



# Constitution assistée d'un parcours de formation s'appuyant sur les outils du web sémantique et l'analyse de corpus

Pierre-Bernard Toubol

## ► To cite this version:

Pierre-Bernard Toubol. Constitution assistée d'un parcours de formation s'appuyant sur les outils du web sémantique et l'analyse de corpus. Septièmes Rencontres Jeunes Chercheurs en EIAH (RJC EIAH 2018), Apr 2018, Besançon, France. 2018. <hal-01769604>

**HAL Id: hal-01769604**

**<https://hal.archives-ouvertes.fr/hal-01769604>**

Submitted on 24 Apr 2018

**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.

# Constitution assistée d'un parcours de formation s'appuyant sur les outils du web sémantique et l'analyse de corpus

Pierre-Bernard Toubol

<sup>1</sup> PRAXILING UMR 5267 CNRS - Université Paul Valéry - Montpellier 3

<sup>2</sup> LIRMM UMR 5506 CNRS - Université de Montpellier

`pierre-bernard.toubol@lr-universites.fr`

**Résumé** Dans cette recherche nous voulons structurer différentes techniques intégrées dans le domaine des Humanités Numériques. Notre but est d'obtenir des connaissances expertes à partir de corpus de textes ou d'entretien. Pour cela, nous ferons appel à des outils de traitement automatique du langage naturel pour alimenter un réseau lexical en français. Des techniques de l'intelligence artificielle seront appliquées. Ainsi nous construirons des modèles pour décrire les connaissances didactiques d'une discipline donnée et les parcours d'apprentissages associés à une typologie d'apprenants. Un outil graphique sera proposé à l'enseignant ou à l'apprenant pour faciliter la construction de ces parcours.

**Mots-clés** : modèle de connaissances, modèle de l'apprenant, scénarisation d'apprentissage, intelligence artificielle, traitement automatique du langage naturel

## 1 Contexte

Les systèmes de gestion numérique d'apprentissage sont aujourd'hui utilisés avec succès dans la communauté de l'enseignement supérieur (ESR). Ils constituent une aide à la création de cours en ligne qu'il est possible de scénariser dans le cadre d'une chaîne éditoriale. Cependant, leur conception reste complexe.

Les modèles de connaissance constituent une médiation permettant l'analyse et la compréhension. Nous proposons de nous appuyer sur une approche issue de l'Intelligence Artificielle (IA) et du domaine de l'Ingénierie des Connaissances (IC). Le champ de recherche portant sur la conception d'un tel outil convoque plusieurs disciplines : l'informatique, la didactique des disciplines, les sciences de l'éducation et la psychologie cognitive. À un niveau plus large, l'IC nécessite de prendre en compte les analyses issues de l'ergonomie, les sciences des organisations, la linguistique ou le traitement automatique des langues (TAL) [2] *"L'enjeu de ces travaux, est de proposer non pas une seule et même méthode liée à des pratiques sociales pédagogiques et scientifiques intégrant l'outil numérique, mais bien plusieurs procédures...Des travaux en sociologie envisagent la science comme un système d'échange entre différents réseaux formés par les [différents acteurs]"* [10] (p. 23). Il s'agira ici des acteurs de l'ESR.

## 2 État de l'art

D'après [3], l'IC <sup>3</sup>s'appuie sur différents types de sources dont par exemple les connaissances individuelles, celles liées aux organisations, les connaissances partagées ou provenant du recoupement de données ou d'information réparties sur le web, etc...

La capture de relations sémantiques entre termes à partir de textes est un moyen privilégié de constituer/alimenter une base de connaissances, ressource indispensable pour l'analyse de textes. Des méthodes de production de relations lexico-sémantiques ont été proposées[8]. Un graphe (réseau lexical) contenant les objets du texte (les mots), des objets identifiés comme pertinents (des syntagmes, des concepts) ainsi que des relations pondérées et typées entre ces objets a été déployé au sein du projet JeuxDeMots (JDM).

La conception d'un Environnements Informatiques pour l'Apprentissage Humain (EIAH) repose sur des modélisations multiples couvrant différents points de vue et différents niveaux d'abstraction [9].

## 3 Propositions

Le recueil des connaissances sera basé sur une analyse en TAL des corpus proposés par l'enseignant et des transcriptions d'entretiens réalisés avec les experts du domaine : ici nous choisirons les bases de la thermodynamique nécessaires à l'étude des technologies énergétiques enseignées dans le supérieur. L'ensemble des concepts et des liens ainsi repérés sera intégré dans un réseau lexical. Nous construirons les outils d'apprentissage profond et les requêtes ad-hoc afin d'alimenter et d'interroger le graphe[7].

