

Production semi-automatisée d'une carte conceptuelle en science et technologie

Maxim Morin et Jean-Guy Blais

Volume 40, numéro 1, 2017

Réception : 1^{er} avril 2016

Version finale : 2 mai 2017

Acceptation : 5 mai 2017

URI : <https://id.erudit.org/iderudit/1041004ar>

DOI : <https://doi.org/10.7202/1041004ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Morin, M. & Blais, J.-G. (2017). Production semi-automatisée d'une carte conceptuelle en science et technologie. *Mesure et évaluation en éducation*, 40(1), 61–99. <https://doi.org/10.7202/1041004ar>

Résumé de l'article

Cet article décrit la conception et la mise à l'essai d'un programme informatique de production semi-automatisée de cartes conceptuelles. Cette technique dérivée de l'extraction d'information a pour but de produire une représentation simple et signifiante du contenu d'un ou de plusieurs textes sous forme d'un schéma de connaissances. Le programme a été mis à l'essai à partir de deux corpus de textes choisis pour couvrir deux thèmes du Programme de formation de l'école québécoise de quatrième secondaire en science et technologie. Les résultats de l'étude sont très encourageants et montrent le potentiel d'une telle approche pour faciliter la construction d'une carte conceptuelle, une tâche qui est généralement réalisée par des humains. Les résultats suscitent également quelques réflexions quant aux modalités courantes d'élaboration et d'évaluation de cartes conceptuelles.

Production semi-automatisée d'une carte conceptuelle en science et technologie

Maxim Morin

Jean-Guy Blais

Université de Montréal

MOTS CLÉS: génération automatique de cartes conceptuelles, carte conceptuelle, analyse de texte, exploration de données, traitement automatique du langage naturel

Cet article décrit la conception et la mise à l'essai d'un programme informatique de production semi-automatisée de cartes conceptuelles. Cette technique dérivée de l'extraction d'information a pour but de produire une représentation simple et signifiante du contenu d'un ou de plusieurs textes sous forme d'un schéma de connaissances. Le programme a été mis à l'essai à partir de deux corpus de textes choisis pour couvrir deux thèmes du Programme de formation de l'école québécoise de quatrième secondaire en science et technologie. Les résultats de l'étude sont très encourageants et montrent le potentiel d'une telle approche pour faciliter la construction d'une carte conceptuelle, une tâche qui est généralement réalisée par des humains. Les résultats suscitent également quelques réflexions quant aux modalités courantes d'élaboration et d'évaluation de cartes conceptuelles.

KEY WORDS: automatic concept mapping, concept map, text analysis, data mining, natural language processing

This article presents the development and testing of a computer program for semi-automatic production of concept maps. This technique derived from information extraction intends to provide a simple and significant representation of the content of one or many texts by a knowledge diagram. The application was tested with two text corpora chosen for covering two themes coming from the Programme de formation de l'école québécoise of Secondary 4 in science and technology. The very encouraging results demonstrate the potential of this approach for easing the construction of concept maps, a process that is usually performed by humans. The results arouse as well some questions about the current methods for creating and assessing concept maps.

PALAVRAS-CHAVE: geração automática de mapas conceituais, mapa conceitual, análise de texto, exploração de dados, processamento automático da linguagem natural

Este artigo descreve a concepção e o teste de um programa informático de produção semiautomatizada de mapas conceituais. Esta técnica derivada da extração de informação destina-se a produzir uma representação simples e significativa do conteúdo de um ou mais textos sob a forma de um esquema de conhecimentos. O programa foi testado a partir de dois corpus de textos selecionados para cobrir dois temas de ciência e tecnologia do Programa de formação do 4.º ano do secundário da escola quebequense. Os resultados do estudo são muito animadores e demonstram o potencial de uma tal abordagem para facilitar a construção de um mapa conceitual, uma tarefa que geralmente é realizada por seres humanos. Os resultados suscitam também algumas reflexões sobre as modalidades atuais de elaboração e de avaliação de mapas conceituais.

Introduction

La cartographie conceptuelle est une technique d'organisation et de représentation des connaissances largement répandue en éducation qui a fait ses preuves à la fois comme stratégie d'enseignement, stratégie d'apprentissage et stratégie d'évaluation des apprentissages (Cañas et al., 2003; Marchand & d'Ivernois, 2004). Jusqu'à tout récemment, l'élaboration d'une carte conceptuelle était une tâche réalisée par un ou des humains, par exemple un expert ou un panel d'experts. Les progrès technologiques en matière d'intelligence artificielle, de traitement automatique du langage naturel (TALN) et d'exploration des données (*data mining*) ont depuis peu ouvert la porte à des techniques de production automatisée ou semi-automatisée de cartes à partir de documents textuels. Les applications potentielles de telles techniques sont nombreuses, que ce soit pour résumer un texte (Chen, Kinshuk, Wei & Chen, 2008), pour corriger et noter automatiquement des questions à réponse élaborée (Clariana & Koul, 2004; Koul, Clariana & Salehi, 2005), pour créer une base de connaissances pour des systèmes tutoriels intelligents (Nielsen, Ward, Martin & Palmer, 2008) ou encore pour servir d'outil de soutien de correction lors de l'évaluation de cartes conceptuelles élaborées par des élèves.

Le présent article se penche sur la conception et la mise à l'essai d'un programme informatique permettant la production semi-automatisée d'une carte conceptuelle. Comme il sera expliqué plus loin, cette démarche s'inscrit dans une recherche plus vaste dont les objectifs sont, de façon générale, d'étudier l'usage de la cartographie conceptuelle comme instrument d'évaluation des apprentissages d'élèves et, de manière plus précise, d'étudier les propriétés psychométriques des données collectées au moyen des cartes conceptuelles.

La suite de cet article se divise en sept sections: le cadre théorique, le contexte de la recherche, la démarche, la méthodologie, les résultats de la mise à l'essai, la discussion et la conclusion.

Cadre théorique

Les appellations «carte conceptuelle» ou «cartographie conceptuelle» sont employées de plusieurs façons dans les écrits scientifiques ou pédagogiques. La première sous-section de ce cadre théorique définit ce qui est entendu par cartographie conceptuelle dans cette recherche. La seconde sous-section décrit plus spécifiquement les divers procédés et techniques qui peuvent être mis à profit lors de la production automatisée, tandis que la troisième sous-section fait état de quelques projets de recherches dans ce domaine. Une synthèse terminera cette section sur le cadre théorique.

Définition de la cartographie conceptuelle

D'abord utilisée comme technique de synthèse d'entretiens de recherche par Joseph D. Novak (1977), la cartographie conceptuelle s'est rapidement révélée comme étant une stratégie d'enseignement et d'apprentissage efficace en sciences (Novak, Gowin & Johansen, 1983) et dans plusieurs autres domaines (Nesbit & Adesope, 2006). Pour Novak, une carte conceptuelle est un réseau composé de nœuds et d'arcs organisés de manière semi-hiérarchique, qui a comme objectif «de représenter les relations significatives entre les concepts sous forme de propositions» (Novak & Gowin, 1984, p. 13). Du point de vue graphique, le nœud est formé d'une figure géométrique et l'arc est un trait fléché. Du point de vue sémantique, le nœud exprime un concept, c'est-à-dire «une régularité perçue (ou un patron) d'évènements ou d'objets, ou de répertoires d'évènements ou d'objets, désignée par une étiquette» (Novak & Cañas, 2008, p. 10). Quant à l'arc, il est accompagné d'une expression de liaison définissant la relation entre deux concepts. La mise en relation de deux ou plusieurs concepts forme des propositions, c'est-à-dire des énoncés de sens significatifs.

La démarche de construction d'une carte conceptuelle est importante aux yeux de Novak. Ce spécialiste des sciences de l'éducation encourage l'apprenant à participer activement à la construction de ses apprentissages, notamment en l'invitant à choisir lui-même ses libellés de concepts et à énoncer les relations entre ceux-ci. Qui plus est, Novak prône une organisation semi-hiérarchique des concepts et la formation de relations dites transversales, c'est-à-dire des relations entre des concepts éloignés dans la carte conceptuelle.

La définition de la cartographie conceptuelle promue par Novak est probablement la plus répandue, mais ce n'est pas la seule. Sa définition ou d'autres sont souvent adaptées, implicitement ou explicitement, au contexte pédagogique ou de recherche. À la fin des années 1990 et au début des années 2000, une équipe de chercheurs du *Center for Research on Evaluation, Standards and Student Testing* (CRESST) a étudié diverses modalités de la cartographie conceptuelle pour évaluer les apprentissages d'élèves, principalement dans le domaine de l'éducation scientifique (Herl, O'Neil, Chung & Schacter, 1999; Klein, Chung, Osmundson & Herl, 2002; Ruiz-Primo, Schultz, Li & Shavelson, 2001; Ruiz-Primo & Shavelson, 1996; Ruiz-Primo, Shavelson, Li & Schultz, 2001; Yin, Vanidez, Ruiz-Primo, Ayala & Shavelson, 2005). Leur définition de la cartographie conceptuelle est essentiellement la même que celle de Novak, sauf que ces chercheurs ne requièrent pas que la carte soit organisée de manière hiérarchique. Leurs travaux de recherche se démarquent aussi par l'éventail de modalités d'élaboration des cartes conceptuelles. Les chercheurs du CRESST s'intéressent aux tâches du type « remplir les trous d'une carte » ainsi qu'aux tâches qui consistent à élaborer une carte à partir de concepts ou de mots de liaison préétablis (Ruiz-Primo, Schultz et al., 2001; Ruiz-Primo, Shavelson et al., 2001; Yin & Shavelson, 2008; Yin et al., 2005). Puisqu'elles sont destinées à l'évaluation des apprentissages, ces modalités sont en général plus contraignantes que celles de Novak. Dans ces conditions, les cartes produites par les élèves sont plus uniformes, ce qui permet d'appliquer ultérieurement des modèles de notation quantitative.

Bref, la carte conceptuelle est un outil d'explicitation et d'organisation des connaissances qui sert essentiellement à l'humain (Valerio & Leake, 2006). Le contenu d'une carte est très souvent exprimé en langage naturel, ce qui en fait une représentation « informelle » des connaissances, au contraire d'autres représentations externes des connaissances comme le graphe conceptuel (Sowa, 2000) qui, lui, est de type « formel » et spécialement conçu pour être traité par un logiciel. La composition graphique et textuelle de la carte conceptuelle en fait un outil tout à fait adapté à la représentation de connaissances conceptuelles sous la forme de relations binaires, c'est-à-dire de triplets « concept, relation, concept ». Elle s'adapte moins bien à d'autres types de relations. Les relations unaires ne sont tout simplement pas permises, tandis que les relations ternaires sont

inévitablement divisées en deux relations binaires (Jean *donne un livre*, livre à Johanne) ou remplacées par une seule relation binaire (p. ex., Jean *donne un livre* à Johanne). La figure 1 illustre des relations de chaque ordre.

