parcours de formation.

# Description du corpus actuel

- Acteurs de l'enseignement supérieur :
  - enseignants et chercheurs (universités françaises et canadiennes)
  - ingénieurs pédagogiques (universités françaises, écoles et grands établissements)
  - étudiants (Québec)
- 21 entretiens
- modalités :
  - présentiel
  - téléphone
  - Skype

# Guides d'entretien

- Deux guides d'entretien dynamiques permettant de s'adapter aux divers interlocuteurs avant et pendant les entretiens (cartes heuristiques)
- L'adaptation des guides se déroule de manière incrémentale selon les catégories émergentes d'une *analyse qualitative ancrée*. (Glaser, al. 1968, Lejeune 2016)



# Guide de codage

- Il s'appuie sur les deux niveaux de catégorie proposés dans le cadre de RQDA
- une "condensation" des catégories d'analyse à partir des étiquettes assignées à des portions de corpus illustrant une typologie de l'aide réellement mise en œuvre.
- Un codage lié à la vue organisationnelle/fonctionnelle
  - identification des acteurs
  - relations entre acteurs
  - identification des actions
  - séquençage des actions
- un codage lié à la vue informationnelle permettant l'identification des "objets traités".



# Première catégorisation (*provisoire*) d'aide

- Catégories d'aide identifiées :
  - Mise en confiance et empathie
  - Exploration : susciter une situation, émettre des hypothèses, faire préciser un mot ou un concept.
  - Nature de la construction à partir de la didactique de la discipline : poser et hiérarchiser les concepts, mettre à jour les contenus.
  - Définition du champ d'intervention dans le dispositif pédagogique proposé : identifier l'accompagnement dans ses objectifs identifier le paramètre institutionnel.
  - Relation entre l'apprenant, son apprentissage et la personne aidée : définir les prérequis, définir les objectifs, gérer les profils d'apprenant.
  - Appui à la pratique pédagogique : positionner une action dans une démarche d'apprentissage, qualifier et évaluer une pratique, relier un propos à une théorie de l'apprentissage.





Catégories d'aide (1/2)	Exemples d'expression
Mise en confiance et empathie	<p>Aider le locuteur dans son expression :  <i>"...C'est votre point de vue qui m'intéresse.."</i></p> <p>Marquer la qualité de son écoute :  <i>"OK," "Je vois", "Tout-à-fait"</i></p> <p>Répéter un mot ou une expression en écho-miroir :  <i>"...donc la didactique c'est essayer de structurer un enseignement ou tout au moins des connaissances pour pouvoir atteindre un objectif de formation...."</i></p>
Exploration	<p>Susciter une situation ou une idée :  <i>"...Ce qui serait intéressant effectivement, c'est qu'il existe un moyen de les inciter à faire ces liens et tout au moins leur montrer de tels liens ..."</i></p> <p>Émettre des hypothèses :  <i>"... et ça cela veut dire que les étudiants ne savent pas travailler.."</i></p> <p>Faire préciser un mot ou un concept :  <i>"...Quelle est votre définition de la didactique ?..."</i></p>
Nature de la construction à partir de la didactique de la discipline	<p>Hiérarchiser les concepts :  <i>"...Et la hiérarchisation des concepts à acquérir : il faut apprendre telle chose avant telle chose. etc..."</i></p> <p>Mettre à jour des contenus :  <i>"...comment vous introduisez les apports de la recherche dans votre enseignement ?..."</i></p> <p>Poser les concepts :  <i>"... pouvez-vous me définir votre définition à vous , celle qui vous semble la plus pertinente de la différence entre didactique et pédagogie ?..."</i></p>






