



REVUE ALGERIENNE DES SCIENCES

Section -A-

- Sciences de la Nature et de la Vie

- Sciences techniques

<http://univ-eltarf.dz/fr/>

Données liées dans l'éducation : un examen des contributions dans les revues SSCI

Chaouki Chemama^{a,*}, Hafed Zarzour^b, Toufik Sari^c

^a Département d'Informatique, Université Chadli Bendjedid El Tarf, El-Tarf, 36000 Algérie.

^b LIM Research, Département d'Informatique, Université de Souk Ahras, Souk Ahras, 41000 Algérie.

^c LabGED laboratory, Université Badji Mokhtar Annaba Annaba, 23000, Algérie.

Informations	Résumé
Mots clés : Données liées Web de données Apprentissage amélioré Données liées dans l'éducation *Correspondance : che_chaouki@yahoo.fr Reçu le 31 mai 2018, Révisé le 07 juin 2018, Accepté le 12 juin 2018	Actuellement la plupart des établissements d'enseignement publient certains de leurs ensembles de données clés en tant que données liées (linked data). Cependant, l'utilisation des technologies de données liées à des fins éducatives est toujours difficile et peu claire. Dans cet article, nous menons un sondage en examinant l'utilisation des données liées dans le contexte éducatif. En effet, nous adoptons deux critères principaux afin de sélectionner et d'inclure tout article publié dans notre revue. (1) le document doit être publié dans l'une des revues reconnues par l'Institut d'Information Scientifique(IIS) et répertorié dans le Social Science Citation Index (SSCI), dans le domaine de la «technologie d'apprentissage amélioré». (2) le document doit clairement montrer comment consommer, produire, ou à la fois consommer et publier des données liées dans le contexte d'apprentissage. Les résultats de l'enquête ont montré l'importance d'utiliser les données liées comme une puissante technologie éducative.

1. Introduction

Récemment, un certain nombre d'établissements d'enseignement publient certains de leurs ensembles de données clés en tant que données liées [1]. Selon [2], la technologie de données liées correspond à un ensemble de bonnes pratiques pour la publication et l'interconnexion de données structurées sur le Web. Ces pratiques peuvent aider les organisations et les institutions à exposer leurs propres ensembles de données via le Web en fournissant des liens vers des éléments de données et des connaissances pour trouver facilement des informations lors de la recherche. En d'autres termes, la technologie de données liées vise à relier les données de différentes sources et fournisseurs dans le but de produire plus de connaissances sur n'importe quel sujet.

Plusieurs technologies du Web sémantique du W3C sont utilisées pour faciliter la description et la publication de données liées telles que RDF (Resource Description Framework) [3] et SPARQL [4]. RDF est basé sur l'idée de décrire une ressource sur le Web comme triplement, un ensemble de trois entités qui expriment une déclaration en termes de sujet, de prédicat et d'objet. SPARQL est un langage de requête sémantique utilisé pour manipuler et interroger des données stockées nativement sous la

forme de RDF. RDF et SPARQL peuvent fournir des outils puissants pour rendre les données réutilisables et partageables via différents systèmes sur le Web. [5]. En outre, deux autres composants intéressants de la pile technologique pour les données liées sont les URI et les URI HTTP [6, 7]. Les URI sont utilisés pour identifier les ressources données, tandis que les URI HTTP sont requis pour interpréter et déferer ces ressources.

Comme les données liées ont le potentiel de créer des ponts entre les silos de données disponibles, elles ont conduit à la création de jeux de données inter-reliés géants appelés Web of Data, qui améliorent principalement la découverte, l'accessibilité et la réutilisabilité des données sur le Web. [8]. Il augmente rapidement pour permettre l'échange au sein de petites organisations ou à l'échelle mondiale.