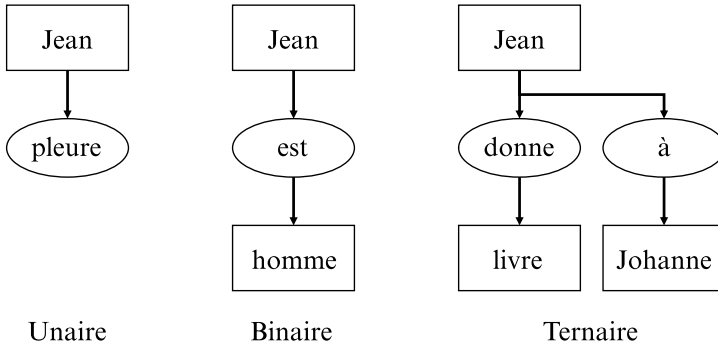


Figure 1. *Exemple de relations unaires, binaires et tertiaires*

Processus d'extraction d'information

La production automatisée d'une carte conceptuelle à partir de documents est essentiellement un problème d'extraction d'information au cours duquel un système extrait et organise des données textuelles en une représentation traitable par l'ordinateur et pertinente pour une application particulière. Typiquement, ces représentations de concepts et relations sont des structures relativement formelles (p. ex., des ontologies informatiques), et servent de bases de données pour d'autres programmes informatiques. En revanche, la production automatisée d'une carte a très souvent une visée éducationnelle et didactique. Selon Villalon et Calvo (2008), la carte générée doit refléter les caractéristiques fondamentales de la cartographie conceptuelle, et doit être une représentation de connaissances simple et compréhensible pour ses utilisateurs.

Une stratégie de production automatisée ou semi-automatisée d'une carte conceptuelle est un assemblage cohérent de techniques et de procédés statistiques ou linguistiques. Cette stratégie peut être divisée en quatre étapes: 1) le prétraitement des textes, 2) la reconnaissance des concepts, 3) la reconnaissance des relations, et 4) la synthèse de la carte.

Étant donné que l'ordinateur n'a pas la capacité de lire un texte comme un humain, il est nécessaire de réaliser une série d'opérations linguistiques permettant de construire une représentation compréhensible des textes pour la machine. Il s'agit de l'étape de prétraitement des textes. Les procédés courants sont les suivants :

- la *tokenisation* : processus de segmentation des textes en phrases et en mots ;
- la *normalisation des mots* : processus de représentation des mots par leur forme canonique (lemmatisation) ou leur racine (racinisation ou, en anglais, *stemming*) ;
- la *désambiguïsation des mots* : processus d'identification du sens d'un mot dans une phrase quand le mot est polysémique. Dans le contexte de la production d'une carte conceptuelle, ce processus se traduit entre autres par l'abstraction d'un concept à partir de plusieurs mots ou groupes de mots qui signifient la même chose, notamment des synonymes ;
- la *résolution d'anaphores* : processus par lequel les pronoms personnels sont remplacés par leur antécédent dans le texte ;
- l'*analyse morphosyntaxique* : processus d'analyse de la phrase qui reconnaît la forme morphologique de chaque mot (catégorie de mot, genre, nombre, personnes, etc.) et la structure de la phrase selon un formalisme donné. Ce formalisme est nommé une grammaire. On distingue généralement les grammaires en constituants et les grammaires de dépendance. La figure 2 donne un exemple d'un arbre syntaxique selon une grammaire de dépendance. Le verbe gouverneur « représentées » est la racine de cet arbre, le mot « relations » est le sujet de la phrase et la préposition « par » introduit le complément d'objet indirect. Chaque mot de la phrase est étiqueté ainsi.

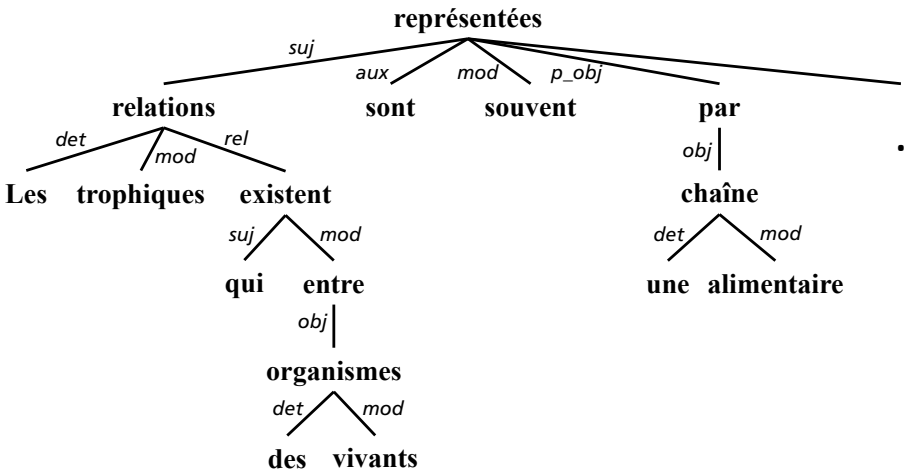


Figure 2. *Arbre de dépendance syntaxique de la phrase*

La seconde étape d'une démarche de construction automatisée d'une carte conceptuelle consiste à identifier les termes composés d'un ou de plusieurs mots qui désignent les concepts ou de potentiels concepts. Selon les besoins de l'utilisateur, le système peut repérer des termes prédéterminés, traiter les variations morphologiques des mots, proposer de nouveaux concepts ou même tenter de regrouper des termes similaires sous un seul concept. Le niveau de sophistication du logiciel croît rapidement selon les mécanismes choisis.

La troisième étape consiste à reconnaître les relations entre les concepts. Alors que des méthodes statistiques (p. ex., compter le nombre de mots séparant deux concepts) permettent de détecter la présence d'une relation, ce sont des méthodes linguistiques, fondées sur l'analyse morphosyntaxique de la phrase, qui reconnaissent les expressions de liaison formant les propositions.

La dernière étape du processus d'extraction a pour but de synthétiser l'information retenue lors des étapes précédentes. Selon Novak et Cañas (2008), une carte conceptuelle doit comporter entre 15 et 25 nœuds, mais il n'est pas rare que l'inventaire fait par le système dépasse ces balises. Lorsque cela se produit, les techniques pour synthétiser l'information sont nombreuses. Elles peuvent consister à sélectionner les concepts et les relations les plus signifiants ou encore à optimiser certaines propriétés struc-

turelles du réseau (p. ex., éliminer le morcèlement de la carte en réseaux distincts). Une fois les propositions choisies, elles peuvent être assemblées en une carte conceptuelle.

Recension des écrits

Un logiciel de production semi-automatisée de cartes conceptuelles est donc une combinaison de différentes techniques en vue d'effectuer une tâche précise. Quatre stratégies issues du domaine de l'éducation ou des sciences informatiques sont présentées dans cette section.

Clariana et Koul (2004) ont mis sur pied une stratégie d'extraction d'une variante d'un réseau de concepts, nommée un PFNet (abréviation de *pathfinder network*), en vue de la notation des productions écrites d'élèves. Dans leur approche, chaque production écrite est transformée en un PFNet, pour ensuite être comparée au réseau PFNet construit à partir du texte d'un expert de contenu. Plus spécifiquement, la stratégie mise en œuvre est la suivante. Un expert de contenu détermine une liste de concepts et de synonymes. Puis, le logiciel ALA-Reader relève ces termes dans les textes et détermine toutes les paires de concepts qui apparaissent dans les mêmes phrases. Enfin, les cooccurrences des paires de concepts sont analysées au moyen d'un logiciel d'analyse de grappes (*cluster analysis*) nommé PCKNOT (Schvaneveldt, 1990) de sorte à produire une représentation simplifiée du texte, le PFNet. Ces réseaux servent enfin à la notation des productions écrites des élèves. Dans leur court article, les auteurs portent leur attention sur l'analyse des propriétés psychométriques des productions écrites, et non sur les performances de leur stratégie. Ainsi, ils donnent très peu d'information sur le rendement du logiciel.

Valerio et ses collaborateurs (Valerio & Leake, 2006; Valerio, Leake & Cañas, 2012) ont développé une application pour assister la synthèse, la compréhension et le référencement de documents. D'un point de vue technique, leur système est plus sophistiqué que celui de Clariana et Koul (2004). Il procède à la reconnaissance de termes composés de plusieurs mots, à la résolution d'anaphores et à l'identification de relations de synonymie. La détection de structures syntaxiques du type «nom, verbe, nom» sert ensuite à reconnaître les expressions de liaison. Enfin, même s'ils explorent de façon informelle certaines méthodes de synthèse des concepts et relations (p. ex., choisir les n concepts les plus fréquents), Valerio et ses collaborateurs n'intègrent aucune méthode dans leur solution finale. Malgré l'absence d'information au sujet de la taille des textes, tout laisse

croire que ceux-ci sont suffisamment courts et que les cartes générées ne comptent pas au-delà de 30 concepts, c'est-à-dire le nombre maximal acceptable de concepts considéré par ces chercheurs.

Dans le domaine des sciences informatiques, Zouaq et Nkambou (2008) ont développé un logiciel de production de cartes conceptuelles, appelée *TEXCOMON*, dans le but de constituer une ontologie informatique. L'analyse morphosyntaxique en grammaire de dépendance est au cœur de leur approche. Ils s'en servent pour identifier des termes composés de plusieurs mots et pour reconnaître les relations de type «prédicat, argument». Pour ce faire, l'analyseur syntaxique repère les mots qui modifient les noms (p. ex., l'adjectif «sémantique» modifie le nom «web» dans le terme «web sémantique») et les triplets du type «sujet, verbe, complément d'objet», couramment appelés des patrons SVO. Le logiciel *TEXCOMON* identifie également divers patrons lexico-syntaxiques (Hearst, 1992) qui témoignent de relations d'hyponymie et d'hyponymie entre les concepts. Enfin, ce logiciel ne requiert pas de procédé de synthèse de l'information étant donné qu'il a pour but de construire une ontologie informatique, et non des cartes conceptuelles. Une ontologie informatique contient généralement beaucoup plus de termes et de relations qu'une carte conceptuelle, et les corpus analysés par *TEXCOMON* sont beaucoup plus volumineux. Sept corpus composés de 10 à 36 textes et de 728 à 1 578 phrases ont été étudiés lors de leur mise à l'essai.

