Une ontologie pour aider à l’apprentissage des statistiques : perspectives et challenges

Jean-Marc Meunier

# Introduction

S’agissant des statistiques, comme bien d’autres domaines, la quantité de ressources pédagogiques libres ou éditoriales disponibles sur le web est colossale. Pour s’en convaincre, il suffit de lancer sur un moteur de recherche une requête pour trouver des centaines de ressources allant de cours complets à des exercices ou des jeux de données sur lesquels les enseignants peuvent élaborer de nouveau exercices et les étudiants s’entraîner ou  parfaire leur compréhension des concepts. Cette opportunité, tant pour l’apprenant que pour l’enseignant nécessite cependant que ces deux types d’usagers possèdent un certain nombre de compétences requises par ces nouveaux médias  pour s’orienter dans toutes ces ressources et en tirer le meilleurs parti (Terras, Ramsay, and Boyle 2013).   Ces compétences sont la capacité à chercher l’information,  à faire le tri parmi les résultats de la recherche puis à extraire de ces documents les éléments utiles et à les traiter pour développer ses connaissances ou, lorsqu’on est enseignant, faciliter l’acquisition à ses étudiants. Accessoirement ces habiletés supposent aussi des compétences à communiquer sur le domaine objet de l’apprentissage entre étudiants, entre enseignants ou entre étudiants et enseignants.  L’ensemble de ces facteurs, disponibilités des ressources et nouvelles compétences requises changent considérablement la façon d’appréhender l’apprentissage et l’enseignement (Barnes 2007).  Ces nouvelles façons de “consommer” les connaissances font passer l’apprenant d’une position passive de réception à une position active de construction de son savoir qui nécessite cependant un étayage (Vygotski et al. 2013) de la part des enseignants et/ou de la part des dispositifs d’apprentissage qu’il a conçu. L’apprentissage à l’aide des ressources éducatives libres est de toute évidence une activité sociale comme en témoigne les communautés d’échanges  sur des forums ou des réseaux sociaux.

A titre d’exemple, nous pouvons citer  le groupe “ Research, Methodology, and Statistics in the Social Sciences ” sur Likedin qui regroupent plus de 120 00 membres et dont l’objet est le partage d’expertise entre des chercheurs et des étudiants.  Pour se repérer dans cette masse de ressources, il est nécessaire de doter celles-ci d’une sémantique qui puisse servir d’appui aux échanges et à la recherche de ressources.

Par ailleurs, la représentation du dispositif de recherche constituent également un facteur important. Ainsi Holman (2011) a montré que ces stratégies de recherche sont par la représentation des relations que les termes de la recherche entretiennent entre eux. (A développer). L’alignement du dispositif sur les capacité cognitive de l’usager est  une condition de son utilisation optimale.  (finir de lire Terras 2013).  chercher  de l’information sur internet est assimilable à une activité de compréhension d’un texe (?)  Intégrer le texte de Barnes 2007 pour développer les différents aspects de la net génération.

Recent findings indicate that reading skills mediate the selection of  hyperlinks (Salmeron & Garcia, 2011) and that semantic link selection results in good inferential comprehension (Amadieu et al, 2009).

# Ontologie et structure de connaissances

Définition épistémologique cf bouzeghoub 2005

## Définition informatiqueOntologie et vocabulaireLes différents types d’ontologie

# Pourquoi structurer les concepts statistiques dans une ontologie ?

## prépondérance du but (voir (Bandalos, Finney, and Geske 2003)Parce que la structure des connaissances de l’expert est une ontologie de concepts (pas de termes)

Propriétés structurales vs propriétés fonctionnelles

Souligner l’importance de la structuration de la connaissances pour la résolution de problèmes. Théorie de la charge cognitive de Sweller : les structures cognitives que sont les schémas permettent de libérer les ressources cognitives et donc d’apprendre plus facilement cf <http://www.instructionaldesign.org/theories/cognitive-load.html>

Exemple de travaux sur la

structure des connaissances en statistiques

# En quoi cela peut-il aider l’étudiant ?

## Activités élaboratives et collaborativesL’évaluation des apprenantsMise à disposition d’une base de connaissances structurée et de  ressources pédagogiquesPersonnalisation de l’enseignement

# En quoi cela peut-il aider l’enseignant ?

## Le design de cours cf concept map designDisposer d’une base de connaissancesPartage de ressources ; Indexation sémantique des ressources

L’utilisation des ontologies pour l’indexation des ressources pédagogiques est un domaine de recherche qui se développe de plus en plus car il permet à la fois de répondre à un certains nombres de limitation des normes,  (Bouzeghoub and Elbyed, n.d.) mais pour donner du sens aux ressources ainsi répertoriées et accompagner de nouveaux usages de ces ressources.

Les OER doivent être décrites par des métadonnées pour être interopérables. Le premier est le LOM  et sa déclinaison française LOM-FR. Ce standard est cependant complexe à utilisé er incomplet pour la description sémantique les dépendances entre ressources. Le SCORM propose un langge de srtucturation des composants. Ces standards ont pour fonctions principales l’indexation pour la recherche mais pas sur l’assemblage et l’adaptation des ressources.

Les auteurs proposent une architecture basée sur trois modèles :

* Celui du domaine grâce à une ontologie du domaine qui sert à indexer sémantiquement les ressources et les apprenants.
* le modèle de l’apprenant  : informations factuelle et ses connaissances, c’est-à-dire les concepts qu’il connait.
* Le modèle de description des ressources avec les caractéristiques éducatives (similaires à LOM et les caractéristiques sémantiques qui se décline en trois composantes : prérequis, contenu et résultats d’apprentisage (fonction d’acquisition selon les auteurs).

Une ressources peut être atomique ou complexe. Dans ce dernier cas, elle peut être décrite par un graphe de composition expliquant comment sont parcouru les composantes (séquentielle, alternative ou parallèlisme).

Un langage d’interrogation SeRQL permet de naviguer dans les trois modèles, de faire des inférences et de définir les les chemins pertinents pour l’apprenant.

# Challenges

## Maintenir, traduire et étendre l’ontologie

Raisonner sur l’ontologie

Adaptative learning et recommandations de ressources

Affiner notre connaissance de la structure de connaissances

# References

Terras, Melody M., Judith Ramsay, and Elizabeth Boyle. 2013. “Learning and Open Educational Resources: A Psychological Perspective”. *E-Learning and Digital Media* 10 (2): 161–73.

Barnes, KassandraMarateo. 2007. “Teaching and Learning with the Net Generation”. *Innovate: Journal of Online Education* 3 (4).

Vygotski, Lev-Sémionovitch, Jean Piaget, Lucien Sève, Yves Clot, and Françoise Sève. 2013. *Pensée Et Langage*. 4E édition. Paris: La Dispute.

Bandalos, Deborah L., Sara J. Finney, and Jenenne A. Geske. 2003. “A Model of Statistics Performance Based on Achievement Goal Theory.”. *Journal of Educational Psychology* 95 (3). American Psychological Association (APA): 604–16. doi:10.1037/0022-0663.95.3.604.

Bouzeghoub, A., and A. Elbyed. n.d. “An Ontology Mapping Algorithm to Share Learning Resources”. In *2006 2nd International Conference on Information & Communication Technologies*. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/ictta.2006.1684442.