Dans les dernières années, de nombreuses recherches antérieures ont rapporté l'efficacité de l'utilisation de données liées dans le développement de méthodes éducatives efficaces pour le domaine de l'apprentissage. [10, 11]. Par exemple, le projet LinkedUp [12] était un projet européen qui a adopté les normes de données liées dans l'éducation en publiant des données d'apprentissage existantes sous une forme

structurée en tant que données liées. Le LinkedUp a fourni à la communauté intéressée par Linked Data pour l'éducation un catalogue de données éducatives pour stocker et mettre à disposition toutes sortes de ressources pertinentes pour l'éducation. mEducator [13] est un autre projet qui se concentre sur des référentiels de données conformes aux exigences de données liées pour résoudre les problèmes d'interopérabilité dans le domaine de l'apprentissage ainsi que pour stocker des contenus éducatifs. Ce projet a utilisé différentes données ouvertes disponibles sur le Web pour permettre la spécialisation des ressources d'apprentissage à découvrir, intégrer, publier et enrichir à travers les institutions académiques.

En effet, l'utilisation de données liées dans le contexte éducatif offre plusieurs avantages, notamment [14] (i) la portabilité et la réutilisation des données car plusieurs fournisseurs de données liées à l'éducation sont disponibles sur le Web. (ii) L'enrichissement sémantique des données liées à l'éducation. (iii) Accessibilité des données liées à l'éducation via des formats standards. (iv) Personnalisation et contextualisation de l'apprentissage via Données liées. (v) Découvrabilité des ressources d'apprentissage pertinentes supplémentaires en examinant les liens entre les données.

Afin de clarifier l'importance de l'adoption de données liées dans le but d'apprentissage en tant que technologie éducative, cet article vise à examiner l'état de l'art des travaux de recherche existants rapportés dans la littérature. Seulement huit revues reconnues dans le domaine de la technologie d'apprentissage amélioré publiées par SOCIAL SCIENCE CITATION INDEX de 2006 à 2017 ont été prises en considération.

Le reste de l'article est organisé comme suit: la section 2 prérèglera la méthodologie suivie. La section 3 donne un aperçu des articles sélectionnés. La section 4 discute des résultats. Enfin la section 5 conclut ce document.

2. Méthodologie

Cette étude vise à examiner l'utilisation des données liées dans le contexte éducatif dans lequel nous suivons les principes d'une revue systématique de la littérature [15]. La revue systématique est considérée comme l'une des techniques les plus largement utilisées pour mener une revue de littérature approfondie en rapport avec des études existantes sur un sujet donné. Pour ce faire, nous avons retenu deux critères principaux afin de sélectionner et d'inclure tout article publié dans notre étude : (1) le document doit être publié dans l'une des revues reconnues par l'Institut d'Information Scientifique (IIS) et répertorié dans le Social Science Citation Index (SSCI), dans le domaine «Apprentissage par la technologie». (2) le document doit montrer clairement comment consommer, produire, ou à la fois consommer et

publier des données liées dans le contexte d'apprentissage.

Par conséquent, huit revues reconnues dans le domaine de la technologie d'apprentissage amélioré indexées et publiées par SOCIAL SCIENCE CITATION INDEX de 2006 à 2017 ont été sélectionnées pour mener cette enquête. Ces revues sont connues pour leur grande qualité dans la publication de recherches liées à l'utilisation de la technologie éducative.

- Computers & Education;
- Educational Technology & Society;
- Educational Technology Research and Development;
- Innovations in Education and Teaching International;
- British Journal of Educational Technology;
- Journal of Computer Assisted Learning;
- Interactive Learning Environments;
- IEEE Transactions on Learning Technologies.

Lors de la révision, nous avons pris en considération uniquement les articles et excluons tous les autres types de publication tels que les lettres et les livres. Les mots clés pertinents suivants ont été utilisés pour effectuer la recherche dans les bases de données associées aux revues sélectionnées: "Linked Data", "linked open data", "Web of data", and "Data Web". Parmi les 22 études récupérées, seulement 13 articles ont été examinés parce qu'ils ont conservé le deuxième critère défini ci-dessus. Comme il n'y a pas de publications dans les revues considérées entre 2006 et 2011, nous présenterons dans la section suivante les résultats allant de 2012 à 2017.