Qasim et ses collaborateurs (2013) se sont inspirés des procédés de *TEXCOMON* pour la création d'une carte conceptuelle portant sur le thème des sciences informatiques. Ils emploient les mêmes stratégies pour reconnaître les concepts et les relations. Par contre, ils échafaudent une solution complexe combinant plusieurs méthodes statistiques pour faire la synthèse du contenu des textes. En comparaison aux trois études précédentes, leur tâche de synthèse est colossale : 65 textes sont analysés simultanément par le logiciel. Pour y arriver, ils recourent à une statistique de pondération de l'information pour déterminer les expressions de liaison les plus significatives et à une technique d'analyse de grappes nommée la propagation d'affinité pour réduire l'inventaire de concepts et de relations. Il n'est pas possible d'évaluer le rendement effectif de ces procédés puisque les auteurs ne mentionnent pas le nombre de concepts et de relations qui ont été repérés dans la phase initiale ni combien ont été retenus dans la carte finale.

Synthèse

Même si elles partagent quelques points en commun, les quatre démarches de production de cartes conceptuelles ont des caractéristiques qui les distinguent. D'abord, chaque démarche adapte quelque peu la définition de la cartographie conceptuelle à son contexte d'utilisation. Dans les études décrites ci-dessus, les réseaux PFNet de l'étude de Clariana et Koul ne comportent pas de mots de relation, tandis que les réseaux générés lors des mises à l'essai de Zouaq et Nkambou (2008) ou encore de Qasim et ses collaborateurs (2013) comptent plus de 25 concepts. Aucune étude n'organise les concepts de manière hiérarchique. Les principes guidant la construction d'une carte conceptuelle dépendent de l'usage qui en est fait.

Deuxièmement, le degré d'automatisation varie d'un système à l'autre. Chacune des solutions discutées comporte des étapes manuelles. Dans Clariana et Koul (2004), c'est l'utilisateur qui fournit la liste de concepts et, dans trois des quatre systèmes, c'est l'humain qui a la responsabilité d'assembler la carte finale. Seulement Valerio et Leake (2006) ont eu recours aux fonctionnalités du logiciel CmapTools pour former une représentation graphique initiale de la carte conceptuelle. Vu que l'utilisateur intervient à une étape ou l'autre de la démarche, il apparaît plus juste de parler de production semi-automatisée d'une carte conceptuelle ou encore de production d'une carte conceptuelle assistée par ordinateur.

Troisièmement, les indicateurs de performance et les niveaux de tolérance ne sont pas les mêmes d'un logiciel à l'autre puisqu'ils ne répondent pas aux mêmes buts. Les programmes informatiques incorporant des techniques de traitement automatique du langage naturel sont inévitablement faillibles. Cela dit, il y a bien des situations où un traitement incomplet ou légèrement erroné d'un texte aura un effet négligeable sur le rendement final du logiciel. Par exemple, pour le référencement de textes, une représentation même partielle du texte est suffisante pour relier un texte à un ensemble de mots clés et pour distinguer les textes entre eux. À l'inverse, il y a d'autres situations plus exigeantes d'un point de vue technique, comme la notation de productions écrites, où il est attendu que le logiciel produise une représentation riche et juste de chaque texte écrit.

Dans le même ordre d'idées, le rendement du logiciel dépend aussi du nombre, de la qualité et de la pertinence des textes analysés. Toutes choses étant égales par ailleurs, le système restitue plus de concepts et de relations lorsque le corpus de textes est volumineux et lorsque les techniques de traitement du langage sont adaptées au style, au format et à la qualité de l'écriture. Les propositions inventoriées sont quant à elles plus significatives lorsque les textes touchent spécifiquement le thème ciblé. En somme, autant la définition de la cartographie conceptuelle que la stratégie d'extraction doivent s'aligner avec les objectifs poursuivis et le contexte d'utilisation du logiciel. La prochaine section s'attarde au contexte dans lequel le présent logiciel a été développé.

Contexte

La production automatisée d'une carte conceptuelle à partir de textes est rarement une finalité en soi ; elle sert d'autres buts. Dans le cadre de ce projet, elle s'inscrit dans une recherche plus large qui étudie les propriétés psychométriques des données collectées au moyen de cartes conceptuelles construites par des élèves en vue de l'évaluation de leurs apprentissages. Plus précisément, deux tâches du type «élaborer une carte conceptuelle à partir d'une liste préétablie de concepts» et conformes avec la définition de la cartographie conceptuelle du CRESST ont été soumises à des élèves de quatrième secondaire en science et technologie.

L'usage d'une carte de référence est une pratique courante dans plusieurs modalités de correction et de notation de cartes conceptuelles. Elle peut servir d'outil d'appoint ou encore de clé de correction et permet d'obtenir d'ailleurs une plus grande fidélité interjuges (McClure, Sonak & Suen, 1999).

L'approche la plus courante pour élaborer une carte de référence est de demander à un ou des experts de construire une carte conceptuelle. Cette approche a pourtant ses lacunes : les représentations externes des connaissances varient d'un expert à l'autre (Acton, Johnson & Goldsmith, 1994) et ne se réconcilient pas nécessairement à celles de novices (Chi, Feltovich & Glaser, 1981 ; Chi, Glaser & Farr, 1988 ; Glaser, 1991), surtout si cette représentation n'est pas construite dans un but «didactique». Il est possible de recourir à un panel d'experts pour réduire les singularités, mais cette stratégie est irréaliste dans plusieurs contextes. La produc-

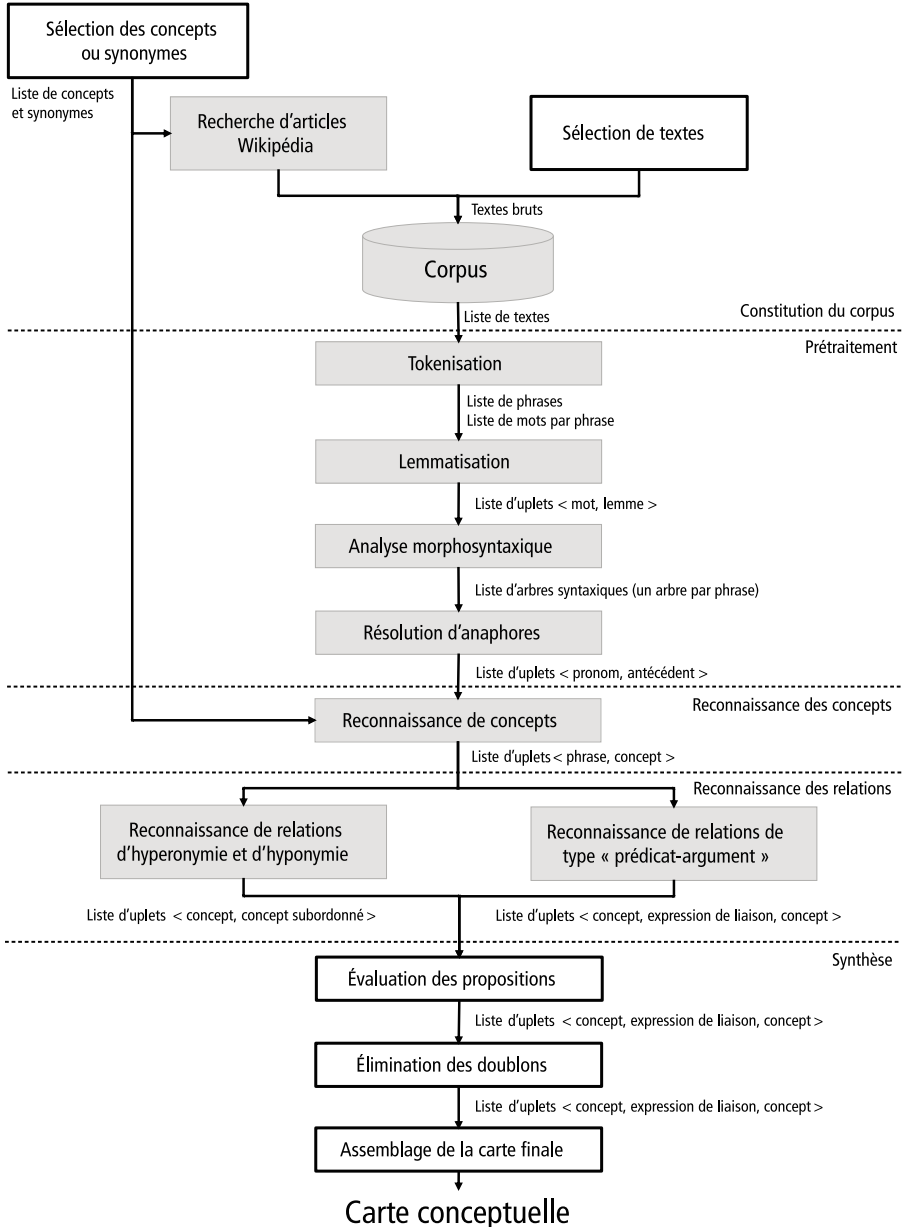
tion d'une carte conceptuelle à partir de textes tente d'offrir une solution de rechange aux approches fondées sur la construction directe d'une carte de référence par des experts de contenu. Comme il sera décrit ci-dessous, la carte conceptuelle sera construite principalement à partir de textes provenant de manuels scolaires. Ces textes sont composés et révisés par des experts du domaine de connaissances, mais, surtout, ils sont rédigés dans un but « didactique » ; leur contenu découle d'un consensus entre des experts et est plus près de ce qui est attendu de l'élève.

Objectif de la recherche

L'objectif du présent article est donc de mettre en place une stratégie d'extraction de propositions assistée par ordinateur en vue de la production d'une carte conceptuelle et dans le but éventuel de créer une carte de référence pouvant servir d'outil de correction des cartes produites par des élèves. Pour y arriver, il est attendu que les propositions extraites soient représentatives du domaine de connaissances, valides et adaptées au niveau de scolarité des élèves. La démarche de production d'une carte conceptuelle, telle que décrite à la section suivante, a donc été développée en conformité avec cet objectif et le type de tâches de cartographie conceptuelle soumises aux élèves.

Démarche

Cette section présente la démarche de production d'une carte conceptuelle à partir de textes (voir Figure 3). Une application informatique a été développée sous le langage de programmation Python afin d'automatiser les processus de sélection des textes, de prétraitement des textes, de reconnaissance des concepts et de reconnaissance des relations. Cette application est intitulée EPTAO et désigne l'extraction (E) de propositions (P) à partir de textes (T) et assistée (A) par ordinateur (O). Les prochaines sous-sections décrivent les composantes techniques de la démarche et d'EPTAO.



Note. Un encadré gris foncé désigne un processus automatisé et un encadré blanc désigne un processus manuel.

Figure 3. Démarche de production de la carte, fonctionnement d'EPTAO et modèle de données.