Nous chercherons à montrer comment une gestion du projet d'élicitation des connaissances (CommonKADS) pourra être utilisée à des fins cognitives. Cette méthode <sup>4</sup> constitue actuellement l'une des plus complètes : acquisition et ingénierie des connaissances pour les systèmes experts. Le méta modèle contextuel décrit les niveaux conceptuels (connaissances didactiques), organisationnels (activité pédagogiques) et agents (enseignants ou apprenants).

Afin de décrire puis de choisir un parcours de formation, nous nous appuierons sur une méthode utilisant la notion de prototype "Successive Approximation Model" (SAM) [1]. Le prototype de scénarisation permettra de matérialiser et de communiquer le contenu de formation aux parties prenantes.

## 4 Nature des résultats attendus

Dans ces réseaux lexicaux, nous modéliserons les apprenants au travers de leurs différents modes d'apprentissage. Nous définirons une méthode permettant la production d'une ontologie didactique bâtie sur les différentes connaissances à

<sup>3</sup> Considérée comme un domaine scientifique à part entière cf. [4]

<sup>4</sup> Un méta modèle d'IC selon [6]

acquérir en fonction des objectifs de formation. Les éléments du réseau lexical modélisant l'apprenant s'appuieront sur les définitions proposées par Myers-Briggs et McCaulley en 1962 (MBTI) puis reprises par les modèles Felder et Silverman [5].

Cette architecture a été partiellement décrite par [11]. Nos travaux viseront à améliorer le dispositif en faisant appel aux techniques de production de réseau lexical et aux techniques d'extraction de relations sémantiques décrites par [8]. Une évaluation sera réalisée sur la fonction scénarisation des parcours de formation et cherchera à répondre à la question : les scénarios proposés par l'outil permettent-ils un gain de temps et de qualité pour les enseignants et les apprenants ?

Le choix et l'articulation des outils associés constitueront une grande part de l'intérêt de cette recherche.

## Références

1. Allen, M.W., Sites, R.H. : Leaving ADDIE for SAM : An Agile Model for Developing the Best Learning Experiences. American Society for Training and Development, Alexandria, Va. (2012), oCLC : ocn817783764
2. Aussenac-Gilles, N. : Méthodes Ascendantes Pour l'ingénierie Des Connaissances. Habilitation à diriger des recherches, Université Paul Sabatier - Toulouse III (Dec 2005)
3. Aussenac-Gilles, N., Charlet, J., Reynaud, C. : Les enjeux de l'Ingénierie des connaissances. In : Information, Interaction, Intelligence-Le Point Sur Le i (3), pp. 244–266. Cépaduès éd., Toulouse (2012), 00000
4. Bachimont, B. : Pourquoi n'y a-t-il pas d'expérience en ingénierie des connaissances ? In : 15èmes Journées Francophones d'ingénierie Des Connaissances. pp. 53–64 (2004)
5. Felder, R.M., Silverman, L.K. : Learning and teaching styles in engineering education. *Engineering education* 78(7), 674–681 (1988)
6. Guin, N. : Des Meta-Modèles Pour Guider l'elicitaiton Des Connaissances En EIAH : Contributions à l'enseignement de Methodes et à La Personnalisation Des Activites. Habilitation à diriger des recherches, Université Claude Bernard - Lyon 1 (Dec 2014)
7. Lafourcade, M. : Lexique et Analyse Semantique de Textes - Structures, Acquisitions, Calculs, et Jeux de Mots. Habilitation à diriger des recherches, Citeseer (2011)
8. Le Brun, N., Lafourcade, M. : Parcourir, reconnaître et réfléchir. Combinaison de méthodes légères pour l'extraction de relations sémantiques. In : 24e Conférence Sur Le Traitement Automatique Des Langues Naturelles (TALN). p. 150 (2017)
9. Nodenot, T. : Contribution à l'Ingénierie Dirigée Par Les Modèles En EIAH : Le Cas Des Situations-Problèmes Coopératives. Habilitation à diriger des recherches, Université de Pau et des Pays de l'Adour (Nov 2005)
10. Pélissier, C. : L'aide En Question : Un Composant Majeur Dans l'éducation et Les Humanités Numériques. Habilitation à diriger des recherches, Paul Valéry, Montpellier 3 (Apr 2017)
11. Zouaq, A. : Une Approche d'ingénierie Ontologique Pour l'acquisition et l'exploitation Des Connaissances à Partir de Documents Textuels : Vers Des Objets de Connaissances et d'apprentissage. Ph.D. thesis, Montréal (Dec 2007)