3. Aperçu des papiers sélectionnés

Dans [16], les auteurs ont développé une base de données liée à l'approche pour permettre la fédération des outils TIC de l'éducation et le partage des descriptions à leur sujet. Ils ont également décrit un jeu de données lié basé sur un certain vocabulaire pour importer automatiquement des descriptions d'outils TIC provenant de sources tierces. Dans [17], les auteurs ont proposé d'utiliser les catégories DBpedia comme données liées afin de réduire la taille des dépôts d'objets d'apprentissage en améliorant la classification des objets d'apprentissage. Dans [18], les auteurs ont développé deux outils en ligne basés sur les données liées: le premier outil appelé Anomation qui permet d'annoter sémantiquement les vidéos en important les vocabulaires de Linked Data Cloud, le second appelé SugarTube qui permet d'explorer sémantiquement les vidéos avec un réseau de données amélioré.

Dans [19], les auteurs ont proposé de résoudre le problème des spectateurs multiples de cours de danse en utilisant le stockage en ligne, et l'application de données liées comme un moyen puissant pour les vidéos sémantiquement liées et ajouter des connexions significatives de données. SEEKAT-WD [20] a été proposé en tant qu'infrastructure visant à tirer parti des Données Liées afin de pérenniser un outil pédagogique TIC dans le Web des Données selon une vision socio-sémantique. mEducator [21] a été développé pour offrir une gamme de ressources éducatives ouvertes pour le domaine médical, publiées via les principes de données liées. Comme l'ensemble de données TED était couramment utilisé par plusieurs applications éducatives, les auteurs de [22] ont exploré le site Web des conférences TED pour le rendre disponible en tant que données liées. Les auteurs de [23] ont proposé une approche pour intégrer les services Cloud à l'apprentissage électronique basé sur les systèmes via les données liées. Dans [24], les auteurs ont démontré l'utilité d'utiliser les principes des données liées pour produire et consommer des données structurées au sein des universités. Ils ont montré quelques exemples sur la façon de fournir des applications éducatives utiles en intégrant des données liées. Les auteurs de [25] ont réalisé une étude sur le référentiel Organique. Edunet, en tant que portail d'apprentissage en ligne dans un contexte agricole. Ils ont mené des expériences pour exposer les métadonnées d'Organique. Edunet en tant que données ouvertes liées et en relier leur contenu à plusieurs nuages de données liées pertinents. Dans [26], les auteurs ont développé un e-TextBooks hybride basé sur EPUB 3 qui utilise des logiciels serveur et client Linked Data Fragments pour enrichir sémantiquement les e-books.

Dans [27], nous avons proposé L2OD, qui permet aux étudiants d'enrichir leurs annotations avec des ressources pertinentes récupérées à partir de nuages de données ouverts liés. L2OD prend en charge deux modes d'annotation, un mode d'annotation privé pour les annotations individuelles de l'utilisateur et un mode partagé pour les annotations de tous les utilisateurs. Les résultats expérimentaux ont montré une amélioration significative des étudiants qui ont utilisé L2OD. De plus dans [28], nous avons proposé une approche d'annotation collaborative basée sur la technologie de données liées, appelée CAALDT pour améliorer les performances d'apprentissage des étudiants. Le CAALDT proposé offrait des possibilités d'apprentissage plus stimulantes qui permettaient aux élèves d'examiner les annotations de leurs pairs ou de leur enseignant et d'exposer leurs annotations en tant que données liées, contribuant ainsi à l'expansion des contenus d'apprentissage.

4. Résultats

4.1. Nombre d'articles publiés dans le domaine d'apprentissage

La figure 1 montre le nombre d'articles publiés par les revues considérées entre 2012 et 2017 sur l'utilisation des données liées dans le contexte éducatif. Il a été constaté que le nombre d'articles dans ce domaine a été augmenté en 2015 et 2017. En divisant les 6 années allant de 2012 à 2017 en deux périodes équivalentes de 3 ans chacune, nous pouvons constater que le nombre d'articles publiés dans les trois dernières années est environ 1,6 fois plus que pendant les trois premières années. Cela suggère que l'utilisation des données liées ne cesse de croître en tant que technologie éducative.