Constitution du corpus de textes

EPTAO offre deux mécanismes pour former un corpus : l'injection de textes choisis par une personne ou la recherche automatisée d'articles Wikipédia associés aux concepts fournis par l'utilisateur. Tous les documents sont formatés en fichiers de texte brut.

Prétraitement des textes

Plusieurs techniques de traitement automatique du langage naturel (TALN) contribuent à la préparation des documents. La tokenisation des textes est faite directement par l'analyseur morphosyntaxique. La lemmatisation, qui sert principalement à définir toutes les flexions morphologiques des concepts, se fait au moyen du *Lexique des formes fléchies du français* (Sagot, 2010). Ce lexique a en général une bonne couverture du français (plus de 530 000 entrées), mais certains mots spécialisés peuvent ne pas s'y trouver (p. ex., les mots « autotrophes » et « hétérotrophes »). Lorsque nécessaire, certains mots ont été ajoutés manuellement au lexique.

Chaque phrase est analysée deux fois par EPTAO : une première fois selon une grammaire en constituants, à l'aide du modèle pour le français de l'analyseur Stanford Parser (Green, De Marneffe, Bauer & Manning, 2011), et une seconde fois selon une grammaire de dépendance, à l'aide d'une adaptation francisée de l'analyseur MSTParser (Candito, Nivre, Denis & Anguiano, 2010). Les annotations de la grammaire en constituants sont utilisées dans le processus de résolution d'anaphores, et celles de la grammaire de dépendance font partie du processus de reconnaissance des relations.

EPTAO comporte également un module de résolution des pronoms de la troisième personne (il, elle, ils, elles, leur, leurs). Ce module est basé sur l'algorithme de Hobbs (1978), et plus spécifiquement, il exploite les relations syntaxiques de la grammaire en constituants et les caractéristiques morphologiques (genre et nombre) des mots pour déterminer l'antécédent qui correspond au pronom à l'examen. Cet algorithme purement linguistique est l'un des plus anciens, mais il offre des performances comparables aux techniques plus récentes et sophistiquées (Tetreault, 1999).

Reconnaissance des concepts

En conformité avec les tâches de cartographie conceptuelle soumises aux élèves, les concepts et leurs synonymes sont fournis à l'avance par un humain dans cette étude, ce qui simplifie grandement le processus d'iden-

tification des concepts. Le logiciel n'a qu'à reconnaître les mots dans les textes ou leurs flexions morphologiques; il n'a pas à construire lui-même une banque de termes candidats.

Au cours de la mise au point du logiciel, il a été constaté que certains concepts sont précédés de locutions telles que «la majorité de» ou «une partie de». Dans une expression comme «la majorité des rayons solaires», c'est donc le nom collectif (majorité) et non le concept (rayons solaires) qui est à la tête du groupe nominal. Pour éviter des embûches lors de la reconnaissance des relations, le logiciel a été modifié afin qu'il reconnaisse autant l'expression «la majorité des rayons solaires» que le concept seul «rayons solaires». Cet ajustement est en fait dérivé de la méthode du plus court parcours, décrite dans la prochaine sous-section.

Reconnaissance des relations

Dans une tâche du type «élaborer une carte à partir d'une liste de concepts préétablie», les élèves ont le libre choix de formuler les expressions de liaison. Le logiciel extrait donc les expressions de liaison le plus naturellement possible, à savoir comment elles sont exprimées dans les textes. Pour y arriver, EPTAO intègre les deux mécanismes proposés par Zouaq et Nkambou (2008), à savoir les patrons lexico-syntaxiques de Hearst pour la reconnaissance des relations d'hyponymie et d'hyperonymie ainsi que les patrons SVO pour la reconnaissance des relations de type «prédicat, argument».

Dans un premier temps, les sept patrons de Hearst ont été adaptés au français et deux patrons supplémentaires ont été ajoutés à la suite d'une analyse préliminaire du corpus. Le tableau 1 dresse la liste de ces patrons, donne leur formulation en expression rationnelle (en anglais, *regular expression*) et les exemplifie. Il est à noter que les patrons n^{os} 1 et 2 sont rassemblés puisqu'ils sont équivalents en français.

Dans un second temps, la méthode SVO a été révisée pour satisfaire aux conditions de la présente étude. Les triplets «sujet, verbe, complément d'objet» ne désignent pas toutes les formes de patrons de dépendance (Stevenson & Greenwood, 2009). Dans la phrase «Un orchestre est un ensemble de musiciens instrumentistes», le sujet «orchestre» est associé au nom collectif et complément d'objet «ensemble». La méthode SVO ne constitue pas de relation entre les concepts d'«orchestre» et de «musiciens». La construction de chaînes de dépendance à l'aide de la méthode

Tableau 1
Adaptation des patrons lexico-syntactiques de Hearst (1992)

N°	Expression rationnelle	Exemples
1-2	NP ₀ tel que [NP ₁ , NP ₂ , ...] [et/ou] NP _n	Un félin tel que le chat ou le lynx
3	NP [, NP]* [,] ou [autre/autres] NP ₀	Le chat, le lynx ou autres félins
4	NP [, NP]* [,] et [autre/autres] NP ₀	Le chat, le lynx et autres félins
5	NP ₀ [,] [comprenant/y compris] [NP,]* [et/ou] NP	Les félins, y compris le chat, ou le lynx
6	NP ₀ [,] en particulier [NP,]* [et/ou] NP	Les félins, en particulier le chat et le lynx
7	NP ₁ [est une sorte de/est un type de] NP ₀	Le chat est une sorte de félin
8	NP ₀ suivant[e][s] : [NP,]* [et/ou] NP	Les félins suivants : le chat, le lynx
9	NP ₀ ([NP,]* [et/ou] NP)	Les félins (le chat, le lynx)

du plus court parcours entre deux nœuds (Bunescu & Mooney, 2005) est une technique permettant d'accroître la reconnaissance de relations. Dans l'exemple précédent, une chaîne de dépendance peut être construite entre «ensemble» et «musiciens» puisque ces termes font partie du même groupe nominal dans l'arbre syntaxique de la phrase. En combinant les résultats du patron SVO et de la chaîne de dépendance, il est donc possible de former la proposition suivante : « un orchestre *est un ensemble de* musiciens ». Des cas similaires se produisent lorsqu'un concept est dans le sous-arbre syntaxique du complément d'objet. Selon cette approche modifiée, l'expression de liaison est définie comme étant le plus court parcours entre deux concepts et qui traverse un verbe dans l'arbre syntaxique de la phrase.

Synthèse de la carte

La dernière étape de la démarche comporte la synthèse des propositions extraites par le logiciel et la création de la carte conceptuelle finale. Cette étape n'est pas traitée par EPTAO et est faite manuellement. La synthèse des propositions consiste essentiellement à regrouper les mots de relation qui portent le même sens et à les représenter par une seule proposition. Le tableau 2 donne un exemple de cette démarche. Dans le corpus sur l'effet de serre, les concepts «couche d'ozone» et «rayons ultraviolets» sont unis entre eux par six relations, regroupées en trois catégories. Les deux premières catégories dénotent respectivement le rôle protecteur

de la couche d’ozone et sa capacité à absorber les rayons ultraviolets. Elles désignent des propositions scientifiquement valides. La proposition de la troisième catégorie partage le même champ sémantique que celle de la première catégorie, mais elle est non valide d’un point de vue scientifique, du moins lorsqu’elle est isolée du reste de la phrase d’où elle est extraite¹.

Tableau 2
Exemple de catégorisation de mots de liaison

Mots de liaison
Sens 1
– protège la biosphère d’une partie du
– joue un rôle essentiel de protection (filtre) contre les
Sens 2
– a pour effet d’absorber la plus grande partie du
– est un filtre chimique qui absorbe les
– désigne la partie de l’atmosphère qui contient une concentration élevée de molécules d’ozone et qui absorbe une partie des
Sens 3
– offrait si peu de protection contre les

Une fois que la liste de propositions distinctes est générée, celles-ci peuvent être représentées graphiquement. Cette opération a été réalisée par le premier auteur, et ce, sans privilégier un mode d’organisation spécifique.

Méthodologie

La mise à l’essai d’un logiciel a pour but de tester ses performances et celles de ses composantes dans diverses conditions. Dans ce qui suit, les performances d’EPTAO sont testées auprès de deux corpus de textes. Ces corpus correspondent aux thèmes des deux tâches de cartographie conceptuelle élaborées par un enseignant participant à l’étude plus vaste dans laquelle s’inscrit cette recherche.

Description des tâches

La première tâche porte sur les mécanismes de l'effet de serre et de phénomènes associés. En s'inspirant du manuel scolaire de son cours, l'enseignant a retenu 11 concepts pour la réalisation de cette tâche, chacun étant accompagné de 1 à 4 synonymes. La liste complète est exposée dans le tableau I en annexe.

La seconde tâche concerne les relations trophiques (c.-à-d. les relations alimentaires) entre les êtres vivants. L'enseignant a formé une liste comprenant 17 concepts pour cette activité (voir Tableau II en annexe).

Création des corpus

Un corpus de textes a été formé pour chacune des deux tâches. Les textes de base sont tirés de manuels scolaires en science et technologie de quatrième secondaire et sont conformes au plus récent curriculum. Seulement trois ouvrages ont été recensés : *Observatoire* (Cyr, Verreault & Forget, 2008), *Science-Tech au secondaire* (Barbeau et al., 2008) et *Synergie* (Chartré et al., 2008).

Seule l'information pertinente aux thèmes des tâches a été extraite de ces manuels. Le Programme de formation de l'école québécoise (MELS, 2007) de science et technologie en quatrième secondaire adopte une approche intégrative des apprentissages en les organisant autour de quatre grands thèmes (p. ex., les changements climatiques). Selon les activités pédagogiques proposées, les mêmes concepts et apprentissages peuvent être réinvestis à travers plusieurs thèmes. Cela fait que certaines informations sont éparpillées et parfois répétées à plusieurs endroits dans le manuel scolaire. Pour les fins de cette étude, toutes les données pertinentes ont été rapatriées dans un même document.

La recherche d'articles Wikipédia se veut quant à elle un moyen d'enrichir le corpus et, ultimement, les résultats de la production d'une carte conceptuelle. Cette recherche a ajouté 17 textes à la première tâche et 15 textes à la seconde tâche. Le tableau 3 décrit quelques caractéristiques des deux corpus soumis au logiciel. Le nombre de documents et le nombre de mots sont beaucoup plus élevés pour les articles Wikipédia que pour les textes tirés des manuels scolaires.

