



HAL
open science

Infrastructure décentralisée et sémantique pour l'apprentissage tout au long de la vie

Hala Skaf-Molli, Patricia Serrano-Alvarado, Sara El Hassad, Emmanuel
Desmontils, Pascal Molli

► **To cite this version:**

Hala Skaf-Molli, Patricia Serrano-Alvarado, Sara El Hassad, Emmanuel Desmontils, Pascal Molli. Infrastructure décentralisée et sémantique pour l'apprentissage tout au long de la vie. Atelier Web des Données (AWD) dans EGC, Jan 2019, Metz, France. hal-01990846

HAL Id: hal-01990846

<https://hal.science/hal-01990846v1>

Submitted on 23 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Infrastructure décentralisée et sémantique pour l'apprentissage tout au long de la vie

Hala Skaf-Molli* Patricia Serrano-Alvarado**
Sara El hassad**, Emmanuel Desmontils**, Pascal Molli**

LS2N – University of Nantes, France

* {Hala.Skaf}@univ-nantes.fr ** {Name.LastName@}univ-nantes.fr

Résumé. L'apprentissage tout au long de la vie joue un rôle fondamental pour le développement professionnel des personnes. Les environnements tels que les portfolios numériques ou les environnements d'apprentissages personnels permettent aux apprenants d'acquérir de l'autonomie, élément essentiel pour cet apprentissage. Cependant, les services proposés par ces environnements se limitent généralement à la simple collecte de faits ou de preuves à propos des apprentissages effectués. De plus, ces services ne permettent pas d'effectuer des activités collaboratives et sont très souvent peu pérennes. Il est donc nécessaire de disposer d'une infrastructure flexible d'apprentissage qui va au-delà de la simple collecte de faits en offrant une intégration continue des données d'apprentissage et qui permet une collaboration en confiance entre les apprenants. Dans ce document, nous montrons comment le Web sémantique pourra aider à créer une infrastructure d'apprentissage respectant ces critères.

1 Introduction

De nos jours, l'apprentissage se déroule durant toute la vie d'une personne. Il commence lors de la formation initiale, de l'école jusqu'à l'université, et se poursuit pendant toute sa carrière, avec les différents emplois qu'il peut occuper. L'autonomisation des apprenants permet à chacun de définir son propre parcours d'apprentissage, ce qui est une condition préalable pour l'autonomie. Les dimensions identifiées de l'apprentissage tout au long de la vie dans les processus d'apprentissage autonome sont : la capitalisation des expériences, la reconnaissance, la gestion des objectifs, la gestion de l'apprentissage personnel et la collaboration. Les différentes infrastructures personnelles d'apprentissage (Van Harmelen, 2006) et les portfolios électroniques permettent de soutenir l'apprentissage autonome. Cependant, ils devraient être fonctionnellement étendus et synchronisés pour rassembler et exploiter les informations pertinentes pour les différentes dimensions de l'apprentissage tout au long de la vie.

Il n'existe actuellement aucune infrastructure fournissant un environnement de confiance pour l'apprentissage à long terme, la capitalisation des données et la gestion de l'apprentissage personnel (El Mawas et al., 2017). Nous croyons que les technologies du Web sémantique profiteraient à l'apprentissage autonome tout au long de la vie

sous plusieurs aspects. Le modèle de données RDF améliorera l'échange de données d'apprentissage sur le Web. Les ontologies des apprenants (Yang et al., 2006; Rezgui et al., 2017; Evangelos Katis et Vassilakis, 2018) permettront une intégration fluide des données d'apprentissage personnelles provenant de sources de données multiples et hétérogènes. Le langage de requête SPARQL permettra d'interroger les données personnelles d'apprentissage et d'autres données grâce aux requêtes fédérées.

Nous croyons que l'apprentissage autonome tout au long de la vie nécessite une infrastructure flexible et décentralisée. La décentralisation permet à chaque apprenant de garder le contrôle de ses données et elle élimine la nécessité d'un tiers qui doit être de confiance. Chaque apprenant doit pouvoir gérer ses données personnelles comme il le souhaite. La construction d'une infrastructure en réseau pair-à-pair dans le contexte de l'éducation a été proposée par Nejdil et al. (2002). Toutefois, ces travaux ne considèrent pas le large éventail des données collectées, le contexte d'une formation continue et d'une collaboration en confiance.