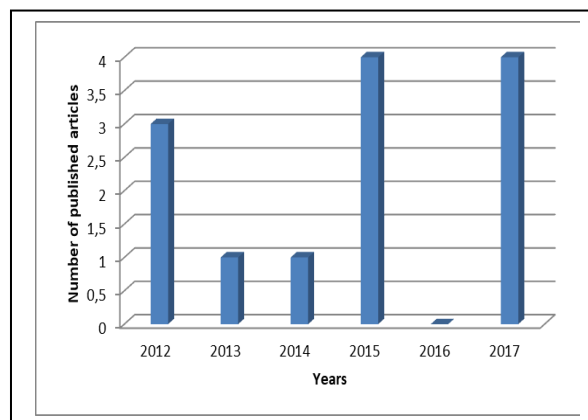


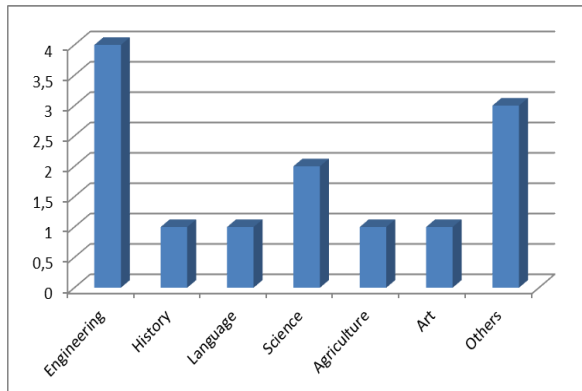
Fig.1 : Nombre d'articles publiés sur les données liées dans l'éducation de 2012 à 2017.

4.2. Classification selon des objectifs pédagogiques

Il existe trois types d'études utilisant les technologies de données liées à des fins éducatives. Le premier type d'études a tenté de consommer les ressources éducatives structurées en ligne existantes à des fins multiples, plus précisément, de créer des solutions innovantes dans l'apprentissage par l'interconnexion des matériaux d'apprentissage disponibles. Le deuxième type d'études était basé sur la création et la publication de ressources éducatives en tant que données liées à des fins d'échange et d'enrichissement des données d'apprentissage. Le troisième type d'études portait sur les principes des données liées pour traiter à la fois la consommation et la production de ressources éducatives.

La figure 2 illustre l'application des technologies de données liées selon ces trois catégories, y compris la consommation, la production et la consommation et la production. Il a été observé que 31% des études examinées utilisent les données liées comme consommateur, 8% comme producteur et 61% comme consommateur et producteur. Par exemple, les auteurs de [24] ont montré leurs expériences lors de la production et de l'exploitation de la technologie des données liées dans le contexte d'un établissement d'enseignement.

Fig.2 : Classification selon le type d'utilisation des données liées dans l'éducation.



4.3. Applications de données liées dans l'éducation

La figure 3 illustre la répartition des études dans sept domaines d'apprentissage, y compris l'ingénierie (y compris l'informatique), l'histoire, la langue, la science, l'art et d'autres. Il a été observé que de 2012 à 2017, il y avait deux tendances clairement de l'application des données liées dans l'éducation, l'ingénierie avec (4/13) et la science avec (2/13). Ces domaines d'apprentissage constituent ensemble près de la moitié de tous les articles publiés. D'un autre côté, le domaine d'apprentissage de l'histoire, de la langue et de l'art n'avait qu'une seule publication.