Tableau 3
Description des corpus des deux mises à l'essai

Corpus	N ^{bre} de documents	N ^{bre} de mots		
		Min	Max	Total
Effet de serre				
Manuels scolaires	3	760	2 251	4 474
Articles Wikipédia	17	170	5 227	30 814
Relations trophiques				
Manuels scolaires	3	869	1 682	3 695
Articles Wikipédia	15	140	5 687	36 112

Stratégies d'évaluation du rendement du logiciel

L'évaluation du rendement du logiciel est une étape importante dans le cycle de développement d'une stratégie de production d'une carte conceptuelle (Villalon & Calvo, 2008). Les stratégies d'évaluation varient d'une étude à l'autre. Qasim et ses collaborateurs (2013) évaluent les performances des sous-composantes (p. ex., la résolution d'anaphores, l'identification des concepts) de leur système, alors que Villalon et Calvo (2008) suggèrent de comparer quantitativement la carte extraite de textes avec une carte produite par un panel d'experts de contenu. Clariana et Koul (2004) évaluent quant à eux les performances prédictives de leur système en établissant dans quelle mesure les scores dérivés des cartes des élèves (en les comparant à la carte de l'expert) sont équivalents aux scores décernés par des correcteurs. Un système doit idéalement être scruté à divers degrés de granularité. Une analyse très fine des performances sert à établir les pistes d'amélioration potentielles, tandis qu'une analyse à l'échelle macroscopique donne un aperçu plus juste du potentiel du programme dans son contexte d'utilisation. Les différentes stratégies d'évaluation mises en œuvre dans cette étude sont détaillées dans les sous-sections suivantes.

Stratégie d'évaluation du processus de reconnaissance des relations

Dans le but d'effectuer une analyse fine des performances du logiciel, tous les textes ont été annotés de façon manuelle par le premier auteur de cet article, également formé en enseignement au secondaire en science et technologie. La codification des textes a été réalisée en deux étapes. Dans un premier temps, toutes les anaphores, pronominales ou non, ont été

signalées, puis remplacées par leur antécédent lorsqu'il exprimait un concept de la liste initiale. Ensuite, tous les concepts et les relations valides ont été codifiés et isolés à l'intérieur des documents. À certaines occasions, il était crucial de tenir compte du contexte de la phrase pour juger de l'exactitude d'une proposition potentielle. Par exemple, il n'aurait pas été juste d'isoler la proposition « *Le dioxyde de carbone est le plus abondant des gaz à effet de serre* » de la phrase « *Le dioxyde de carbone est le plus abondant des gaz à effet de serre produits par les activités humaines* ». Une proposition valide est donc une proposition qui, prise isolément, est valide d'un point de vue scientifique.

Les annotations manuelles ont ensuite servi de critère pour l'évaluation du processus de reconnaissance des concepts et relations. Dans le cadre des mises à l'essai, deux statistiques, soit le taux de restitution et le taux de précision, sont calculées pour évaluer les performances du programme. Ces statistiques sont définies comme suit:

[Formule mathématique 1]

$$\text{Restitution} = \frac{\text{Nombre de propositions valides reconnues par le programme}}{\text{Nombre de propositions reconnues par l'expert}}$$

[Formule mathématique 2]

$$\text{Précision} = \frac{\text{Nombre de propositions valides reconnues par le programme}}{\text{Nombre de propositions reconnues par le programme}}$$

Les valeurs de ces deux statistiques se situent entre 0 et 1. Un taux de restitution de 1 indique une adéquation parfaite entre les propositions restituées par le système et celles reconnues par l'expert. Un taux de précision de 1 indique que toutes les entités reconnues par le système sont valides.

Stratégie d'évaluation du processus de remplacement des anaphores pronominales

EPTAO comporte un module de substitution des anaphores pronominales de la troisième personne basé sur l'algorithme de Hobbs (voir la sous-section concernant les mécanismes de prétraitement des textes). Dans le but d'évaluer le rendement réel et attendu de cet algorithme, trois variantes du taux de restitution sont rapportées dans les résultats. Les deux premières variantes servent à comparer les taux de restitution des propositions avant et après la résolution d'anaphores. Pour la troisième

variante, une correction est appliquée au calcul du taux de restitution comme si tous les pronoms personnels de la troisième personne avaient été remplacés correctement par le système. Cette dernière variante permet ainsi de mesurer les gains potentiels dans le cas où le module aurait fonctionné parfaitement.

Stratégie d'évaluation des sources d'erreurs

À certaines occasions, le logiciel commet des erreurs et suggère des relations qui ne font pas partie du texte. Une analyse préliminaire des faux positifs a permis de les catégoriser en quatre sources d'erreurs : 1) le remplacement d'un pronom par un mauvais antécédent, 2) l'équivocité des concepts et de leurs synonymes, 3) l'assemblage ambigu d'un patron SVO et d'une chaîne en dépendance, ou 4) une simplification abusive de la phrase.

Stratégie d'évaluation du processus de la synthèse des propositions

La sélection de textes est une étape importante de la démarche proposée. Un examen approfondi de l'inventaire des propositions extraites par EPTAO offre alors la possibilité d'évaluer la pertinence des textes choisis au début de la démarche. En particulier, deux indicateurs sont étudiés : la répétition et la pertinence des propositions distinctes. Le premier peut être considéré comme un indicateur, certes imparfait, d'un consensus au sujet des connaissances importantes dans un domaine spécifique. Le second indique à quel degré le texte contient des informations qui sont proches ou éloignées du thème principal de la carte conceptuelle.

Stratégie d'analyse de la carte finale

Lors de la dernière étape de la démarche, les informations consignées dans l'inventaire des propositions distinctes sont assemblées sous forme d'une carte conceptuelle. Cette carte issue des propositions extraites par EPTAO peut alors être comparée à une carte de référence. Villalon et Calvo (2008) proposent de former une carte de référence à l'aide d'un panel d'experts. En conformité avec la démarche proposée dans cette étude, la carte de référence est tirée des manuels scolaires puisque ces documents sont jugés plus conformes aux attentes quant aux apprentissages réalisés par les élèves. Cette comparaison est faite sur la base d'indicateurs quantitatifs, comme le taux de restitution des propositions distinctes, et d'indicateurs qualitatifs, comme le respect des principes de construction d'une carte conceptuelle.

En résumé, ce cadre méthodologique propose d'étudier sous plusieurs angles le rendement du logiciel. La prochaine section présente les résultats de cette mise à l'essai.

Résultats

Les résultats de chaque stratégie d'analyse sont présentés séparément dans les sous-sections suivantes.

Évaluation du processus de reconnaissance des propositions

Le tableau 4 donne les taux de restitution et de précision pour les deux mises à l'essai. Globalement, les taux réels de restitution (39%) et de précision (66%) sont assez modestes pour la tâche sur l'effet de serre, tandis que les résultats sont plus encourageants pour la seconde tâche, avec des taux réels respectivement de 60% et 76%. Comme il sera expliqué plus loin, le thème de l'effet de serre est peut-être en lui-même une source de difficultés pour le processus d'extraction.

Les résultats varient selon la source d'information. Les manuels scolaires renferment un plus grand nombre de propositions, et ce, même s'ils comptent beaucoup moins de mots que les articles provenant de Wikipédia

Tableau 4
Taux de restitution et taux de précision de l'extraction des propositions des deux mises à l'essai

Source	N ^{bre} de propositions	Taux de restitution			Taux de précision
		Avant	Après (réel)	Corrigé	
Effet de serre					
Manuels scolaires	38	0,39	0,42	0,47	0,84
Articles Wikipédia	32	0,34	0,34	0,44	0,50
Total	70	0,37	0,39	0,46	0,66
Relations trophiques					
Manuels scolaires	52	0,58	0,62	0,65	0,89
Articles Wikipédia	39	0,56	0,59	0,69	0,64
Total	91	0,57	0,60	0,67	0,76

Note. Avant = avant la résolution des anaphores; après (réel) = après la résolution des anaphores. Ce taux correspond au taux réel de restitution; corrigé = comme si toutes les anaphores pronominales de la troisième personne avaient été remplacées correctement par le logiciel.

(voir Tableau 3). Les taux de restitution des deux tâches sont relativement stables entre les textes des manuels scolaires et les articles Wikipédia, mais le taux de précision des textes provenant des manuels scolaires est toujours plus élevé que celui des articles Wikipédia. Le contenu, le style d'écriture et l'organisation des articles ne sont pas nécessairement uniformes, ce qui occasionne certaines difficultés lors de leur traitement.

Évaluation du processus de résolution d'anaphores

Toujours selon les résultats du tableau 4, l'algorithme de Hobbs contribue très peu à l'extraction de propositions. D'une part, le système n'a pas réussi à remplacer tous les pronoms de la troisième personne par leurs antécédents. Un seul pronom a été remplacé lors de la première tâche, tandis que seulement trois pronoms l'ont été dans le corpus de la seconde tâche. Il reste un grand nombre de pronoms qui n'a pas été traité par le système. Comme l'indique la différence entre le taux corrigé et le taux après résolution, des gains additionnels de 7% auraient pu être enregistrés si tous ces pronoms avaient été traités. Les valeurs modérées des taux corrigés (entre 0,46 et 0,67) indiquent également que la prise en charge d'autres formes d'anaphores (p. ex., les pronoms démonstratifs) pourrait accroître grandement le taux de restitution.

Évaluation des sources d'erreurs

Comme nous l'avions anticipé, le logiciel fournit des propositions qui sont fausses. Le tableau 5 donne la fréquence d'apparition des quatre sources d'erreurs pour les propositions faussement identifiées. Les deux corpus ont été affectés par différentes sources d'erreurs et ce sont les articles Wikipédia qui ont mené au plus grand nombre d'erreurs. Le corpus sur l'effet de serre a mené à 9 erreurs de «simplification abusive de la phrase». Ces erreurs se produisent lorsqu'une relation binaire est extraite d'une phrase qui comporte en fait une relation ternaire. Le corpus sur les relations trophiques est plutôt marqué d'erreurs d'équivocité des concepts et synonymes. À 10 occasions, le concept identifié dans une proposition ne correspond pas exactement au concept de la liste de départ. Cela se produit entre autres lorsque les termes «champignon» ou «bactérie» pris isolément ne désignent pas «champignon et bactérie».

Alors que les erreurs de simplification abusive de la phrase émanent de la structure des phrases, les erreurs d'équivocité sont issues d'une liste de concepts et de synonymes ambigus. C'est pourquoi les secondes sont beaucoup plus faciles à corriger que les premières. Il suffit de réviser la liste de concepts et synonymes. Quant aux erreurs de simplification abusive, il faut probablement revoir le programme informatique afin qu'il puisse analyser efficacement les relations ternaires.