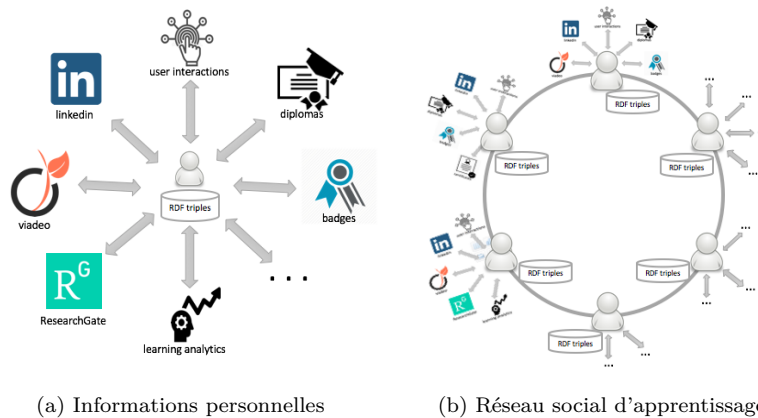


FIG. 1: Infrastructure d'apprentissage décentralisée et sémantique

2 Infrastructure décentralisée et sémantique

Nous envisageons une *infrastructure d'apprentissage décentralisée et sémantique* où chaque apprenant intègre ses données d'apprentissage personnelles dans un espace privé et collabore d'une manière personnalisée avec d'autres apprenants en partageant des données par le biais d'un ensemble de services fiables. La figure 1 montre une vision globale de l'infrastructure. L'espace privé présenté dans la figure 1a est composé de l'ensemble des informations pertinentes collectées par un apprenant. Chaque apprenant peut composer son espace privé différemment des autres et contrôler totalement les informations qui seront partagées avec les autres apprenants du fait de l'architecture décentralisée, comme illustré dans la figure 1b.

Nous proposons d'utiliser Solid¹ comme infrastructure de stockage. Chaque apprenant stocke ses données RDF dans une banque de données personnelle accessible sur le Web (Mansour et al., 2016). Les données sont collectées pendant les processus d'apprentissage, par exemple, les diplômes d'un apprenti récupérés de l'université et ses emplois collectés de LinkedIn et Viadeo. Pour gérer l'hétérogénéité des données collectées, nous utilisons les techniques d'intégration de données sémantiques. Pour la flexibilité, nous proposons deux services d'intégration de données : data warehouse et l'intégration virtuelle. Dans data warehouse, les données personnelles de chaque apprenant provenant des différentes sources de données sont chargées et matérialisées. Dans l'intégration virtuelle, les données restent dans les sources et elles sont accessibles à la demande au moment de la requête, par exemple via les API accessibles (Moreau et al., 2017). Elles seront transformées avec le schéma global par les wrappers et interrogées par les médiateurs. Nous proposons aussi un service d'intégration sémantique (Montoya et al., 2014) adapté à l'apprentissage tout au long de la vie, et qui permet d'intégrer facilement de nouvelles sources de données Web.

Les apprenants partageront leurs données via des requêtes fédérées combinant des données locales (personnelles) et des données distantes (des autres apprenants). Dans notre infrastructure, le traitement des requêtes fédérées nécessite que chaque apprenant expose un serveur, si possible peu coûteux (Verborgh et al., 2016), capable de traiter les requêtes SPARQL. Les services exécuteront des requêtes fédérées pour des finalités différentes, par exemple pour trouver *les cinq MOOC les plus suivis par les apprenants qui ont suivi le MOOC Web sémantique*. Pour trouver des données pertinentes, chaque apprenant tiendra une liste de serveurs fiables qui peuvent être organisés par un index guidé par des politiques d'usage. Les sources de données pertinentes peuvent également être découvertes d'une manière dynamique lors du traitement des requêtes grâce aux réseaux superposés sémantiques (Grall et al., 2018).