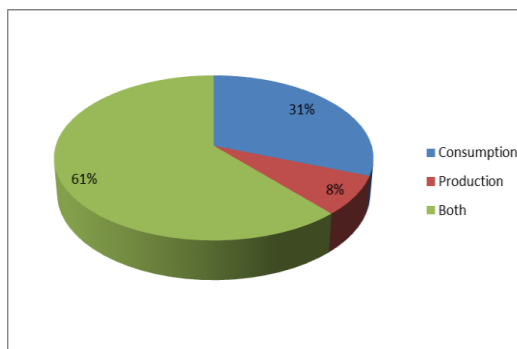


Fig.3 : Différentes applications de données liées dans l'éducation

4.4. Utilisations des jeux de données LD

Lorsqu'ils travaillent sur des données liées à des fins éducatives, les chercheurs ont utilisé différents types d'ensembles de données disponibles sur le réseau de données pour interconnecter leurs propres sources d'apprentissage. En effet, nous avons classé les ensembles de données référencés en quatre catégories: DBpedia (un ensemble de données extrait automatiquement de Wikipedia en tant que données liées), ensemble de données TED (un ensemble de métadonnées et de transcriptions automatiquement extraites de TED en tant que données liées), LD Cloud de tous les ensembles de données disponibles interconnectés selon les principes des données liées) et d'autres (y compris certains des ensembles de données fournis par certaines universités). La figure 4 montre la distribution des études en fonction de ces jeux de données LD. Il a été observé que la plupart des études (7/13) utilisaient LD Cloud comme sources

principales de publication ou de création et d'interconnexion de leurs ressources éducatives, suivies par DBpedia et TED ensemble de données avec 3 et 1 de l'ensemble des études, respectivement. Par exemple, dans [27], nous avons proposé d'utiliser DBpedia afin d'enrichir le contenu d'apprentissage des étudiants en trouvant et en liant des ressources de données liées connexes pour une annotation donnée. Il a également été constaté que 2 des 13 études utilisaient d'autres ensembles de données spécifiques tels que ceux fournis par l'Open University pour produire ou exposer les données en tant que données liées [24].

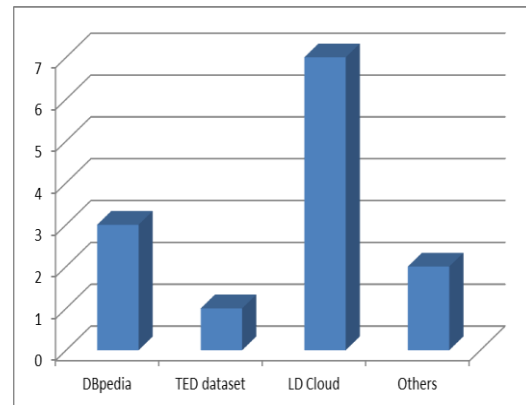


Fig.4 : Répartition des études en fonction des jeux de données LD utilisés

4.5. Principales contributions des pays

La figure 5 montre les principales contributions des pays aux données liées à des fins éducatives, conformément aux articles publiés dans ce domaine. Il a été observé que le sommet des articles de contribution faite par les auteurs espagnols et britanniques avec 4 publications chacun; suivi par les auteurs allemands et grecs avec (3) chacun; suivi d'auteurs algériens, italiens et belges avec (2) chacun; suivis par les auteurs indiens et mexicains avec une seule contribution chacun. Cependant, nos deux contributions [27] et [28] en tant qu'auteurs algériens ont été les recherches les plus récentes dans ce domaine, publiées en 2017.

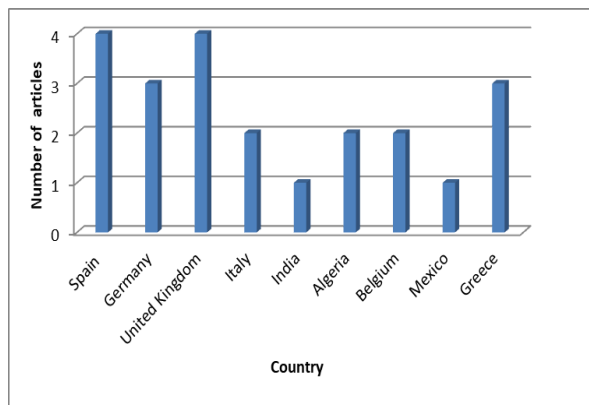


Fig.5 : Contributions des pays dans les données liées à des fins éducatives.