Tableau 5
Fréquence des erreurs d'extraction dans les deux corpus

Corpus/Documents	Remplacement d'un pronom par un mauvais antécédent	Équivocité des concepts et synonymes	Assemblage ambigu d'un patron SVO et d'une chaîne en dépendance	Simplification abusive de la phrase	Total
Effet de serre					
Manuels scolaires	0	0	0	3	3
Articles Wikipédia	0	2	3	6	11
Total	0	2	3	9	14
Relations trophiques					
Manuels scolaires	0	0	3	1	4
Articles Wikipédia	1	10	2	0	13
Total	1	10	5	1	17

Évaluation de la redondance et de la pertinence des idées

La dernière étape de la stratégie d'extraction consiste à faire une synthèse des propositions identifiées. Au cours de cette étape, un expert de contenu, dans ce cas-ci le premier auteur, regroupe les propositions extraites par le logiciel en idées communes, puis il détermine si ces idées sont pertinentes d'un point de vue didactique, c'est-à-dire à l'égard de la tâche et du niveau de scolarité des élèves. Deux propositions représentent une même idée si elles partagent le même sens d'un point de vue scientifique (voir Tableau 2 pour un exemple). Dans le but d'évaluer le rendement du logiciel, ces deux opérations ont été effectuées avec les propositions extraites par le logiciel ainsi qu'avec les propositions extraites manuellement des textes. Le tableau 6 présente la fréquence d'apparition des propositions non pertinentes ou pertinentes dans le corpus et dans la carte de référence. Il est important de noter que le nombre d'idées provenant du corpus

entier ne correspond pas au nombre d'idées identifiées dans les manuels scolaires et dans les articles Wikipédia. Ces deux ensembles de textes ne sont pas mutuellement exclusifs. Une même idée est comptabilisée une seule fois dans le corpus entier, même si elle est répétée dans les deux ensembles pris séparément. Par exemple, l'idée «les gaz à effet de serre emprisonnent une partie des rayons infrarouges» est répétée quatre fois dans les manuels scolaires et deux fois dans les articles Wikipédia.

Comme l'illustre le tableau 6, seul le corpus sur les relations trophiques a généré des propositions qui sont scientifiquement valides, mais non pertinentes d'un point de vue didactique. Un exemple d'une telle proposition est «les animaux *sont l'un des types d'eucaryotes à s'être développés sur un mode multicellulaire comme les végétaux*». Il n'y a rien de faux dans cette proposition, mais celle-ci est trop précise pour des élèves de quatrième secondaire en science et technologie. Le contenu des ouvrages didactiques est en effet plus homogène et plus ciblé que celui des articles Wikipédia. Ces articles contiennent parfois des propositions trop pointues pour le groupe d'élèves ciblé par la tâche. Dans un contexte scolaire, l'enseignant devra vraisemblablement modifier ou éliminer les propositions qui ne correspondent pas aux attentes du programme de formation.

Tableau 6
Pertinence et fréquence d'apparition des propositions distinctes dans le corpus et dans les inventaires de propositions extraites par EPTAO

Sources	Nbre de propositions distinctes	Non pertinente		Pertinente	
		1 fois	2 fois ou plus	1 fois	2 fois ou plus
Effet de serre					
Corpus entier	43	0	0	31	12
Manuels scolaires	21	0	0	10	11
Articles Wikipédia	27	0	0	21	6
Inventaire de propositions	17	0	0	11	6
Relations trophiques					
Corpus entier	49	4	0	24	21
Manuels scolaires	31	0	0	12	19
Articles Wikipédia	27	2	0	12	13
Inventaire de propositions	33	5	0	15	15

Analyse de la carte finale

Une fois que la révision de l'inventaire des idées communes est terminée, les propositions sont assemblées en une carte conceptuelle. Par souci d'économie d'espace, seules les cartes portant sur les relations trophiques sont présentées et décrites dans ce qui suit. Pour plusieurs raisons, certaines ayant déjà été évoquées, cette tâche a mené à de meilleurs résultats. Les figures 4 et 5 illustrent respectivement la carte de référence, obtenue à l'aide des annotations manuelles des textes, et la carte avec les propositions extraites par EPTAO. Les concepts et relations ont été organisés de manière à faire ressortir les ressemblances et les différences entre ces deux cartes. En prenant la carte de référence comme critère de comparaison, le logiciel a extrait 24 des 31 idées pertinentes (77%) contenues dans les textes didactiques sur les relations trophiques². Un nœud (Animaux) de la liste initiale de concepts est totalement absent des deux cartes. Même si ce concept ou ses synonymes étaient présents dans les documents, aucune relation valide n'a été découverte.

Un code de couleurs a été utilisé dans les figures 4 et 5 pour discerner les différences entre les deux cartes. Plus précisément, les expressions de liaison encadrées en gris foncé désignent celles qui se trouvent dans une seule des deux cartes conceptuelles. Dans la carte de la figure 4, ces relations uniques découlent de toutes de situations où un pronom n'est pas remplacé par son antécédent. Dans la carte de la figure 5, les relations distinctes sont toutes issues des articles Wikipédia.

Selon Villalon et Calvo (2008), les cartes produites automatiquement devraient ressembler autant que possible aux cartes élaborées à la main. Les cartes des figures 4 et 5 s'éloignent un peu des principes généralement admis de construction d'une carte conceptuelle. Premièrement, les deux cartes renferment plus d'une proposition pour certaines paires de concepts. Deuxièmement, elles renferment des relations redondantes. Par exemple, les propositions «les consommateurs *peuvent se nourrir de* décomposeurs» et «les champignons et bactéries *peuvent être mangés par des* consommateurs» signifient essentiellement la même chose. Troisièmement, la construction de propositions en combinant les chaînes de dépendance et les patrons SVO peut générer des expressions de liaison plus longues qu'à l'accoutumée (voir les exemples dans Novak & Cañas, 2008). Les répercussions de ces observations sont discutées plus longuement dans la discussion, présentée dans la prochaine section.