Avant de partager les données, l'apprenant associera différentes politiques d'usage (ou licences) à ses données, en fonction de ses préférences de confidentialité. Les politiques d'usage spécifient avec précision les conditions de réutilisation des données (autorisations, interdictions et obligations) et sont définies à l'aide de ODRL². Par exemple, un apprenant peut décider que ses données collectées sur LinkedIn peuvent être *lues* par les autres apprenants avec l'interdiction de les *distribuer*. Les résultats de la requête seront protégés par une politique d'usage conforme à toutes les politiques d'usage des données contribuant aux résultats de la requête Moreau et al. (2018). Toutes ces politiques permettront un contrôle d'usage fort dans notre infrastructure.

3 Conclusion

Nous commençons l'expérimentation de ces services dans le contexte du projet SEDELA (El Mawas et al., 2017). Des expériences préliminaires avec des étudiants du deuxième cycle en sciences de l'éducation viennent de commencer. L'objectif est d'explorer de nouvelles méthodes et de nouveaux outils pour renforcer l'autonomie, l'auto-apprentissage et le développement personnel.

1. <https://solid.mit.edu/>

2. www.w3.org/TR/odr1-vocab/

Remerciement Ce travail fait partie du projet multidisciplinaire SEDELA, financé par CominLabs.

Références

- El Mawas, N., J.-M. Gilliot, S. Garlatti, P. Serrano-Alvarado, H. Skaf-Molli, J. Eneau, G. Lameul, J.-F. Marchandise, et H. Pentecouteau (2017). Towards a Self-Regulated Learning in a Lifelong Learning Perspective. In *CSEDU*.
- Evangelos Katis, H. K. et K. Vassilakis (2018). Developing an Ontology for Curriculum & Syllabus. In *ESWC*. Poster paper.
- Grall, A., P. Molli, et H. Skaf-Molli (2018). SPARQL Query Execution in Networks of Web Browsers. In *ISWC*.
- Mansour, E., A. V. Sambra, S. Hawke, M. Zereba, S. Capadisli, A. Ghanem, A. Aboul-naga, et T. Berners-Lee (2016). A Demonstration of the Solid Platform for Social Web Applications. In *WWW*. Demo paper.
- Montoya, G., L. D. Ibáñez, H. Skaf-Molli, P. Molli, et M.-E. Vidal (2014). SemLAV : Local-As-View Mediation for SPARQL. *TLSDKCS*.
- Moreau, B., P. Serrano-Alvarado, et E. Desmontils (2018). CaLi : A Lattice-Based Model for License Classifications. In *BDA*.
- Moreau, B., P. Serrano-Alvarado, E. Desmontils, et D. Thoumas (2017). Querying non-RDF Datasets using Triple Patterns. In *ISWC*. Demo paper.
- Nejdl, W., B. Wolf, C. Qu, S. Decker, M. Sintek, A. Naeve, M. Nilsson, M. Palmér, et T. Risch (2002). EDUTELLA : a P2P Networking Infrastructure Based on RDF. In *WWW*.
- Rezgui, K., H. Mhiri, et K. Ghédira (2017). Ontology-based e-Portfolio Modeling for Supporting Lifelong Competency Assessment and Development. *Computer Science*.
- Van Harmelen, M. (2006). Personal Learning Environments. In *ICALT*.
- Verborgh, R., M. Vander Sande, O. Hartig, J. Van Herwegen, L. De Vocht, B. De Meester, G. Haesendonck, et P. Colpaert (2016). Triple Pattern Fragments : A low-cost Knowledge Graph Interface for the Web. *J. Web Sem.*
- Yang, S. J. et al. (2006). Context Aware Ubiquitous Learning Environments for Peer-to-Peer Collaborative Learning. *Educational Technology & Society*.

Summary

Lifelong learning is a crucial support for experienced workers in their professional development. Existing environments help learners to gain autonomy. However, they are usually restricted to collecting simple facts or evidence about personal learning. Moreover, they are not collaborative and lifelong available. In this vision paper, we show how the Semantic Web helps in building a flexible and decentralized lifelong learning infrastructure that enables trusted collaboration among learners.