5. Conclusion

Dans cet article, nous avons examiné l'utilisation de la technologie des données liées à des fins d'apprentissage de 2012 à 2017, en fonction des contributions publiées dans huit revues SSCI bien reconnues. Deux critères ont été définis pour prendre en compte les articles publiés. Le premier critère consistait en la qualité des revues qui publiaient l'article, tandis que le second critère consistait à montrer clairement comment les données liées sont utilisées dans le contexte d'apprentissage. Il a été constaté que le nombre d'articles publiés dans ce domaine a été augmenté en 2015 et 2017; de plus, l'application des technologies de données liées à de multiples fins d'apprentissage variait de la consommation, de la production, de la consommation et de la production. En outre, les chercheurs qui ont contribué à l'apprentissage des données liées provenaient de différents pays ces dernières années. En outre, il a été constaté que LD Cloud, qui comprend de nombreux ensembles de données interconnectés, était la principale source de publication et d'utilisation des ressources d'apprentissage. Ces résultats impliquent l'importance de données liées à être adopté par toutes les universités dans leur apprentissage stratégique en tant que technologie éducative puissante.

Référence :

- [1] C. Keßler, M. d'Aquin, and S. Dietze, "Linked Data for science and education," *Semantic Web*, vol. 4, no. 1, 2013.
- [2] C. Bizer, T. Heath, and T. Berners-Lee, "Linked Data - The Story So Far," *International Journal on Semantic Web and Information Systems*, vol. 5, no. 3, pp. 1–22, 2009.
- [3] R. Cyganiak, D. Wood, M. Lanthaler, *RDF 1.1 concepts and abstract syntax*. W3c Recommendation, vol. 25, no. 02, 2014.
- [4] D. Calvanese, B. Cogrel, S. Komla-Ebri, R. Kontchakov, D. Lanti, M. Rezk, M. Rodriguez-Muro, and G. Xiao, "Ontop: Answering SPARQL queries over relational databases," *Semantic Web*, vol. 8, no. 3, pp. 471–487, 2017.
- [5] S. YUMUSAK, E. DOGDU, H. KODAZ, A. KAMILARIS, and P.-Y. VANDENBUSSCHE, "SpEnD: Linked Data SPARQL Endpoints Discovery Using Search Engines," *IEICE Transactions on Information and Systems*, vol. E100.D, no. 4, pp. 758–767, 2017.
- [6] X. SHI and L. YANG, "Privacy-Preserving Multi-Keyword Fuzzy Search over Outsourced Encrypted Cloud Data," *Computer Science and Technology*, vol. 12, no. 8, pp.1874-1884, 2017.
- [7] O. Yurur, C. H. Liu, Z. Sheng, V. C. M. Leung, W. Moreno, and K. K. Leung, "Context-Awareness for Mobile Sensing: A Survey and Future Directions," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 68–93, 2016.
- [8] V. Christophides, V. Efthymiou, and K. Stefanidis "Entity resolution for hidden Web data," *Synthesis Lectures on the Semantic Web*, vol. 5, no. 3, pp. 1-122, 2015.
- [9] M. Guy, "The Open Education Working Group: Bringing People, Projects and Data Together," *Lecture Notes in Computer Science*, pp. 166–187, 2016.
- [10] M. d' Aquin, "On the Use of Linked Open Data in Education: Current and Future Practices," *Lecture Notes in Computer Science*, pp. 3–15, 2016.
- [11] E. Daga, M. d'Aquin, A. Adamou, and S. Brown, "The Open University Linked Data-data. open. ac. uk," *Semantic Web*, vol. 7, no. 2, pp. 183-191, 2016.
- [12] A. Motzo, "The Linkedup Project: An Online Repository for Language Teachers," *Case Studies of Openness in the Language Classroom*, pp. 45–56, Sep. 2013.
- [13] S. T. Konstantinidis, L. Ioannidis, D. Spachos, C. Bratsas, and P. D. Bamidis, "mEducator 3.0: Combining Semantic and Social Web Approaches in Sharing and Retrieving Medical Education Resources," *2012 Seventh International Workshop on Semantic and Social Media Adaptation and Personalization*, Dec. 2012.
- [14] H. Zarzour and M. Sellami, "Effects of a linked data-based annotation approach on students' learning achievement and cognitive load," *Interactive Learning Environments*, pp. 1–10, Mar. 2018..
- [15] C. Cronin, "Doing your literature review: traditional and systematic techniques," *Evaluation & Research in Education*, vol. 24, no. 3, pp. 219–221, Sep. 2011.
- [16] A. Ruiz-Calleja, G. Vega-Gorgojo, J. I. Asensio-Pérez, M. L. Bote-Lorenzo, E. Gómez-Sánchez, and C. Alario-Hoyos, "A Linked Data approach for the discovery of educational ICT tools in the Web of Data," *Computers & Education*, vol. 59, no. 3, pp. 952–962, Nov. 2012.
- [17] M. Lama, J. C. Vidal, E. O. Garcia, A. Bugarin, and S. Barro, "Semantic Linking of a Learning Object Repository to DBpedia," *Educational Technology & Society*, vol. 15, no. 4, pp. 47-61.
- [18] Hong Qing Yu, C. Pedrinaci, S. Dietze, and J. Domingue, "Using Linked Data to Annotate and Search Educational Video Resources for Supporting Distance Learning," *IEEE Transactions on Learning Technologies*, vol. 5, no. 2, pp. 130–142, Apr. 2012
- [19] S. Morris, "Using Semantic Web technologies to support the teaching of dance," *British Journal of Educational Technology*, vol. 44, no. 6, pp. E213–E216, Oct. 2013.
- [20] A. Ruiz-Calleja, G. Vega-Gorgojo, J.I. Asensio-Pérez, E. Gómez-Sánchez, M. L. Bote-Lorenzo and C. Alario-Hoyos, "SEEK-AT-WD: A social-semantic infrastructure to sustain educational ICT tool descriptions in the Web of Data," *Journal of Educational Technology & Society*, vol. 17, no. 2, pp. 321-332, 2014.
- [21] S. Dietze, D. Taibi, H. Q. Yu, and N. Dovrolis, "A Linked Dataset of medical educational resources," *British Journal of Educational Technology*, vol. 46, no. 5, pp. 1123–1129, Apr. 2015.
- [22] D. Taibi, S. Chawla, S. Dietze, I. Marenzi, and B. Fetahu, "Exploring TED talks as Linked Data for