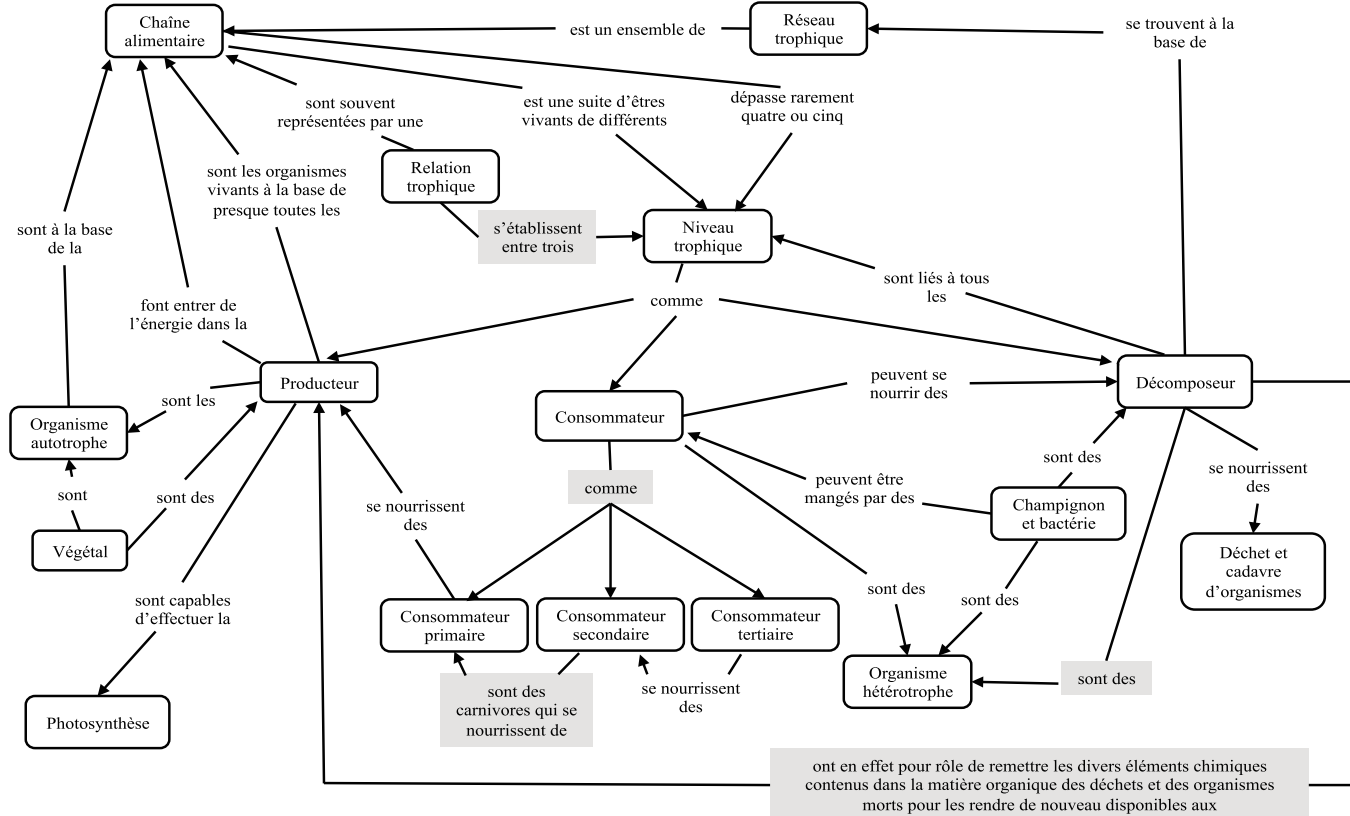


Figure 4. Carte conceptuelle de référence sur le thème des relations trophiques

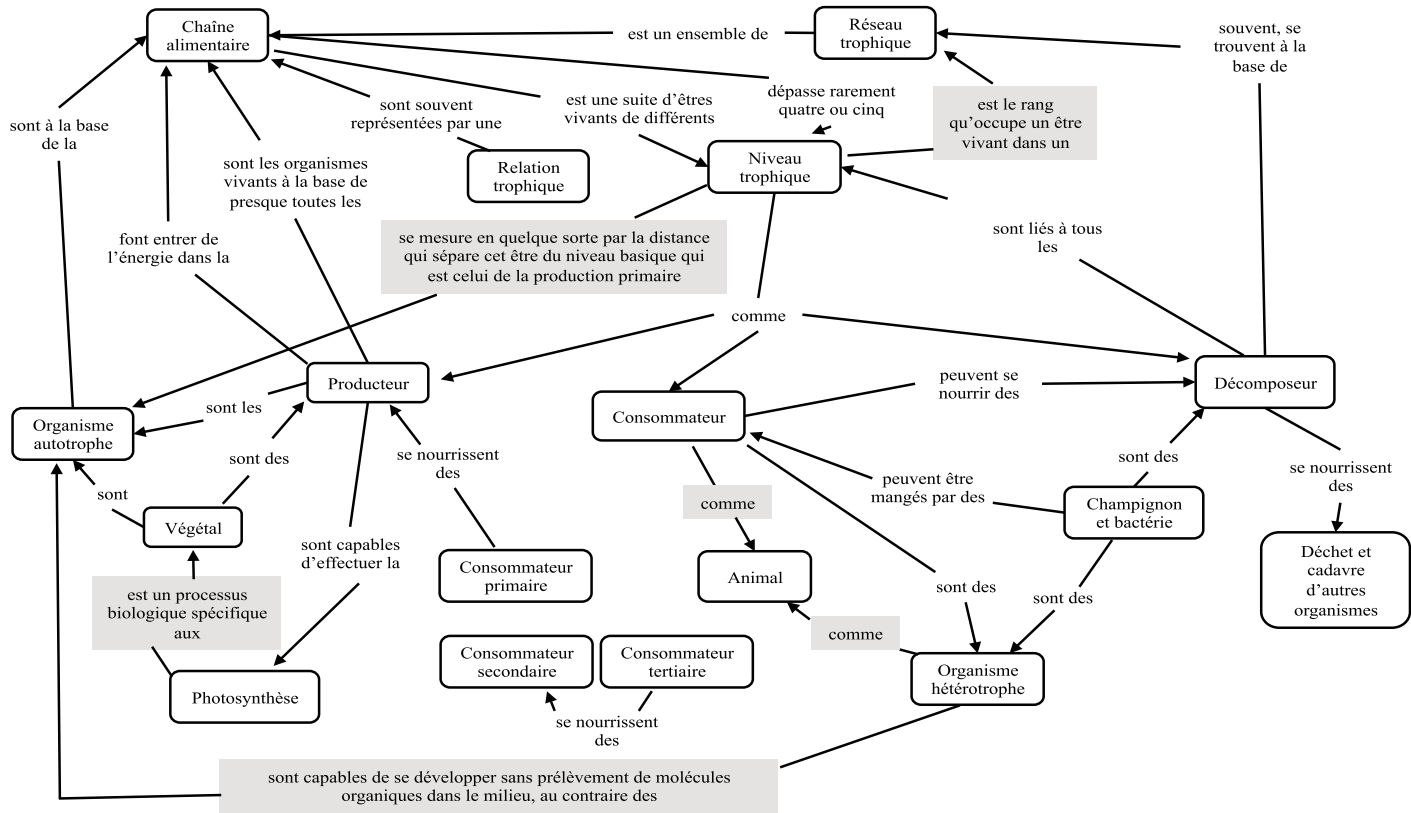


Figure 5. Carte produite sur la base des propositions extraites par EPTAO

Discussion

La section précédente a décrit les performances globales et spécifiques d'EPTAO et de ses composantes. Globalement, les taux réels de restitution (entre 39 et 60%) et de précision (entre 66 et 76%) sont relativement modestes. Cela dit, il est difficile de situer les performances d'EPTAO par rapport à celles d'autres recherches similaires parce qu'aucune d'entre elles n'a, autant que nous sachions, évalué de manière systématique le taux de restitution et de précision des relations extraites³. En comparaison au taux de restitution des propositions, le taux de restitution des propositions distinctes (idées) est plus encourageant. Les mêmes idées se répètent d'un texte à l'autre de telle sorte qu'au moment de la synthèse des propositions, une grande partie des idées contenues dans les textes se retrouvent dans la carte conceptuelle finale.

Dans une démarche de recherche et développement, l'évaluation de la performance des composantes spécifiques du logiciel est une source d'information capitale pour établir les améliorations potentielles de la démarche. Du point de vue technique, il y a lieu de se questionner au sujet du faible rendement de l'algorithme de Hobbs (2 à 3%) pour remplacer les pronoms de la troisième personne. En comparaison, Qasim et ses collaborateurs (2013) ont obtenu un rendement de 9% avec l'algorithme de Lappin et Leass (1994). De plus, une analyse plus détaillée des textes laisse aussi croire qu'il faudrait se pencher sur la résolution d'autres types d'anaphores, dont celles composées de pronoms démonstratifs (celui, celle, etc.) ou les anaphores nominales (p. ex., «[...] des rayons solaires. Ces rayons [...]»). La résolution d'anaphores est une composante essentielle de la reconnaissance des concepts, même s'ils sont fournis par l'utilisateur.

Certains choix méthodologiques peuvent aussi être revus à la lumière des résultats de la recherche. Entre autres, la contribution des articles Wikipédia est plutôt mitigée. Certes, ces articles ont généré quelques propositions pertinentes qui ne se trouvent pas dans les manuels scolaires. En revanche, ils sont à l'origine d'un plus grand nombre de propositions non pertinentes à l'égard du niveau de scolarité ciblé et ils occasionnent plus d'erreurs lors de la reconnaissance des relations. Les articles Wikipédia sont des textes de nature encyclopédique et ne sont pas rédigés dans une visée didactique comme les textes tirés des manuels scolaires. Le style

d'écriture, la longueur des phrases et le contenu HTML sont tous des facteurs susceptibles d'influencer le processus d'extraction des propositions contenues dans ces articles.

La présente étude soulève aussi quelques pistes de réflexion d'ordre pédagogique. La première piste concerne un principe de construction d'une carte conceptuelle généralement admis et peu contesté, soit celui de formuler une et une seule relation à partir de deux concepts. L'analyse systématique du contenu des textes montre pourtant qu'il est possible de formuler plus d'une idée indépendante à partir d'une même paire de concepts. D'un point de vue pédagogique, il ne semble donc pas y avoir de raisons d'empêcher un élève d'exprimer deux relations ou plus entre les deux mêmes concepts.

La seconde piste de réflexion concerne la représentation d'une même idée à partir de paires de concepts différentes. Dans certaines modalités d'évaluation de cartes conceptuelles (p.ex., voir Ruiz-Primo, Schultz et al., 2001), seules les propositions qui se trouvent conjointement dans la carte de l'élève et dans la carte de référence sont jugées valides. Qu'arrive-t-il lorsque l'élève choisit une paire de concepts différente pour exprimer une proposition qui se trouve dans la carte de référence?

La troisième piste de réflexion porte sur la longueur et le contenu des mots de relation. Normalement, une expression de liaison comporte quelques mots et est composée d'un verbe. Parmi les expressions extraites par le logiciel, certaines sont très longues (plus de 30 mots) et sont composées de compléments d'objet dont certains comportent des groupes nominaux. Ces groupes du nom auraient, eux-mêmes, pu faire partie de la liste de concepts. Puisque cette liste est établie à l'avance et figée, EPTAO cherche tous les parcours raisonnables (dans les arbres syntaxiques) entre deux concepts contenus dans la même phrase. Dans la pratique, de telles relations, si jugées trop longues, peuvent tout simplement être ignorées ou reformulées par un élève ou un enseignant. Cela dit, cette approche offre une autre option pour représenter les relations ternaires (p. ex., les producteurs *font entrer de l'énergie dans* la chaîne alimentaire). Elle évite de les reformuler comme une suite de deux relations binaires (p. ex., les producteurs *font entrer de l'énergie dans la chaîne alimentaire*) et réduit le risque d'interpréter ces deux relations binaires de manière indépendante.

Une dernière piste de réflexion, plus large, concerne la pertinence de la cartographie conceptuelle pour représenter certains domaines de connaissances. La cartographie conceptuelle est assujettie à une « syntaxe » particulière qui ne s'adapte peut-être pas à toutes les situations. La représentation de connaissances sous forme de relations binaires est un premier exemple des limites de cet outil. L'ajout d'un quantificateur venant modifier le concept est un second exemple. La carte conceptuelle, du moins dans sa forme courante, n'est pas adaptée à la représentation de relations dynamiques entre les concepts (Safayeni, Derbentseva & Cañas, 2005), ce qui explique pourquoi certains ajustements ont dû être apportés à EPTAO pour extraire des propositions dans les textes concernant l'effet de serre.

Enfin, le rendement d'un logiciel de production automatisée de cartes conceptuelles n'est pas seulement tributaire du logiciel, mais aussi des domaines de connaissances représentés dans les textes qui sont analysés par le système. Le but de l'activité, le thème et les contenus d'apprentissage sont ainsi tous des facteurs à considérer lors du choix d'un outil de représentation des connaissances.

Conclusion

Cet article a présenté la conception et la mise à l'essai d'une démarche de production semi-automatisée d'une carte conceptuelle à partir de documents textuels. Le développement d'une telle stratégie a pour but d'offrir une solution de rechange à la méthode de constitution d'une carte conceptuelle par un expert ou un panel d'experts de contenu.

Les résultats de cette étude sont prometteurs. Le logiciel EPTAO, conçu et testé dans cette recherche, a démontré qu'il est capable d'extraire les connaissances contenues dans les textes et ainsi d'assister un humain dans la création d'une carte conceptuelle. Pour l'instant, les propositions extraites par EPTAO n'aboutissent peut-être pas à une représentation suffisante du domaine de connaissances pour des usages contraignants, comme la comparaison de cartes produites par des élèves et d'une carte de référence en vue de l'évaluation des apprentissages. Cependant, la carte finale construite à partir des propositions extraites par le logiciel offre un bon point de départ pour divers autres usages (la rétroaction, la synthèse des apprentissages, etc.).

Cela dit, il y a encore plusieurs avenues à explorer pour améliorer le rendement d'EPTAO. D'un point de vue technique, quelques composantes du logiciel peuvent être perfectionnées ou encore d'autres composantes de la démarche peuvent être automatisées. D'un point de vue méthodologique, il y a certainement un intérêt à mettre en place de nouvelles stratégies pour augmenter le corpus, car EPTAO n'a quasiment pas de limites quant au nombre de textes qu'il peut traiter. Ces améliorations pourront à leur tour faciliter le travail des utilisateurs d'EPTAO et accroître potentiellement les usages d'un tel logiciel.

Réception: 1^{er} avril 2016

Version finale: 2 mai 2017

Acceptation: 5 mai 2017

NOTES

1. En effet, cette proposition est un exemple d'une simplification abusive de la phrase suivante: «Le 9 octobre 2003, au sud du Chili, la couche d'ozone, qui avait perdu 50% de son épaisseur habituelle, offrait si peu de protection contre les ultraviolets que, ce jour-là, il suffisait de passer cinq minutes dehors sans protection pour attraper un coup de soleil.»
2. En guise de comparaison, 11 des 21 idées (52%) contenues dans les textes traitant de l'effet de serre ont été identifiées correctement.
3. Qasim et ses collaborateurs (2013), par exemple, évaluent le taux de restitution et de précision des concepts, mais pas des relations.