Catégories d'aide (2/2)	Exemples d'expression
<p>Définition du champ d'intervention dans le dispositif pédagogique proposé</p>	<p>Identifier l'accompagnement dans ses objectifs :  <i>"...Ce que l'on attends d'un ingénieur pédagogique c'est ça ? C'est uniquement : j'arrive mon corpus , débrouille toi pour mettre les bons outils qui vont me permettre de faire passer ce corpus...."</i></p> <p>Identifier le périmètre institutionnel de l'accompagnement proposé :  <i>"... C'est de ça dont vous parliez en parlant de "cadre"...."</i></p>
<p>Relation entre l'apprenant, son apprentissage et la personne aidée</p>	<p>Définir des prérequis :  <i>"On devrait être capable de dire que tel concept on devrait avoir tel ou tel prérequis pour pouvoir le comprendre"</i></p> <p>Définir un objectif de formation :  <i>"Un expert du domaine qui est capable de dire que la notion d'enthalpie doit être apprise que parce que on a appris quelque part ce que c'était que l'entropie et l'entropie que parce qu'on a su quelque part ce que c'était que l'énergie et que sais-je encore ?"</i></p> <p>Gérer les profils d'apprenant :  <i>"Ça veut dire qu'il faut déjà faire une analyse de l'apprenant et du profil de l'apprenant avant la formation ?"</i></p>
<p>Appui à la pratique pédagogique</p>	<p>Positionner une action dans une démarche d'apprentissage :  <i>"Vous dites que les enseignants de manière général connaissent le fond, que vous n'intervenez pas sur le fond, cela dit comment adapter les outils au fond."</i></p> <p>Qualifier une pratique :  <i>"Donc le positionnement de l'ingénieur pédagogique c'est un accompagnateur de l'ensemble des acteurs de la formation: aussi bien aussi bien apprenant qu'enseignant"</i></p> <p>Relier un propos à une théorie de l'apprentissage :  <i>"tu t'appuies beaucoup sur les échanges et la découverte. Tu fais du socioconstructivisme"</i></p>

Ces catégories constituent des éléments clés pour :

- une base à l'alimentation du guide visant à faire expliciter par la personne interviewée la nature des activités qu'elle mène en tant qu'enseignant produisant seul un cours ou en tant qu'ingénieur pédagogique aidant un enseignant à structurer son cours.
- une mise en évidence d'éventuelles ruptures dans la chaîne de l'organisation des connaissances en vue de les transmettre.
- alimenter la démarche de production d'un outil informatique intégré destiné aux enseignants pour les aider à structurer les connaissances à transmettre et aux ingénieurs pédagogiques pour les aider dans leur travail d'aide à l'enseignant.



## Bibliographie sommaire

- 
- ARTIGUE, Michèle, 1989. \*Épistémologie et didactique.\* [en ligne]. S.I. : IREM. Cahier de DIDIREM, 3. Disponible à l'adresse : <<http://docs.irem.univ-paris-diderot.fr/up/publications/IPS97066.pdf>>.
  - BALACHEFF, Nicolas et MARGOLINAS, Claire, 2003. cK Modèle de connaissances pour le calcul de situations didactiques. In : \*XII° école d'été de didactique des mathématiques\* [en ligne]. S.I. : La Pensée Sauvage éditions. 2003. p. 1–32. [Consulté le 20 juin 2017]. Disp
  - BROUSSEAU G., 1998, Théorie des situations didactiques, 395p, ed. La Pensée Sauvage, Grenoble disponible à l'adresse : <<https://hal.archives-ouvertes.fr/hal-01139408/>>.
  - CHAACHOUA, Hamid, 2010. La praxéologie comme modèle didactique pour la problématique EIAH. Etude de cas: la modélisation des connaissances des élèves. HDR. S.I. : Université de Grenoble.
  - CHEVALLARD, Yves, 1991. La transposition didactique. S.I. : La pensée sauvage Grenoble. ISBN 2-85919-078-3.
  - CHEVALLARD, Yves, 2003. Didactique et formation des enseignants. In : Journées d'études INRP-GÉDIAPS. S.I. : s.n. 20 mars 2003. p. 14.
  - GIORDAN, André et DE VECCHI, Gérard, 1987. Les origines du savoir. In : \*Neuchâtel, Delachaux et Niestlé\*. 1987.
  - GLASER, Barney G. and Strauss, Anselm L. and Strutzel, Elizabeth, The Discovery of Grounded Theory; Strategies for Qualitative Research, 1968
  - LAFOURCADE, Mathieu et JOUBERT, Alain, 2008. Jeuxdemots: un prototype ludique pour l'émergence de relations entre termes. In : \*JADT'08: Journées internationales d'Analyse statistiques des Données Textuelles\* [en ligne]. S.I. : s.n. 2008. p. 657–666. [Consulté le 9 mars 2016]. Disponible à l'adresse : <<https://hal.archives-ouvertes.fr/lirmm-00358848/>>.
  - LAFOURCADE, Mathieu et JOUBERT, Alain, 2018. Si beaucoup d'oiseaux volent, peut-on en déduire que tous les oiseaux volent ? Production endogène de règles déductives dans un réseau lexico-sémantique. In : \*Proceedings TALIA 2018\*. Nancy : s.n. 6 juillet 2018. p. 10.
  - LEJEUNE, Christophe, Manuel d'analyse qualitative: analyser sans compter ni classer, 2016
  - PELISSIER, Chrysta, Des Aides à l'apprentissage Pour La Formation à l'Université, 2011



Merci de votre attention...

... questions, précisions, suggestions ?