- education," *British Journal of Educational Technology*, vol. 46, no. 5, pp. 1092–1096, Apr. 2015.
- [23] G. Gutiérrez-Carreón, T. Daradoumis and J. Jorba, "Integrating learning services in the cloud: An approach that benefits both systems and learning," *Journal of Educational Technology & Society*, vol. 18, no.1, pp. 145-157, 2015.
- [24] F. Zablith, M. Fernandez, and M. Rowe, "Production and consumption of university Linked Data," *Interactive Learning Environments*, vol. 23, no. 1, pp. 55–78, Dec. 2015.
- [25] E. Rajabi, S. Sanchez-Alonso, M.-A. Sicilia, and N. Manouselis, "A linked and open dataset from a network of learning repositories on organic agriculture," *British Journal of Educational Technology*, vol. 48, no. 1, pp. 71–82, Sep. 2017.
- [26] H. Ghaem Sigarchian, S. Logghe, R. Verborgh, W. de Neve, F. Salliau, E. Mannens, R. Van de Walle, and D. Schuurman, "Hybrid e-TextBooks as comprehensive interactive learning environments," *Interactive Learning Environments*, pp. 1–20, 2017.
- [27] H. Zarzour and M. Sellami, "A Linked Data-based collaborative annotation system for increasing learning achievements," *Educational Technology Research and Development*, vol. 65, no. 2, pp. 381–397, 2016.
- [28] H. Zarzour and M. Sellami, "An investigation into whether learning performance can be improved by CAALDT," *Innovations in Education and Teaching International*, pp. 1–8, Feb. 2017.