Annexe

Tableau I
Concepts et synonymes acceptés pour le corpus sur l'effet de serre

Concepts	Synonymes acceptés
Chaleur du Soleil	Énergie émise par le Soleil, énergie solaire
Couche d'ozone	Ozone stratosphérique
Dioxyde de carbone	CO ₂
Énergie thermique de la Terre	Chaleur de la Terre
Gaz à effet de serre	GES
Lumière du Soleil	Rayonnement solaire, rayons du Soleil, rayon solaire, lumière solaire
Processus naturel	
Rayon infrarouge	Rayonnement infrarouge, infrarouge, IR
Rayon ultraviolet	Rayonnement ultraviolet, ultraviolet, UV
Surface de la Terre	Surface terrestre, sol
Température de la Terre	Température moyenne de la Terre, température sur la Terre

Tableau II
Concepts et synonymes acceptés pour le corpus sur les relations trophiques

Concepts	Synonymes acceptés
Animal	
Chaîne alimentaire	Chaîne trophique
Champignon et bactérie	Champignon, bactérie
Consommateur	Consommateur
Consommateur primaire	Consommateur de premier ordre
Consommateur secondaire	Consommateur de second ordre
Consommateur tertiaire	Consommateur de troisième ordre
Déchet et cadavre d'autres organismes	Détritus
Décomposeur	
Organisme autotrophe	Autotrophe
Organisme hétérotrophe	Hétérotrophe
Niveau trophique	
Photosynthèse	
Producteur	
Relation trophique	
Réseau trophique	Réseau alimentaire
Végétal	Plante

RÉFÉRENCES

- Acton, W. H., Johnson, P. J., & Goldsmith, T. E. (1994). Structural knowledge assessment: Comparison of referent structures. *Journal of Educational Psychology*, 86(2), 303-311. doi: 10.1037/0022-0663.86.2.303
- Barbeau, C., Groulx, D., Schepper, C., Valiquette, C., Lévesque, N. et Dignard, C. (2008). *Science-tech au secondaire: un regard sur l'environnement – Manuel de l'élève*. Laval, Québec: Grand Duc.
- Bunescu, R. C., & Mooney, R. J. (2005, October). *A shortest path dependency kernel for relation extraction*. Communication presented at the Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Vancouver, BC. doi: 10.3115/1220575.1220666
- Cañas, A. J., Coffey, J. W., Carnot, M. J., Feltoich, P. J., Hoffman, R. R., Feltoich, J., & Novak, J. D. (2003). *A summary of literature pertaining to the use of concept mapping techniques and technologies for education performance support*. Pensacola, FL: The Institute for Human and Machine Cognition. Retrieved from http://cmap.bradercomm.net:8001/rid=1186721000045_1491728413_5664/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf
- Candito, M., Nivre, J., Denis, P., & Anguiano, E. H. (2010, August). *Benchmarking of statistical dependency parsers for French*. Communication presented at the 23rd International Conference on Computational Linguistics, Beijing, China. Retrieved from <http://dl.acm.org/citation.cfm?id=1944579>
- Chartré, C., Levert, I., Robert, O.-L., Caillier, M., Racine, C., Athlan, É., ... Tougas, M.-H. (2008). *Synergie: manuel de l'élève – 2^e cycle du secondaire*. Montréal: Graficor-Chenelière Éducation.
- Chen, N.-S., Kinshuk, Wei, C.-W., & Chen, H.-J. (2008). Mining e-learning domain concept map from academic articles. *Computers & Education*, 50(3), 1009-1021. doi: 10.1016/j.compedu.2006.10.001
- Chi, M. T. H., Feltoich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121-152. doi: 10.1207/s15516709cog0502_2
- Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- Clariana, R. B., & Koul, R. (2004, September). *A computer-based approach for translating text into concept map-like representations*. Communication presented at the 1st International Conference on Concept Mapping, Navara, Spain. Retrieved from <http://cmc.ihmc.us/papers/cmc2004-045.Pdf>
- Cyr, M.-D., Verreault, J.-S. et Forget, D. (2008). *Observatoire: l'environnement – Manuel de l'élève*. Saint-Laurent, Québec: ERPI.
- Glaser, R. (1991). Expertise and assessment. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and cognition* (pp. 17-30). Englewood Cliffs, NJ: Prentice-Hall.

- Green, S., De Marneffe, M.-C., Bauer, J., & Manning, C. D. (2011, July). *Multiword expression identification with tree substitution grammars: A parsing tour de force with French*. Communication presented at the Conference on Empirical Methods in Natural Language Processing, Edinburgh, United Kingdom. Retrieved from <http://dl.acm.org/citation.cfm?id=2145516>
- Hearst, M. A. (1992, August). *Automatic acquisition of hyponyms from large text corpora*. Communication presented at the 14th Conference on Computational Linguistics (Vol. II), Nantes, France. doi: 10.3115/992133.992154
- Herl, H. E., O'Neil, H. F., Chung, G. K. W. K., & Schacter, J. (1999). Reliability and validity of a computer-based knowledge mapping system to measure content understanding. *Computers in Human Behavior, 15*, 315-333. doi: 10.1016/S0747-5632(99)00026-6
- Hobbs, J. R. (1978). Resolving pronoun references. *Lingua, 44*(4), 311-338. doi: 10.1016/0024-3841(78)90006-2
- Klein, D. C. D., Chung, G. K. W. K., Osmundson, E., & Herl, H. E. (2002). *Examining the validity of knowledge mapping as a measure of elementary students' scientific understanding*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- Koul, R., Clariana, R. B., & Salehi, R. (2005). Comparing several human and computer-based methods for scoring concept maps and essays. *Journal of Educational Computing Research, 32*(3), 227-239. doi: 10.2190/5X9Y-0ETN-213U-8FV7
- Lappin, S., & Leass, H. J. (1994). An algorithm for pronominal anaphora resolution. *Computational Linguistics, 20*(4), 535-561. Retrieved from <http://acl-arc.comp.nus.edu.sg/archives/acl-arc-090501d3/data/pdf/anthology-PDF/J/J94/J94-4002.pdf>
- Marchand, C. et d'Ivernois, J.-F. (2004). Les cartes conceptuelles dans les formations en santé. *Pédagogie médicale, 5*(4), 230-240. doi: 10.1051/pmed:2004031
- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity and logistical practicality. *Journal of Research in Science Teaching, 36*(4), 475-492. doi: 10.1002/(SICI)1098-2736(199904)36:4<475::AID-TEA5>3.0.CO;2-O
- Ministère de l'Éducation, du Loisir et du Sport (MELS). (2007). *Programme de formation de l'école québécoise: secondaire, deuxième cycle*. Québec, gouvernement du Québec. Repéré à www.mels.gouv.qc.ca/sections/programmeFormation/secondaire2/index.asp?page=introduction
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413-448. doi: 10.3102/00346543076003413
- Nielsen, R. D., Ward, W., Martin, J. H., & Palmer, M. (2008, June). *Extracting a representation from text for semantic analysis*. Communication presented at the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies, Columbus, OH. Retrieved from <http://dl.acm.org/citation.cfm?id=1557760>
- Novak, J. D. (1977). An alternative to piagetian psychology for science and mathematics education. *Science Education, 61*(4), 453-477. doi: 10.1002/sci.3730610403
- Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them: IHMC CmapTools*. Pensacola, FL: Institute for Human and Machine Cognition.

- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge: Cambridge University Press.
- Novak, J. D., Gowin, D. B., & Johansen, G. T. (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education*, 67(5), 625-645. doi: 10.1002/sce.3730670511
- Qasim, I., Jeong, J.-W., Heu, J.-U., & Lee, D.-H. (2013). Concept map construction from text documents using affinity propagation. *Journal of Information Science*, 39(6), 719-736. doi: 10.1177/0165551513494645
- Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569-600. doi: 10.1002/(SICI)1098-2736(199608)33:6<569::AID-TEA1>3.0.CO;2-M
- Ruiz-Primo, M. A., Shavelson, R. J., Li, M., & Schultz, S. E. (2001). On the validity of cognitive interpretations of scores from alternative concept-mapping techniques. *Educational Assessment*, 7(2), 99-141. doi: 10.1207/S15326977EA0702_2
- Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching*, 38(2), 260-278. doi: 10.1002/1098-2736(200102)38:2<260::AID-TEA1005>3.0.CO;2-F
- Safayeni, F., Derbentseva, N., & Cañas, A. J. (2005). A theoretical note on concepts and the need for cyclic concept maps. *Journal of Research in Science Teaching*, 42(7), 741-766. doi: 10.1002/tea.20074
- Sagot, B. (2010, May). *The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French*. Communication presented at the 7th International Conference on Language Resources and Evaluation, Valletta, Malta. Retrieved from <https://hal.inria.fr/inria-00521242>
- Schvaneveldt, R. W. (1990). *Pathfinder associative networks: Studies in knowledge organization*. New York, NY: Ablex Publishing.
- Sowa, J. F. (2000). *Knowledge representation: Logical, philosophical, and computational foundations*. Pacific Grove: Brooks/Cole.
- Stevenson, M., & Greenwood, M. A. (2009). Dependency pattern models for information extraction. *Research on Language and Computation*, 7(1), 13-39. doi: 10.1007/s11168-009-9061-2
- Tetreault, J. R. (1999, June). *Analysis of syntax-based pronoun resolution methods*. Communication presented at the 37th Annual Meeting of the Association for Computational Linguistics on Computational Linguistics, College Park, MD. Retrieved from <http://dl.acm.org/citation.cfm?id=1034688>
- Villalon, J. J., & Calvo, R. A. (2008, December). *Concept map mining: A definition and a framework for its evaluation*. Communication presented at the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Sydney, Australia. doi: 10.1109/WIAT.2008.387
- Valerio, A., & Leake, D. (2006, September). *Jump-starting concept map construction with knowledge extracted from documents*. Communication presented at the 2nd International Conference on Concept Mapping, San José, Costa Rica. Retrieved from <http://cmc.ihmc.us/cmc2006Papers/cmc2006-p181.pdf>

- Valerio, A., Leake, D. B., & Cañas, A. J. (2012, September). *Using automatically generated concept maps for document understanding: A human subjects experiment*. Communication presented at the 5th International Conference on Concept Mapping, Valletta, Malta. Retrieved from <http://cmc.ihmc.us/cmc2012papers/cmc2012-p102.pdf>
- Yin, Y., & Shavelson, R. J. (2008). Application of generalizability theory to concept map assessment research. *Applied Measurement in Education, 21*(3), 273-291. doi: 10.1080/08957340802161840
- Yin, Y., Vanidez, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation and use. *Journal of Research in Science Teaching, 42*(2), 166-184. doi: 10.1002/tea.20049
- Zouaq, A., & Nkambou, R. (2008). Building domain ontologies from text for educational purposes. *IEEE Transactions on Learning Technologies, 1*(1), 49-62. doi: 10.1109/TLT.2008.12