



HAL
open science

Architecture d'une Application pour la Gestion Consolidée de Ressources Numériques

Daouda Sawadogo

► **To cite this version:**

Daouda Sawadogo. Architecture d'une Application pour la Gestion Consolidée de Ressources Numériques. Rencontre Jeunes Chercheurs en Environnements Informatiques pour l'Apprentissage Humain, Jun 2014, La Rochelle, France. hal-01086277

HAL Id: hal-01086277

<https://hal.science/hal-01086277>

Submitted on 23 Nov 2014

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.

Architecture d'une application pour la gestion consolidée de ressources numériques

Daouda Sawadogo

2^{ème} année de doctorat

Laboratoire L3i, Avenue M. Crépeau, 17000 La Rochelle, France

daouda.sawadogo@univ-lr.fr

Résumé

Dans cet article, nous nous intéressons à la pertinence et à la cohérence des ressources numériques dans le système d'information d'un utilisateur. L'accroissement de la production des données numériques ces deux dernières années soulève plusieurs problématiques quant à la gestion de ces données multi-sources et hétérogènes. Nos travaux consistent à proposer une architecture permettant de gérer des données de qualité, structurées, adaptées aux besoins de ses utilisateurs et validées par ceux-ci.

Dans la première partie, nous nous intéressons à la caractérisation des ressources numériques (pour un usage adaptatif).

Dans la deuxième, partie nous nous intéressons aux processus et mécanismes pour une gestion consolidée de ces ressources numériques dans un système d'information. L'objectif visé dans ces travaux, est la proposition d'une architecture pertinente d'une application interactive pour la gestion consolidée et assistée de ressources numériques.

Dans cet article, notre contribution concerne notre approche pour caractériser la pertinence d'une ressource numérique. Cette approche est basée sur le profil de l'utilisateur et les métadonnées associées aux ressources. Cette contribution permet aux utilisateurs d'avoir des environnements utilisant des ressources pertinentes en fonction de leurs profils. Notre expérimentation concerne une étude de cas d'une application pour les chercheurs. Cette application, leur permet de mieux caractériser leurs ressources.

Introduction

L'avènement des Technologies de l'Information et de la Communication (TIC) représente une véritable opportunité dans la diffusion des ressources et des connaissances pour tous. De nombreux résultats ont déjà été obtenus dans le cadre de l'étude d'Environnements Informatiques pour l'Apprentissage Humain. Ils visent essentiellement à faciliter la mise à disposition de contenus pédagogiques, accompagner l'apprenant en offrant des modalités d'interactions avec le formateur et valider les connaissances acquises. Le déploiement large et massif des Environnements Numériques de Travail (ENT) constitue une première génération d'outils d'aide à la formation découlant de

ces travaux. Le développement des technologies du multimédia conjugué à celui de l'Internet et la démocratisation du haut débit, rendent désormais possibles la formation à distance d'apprenants situés dans des classes virtuelles géographiquement distribuées. Dans ce contexte, nous nous intéressons aux systèmes de gestion de ressources numériques. L'usage de ces ressources est adapté par les connaissances que le système déduit des interactions, du profil et du parcours de l'utilisateur.

Contexte et Problématiques

D'après une étude de l'IDC (Gantz & Reinsel, 2012) et dans le rapport EGC-2013, nous constatons que « Depuis les quatre dernières années, le nombre de sources de données structurées rendues disponibles sur le Web est en croissance fulgurante aboutissant à un espace global de données de l'ordre de milliards d'assertions. » (PFIA, 2013) Dans ce nouveau écosystème, les utilisateurs de système d'information sont confrontés à plusieurs difficultés majeures sur l'organisation, la qualité (Mombrun & Pauchet, 2010) de l'information récupérée à partir de ces sources de données (Jérôme et al., 2011). Ils souhaitent travailler dans un environnement cohérent, avec des informations fiables malgré la multiplicité des sources des données et leurs hétérogénéités. Ces données utilisées doivent également être pertinentes par rapport à leurs contextes d'utilisation et adaptées à leurs profils (Saint-Réquier, 2012).

Notre problème est de savoir comment adapter l'usage de ces données hétérogènes et multi-sources à leurs utilisateurs. Il est très difficile voire impossible que plusieurs utilisateurs ayant des profils différents et évoluant dans des contextes distincts (Theodorakis, 2001) puissent avoir un même point de vue sur des données hétérogènes. Pour tenter de répondre à cette problématique, nous avons regardé de plus près l'expérience dans le monde de l'éducation (EIAH). Dans ce domaine, chaque objet d'apprentissage est transformé en une ressource d'apprentissage afin de mieux s'intégrer à l'environnement d'apprentissage et adapté aux besoins des utilisateurs (Pernin, 2003). Aujourd'hui il existe plusieurs propositions sur :

- l'indexation des ressources numériques (LOM)
- l'interopérabilité (SCORM, EDUPUB) et la réutilisation (OER) des ressources

numériques.

Mais il existe peu de propositions sur l'adaptabilité et la pertinence des ressources numériques. Ainsi, comment mettre en place ce processus de caractérisation des données en ressources, pour que cette ressource soit adaptée à son utilisateur (à son contexte dans l'environnement, à son profil et à son mode d'interaction).

Dans cet article, nous proposons une approche pour caractériser la pertinence d'une ressource numérique. Cette approche se base sur les informations du profil de l'utilisateur, les métadonnées de la ressource et la méthode de mesure de distance TF-IDF pour calculer la pertinence.

Caractérisation d'une Ressource Numérique

Dans le contexte de nos travaux, il est important de caractériser le terme ressource numérique pour déterminer les processus de gestion et d'usage qu'elle implique.

Le groupe de travail du CNRS (Pedauque, 2006), les travaux de (Bachimont, 2004) et de Davenport et Prusak (Davenport, 1997) nous ont principalement servi pour retenir qu'« un document numérique est un ensemble de données organisées selon une structure stable associée à des règles de mise en forme permettant une lisibilité partagée entre son concepteur et ses lecteurs » (Pedauque, 2006) et une ressource numérique désigne des informations construites dans une logique de médiation et d'usage (réception), évolutives (susceptibles d'être mises à jour), et éventuellement adaptables (personnalisables). Leur fonction est d'être utile et de rendre des services. Elles fournissent du renseignement (instantané) mais non de la preuve. (Lainé-Cruzé, 2004)

La différence entre un document numérique et une ressource numérique est selon Sylvie Lainé-Cruzé : Une ressource numérique est un objet dont la logique de conception est entièrement tournée vers l'usage. Un document numérique est un objet dont la logique de conception est entièrement tournée vers la production (fournir « une preuve ou un renseignement »).

La connaissance représente un ensemble de notions et de principes qu'une personne acquiert par l'étude, l'observation ou l'expérience, et qu'elle peut intégrer à ses acquis antérieurs. On comprend alors que la connaissance est une agrégation d'informations analysées et interprétées (résultat du traitement de l'information) comme des brevets, développements informatiques, bases d'études et de recherches, procédures/règles, etc. (Interactif, 2003). La gestion des connaissances (Knowledge Management) représente « un ensemble de modes d'organisation et de technologies visant à créer, collecter, organiser, stocker, diffuser, utiliser et transférer la connaissance. Connaissance matérialisée par des documents internes ou externes mais aussi sous la forme de capital intellectuel et d'expériences détenues par les

collaborateurs ou les experts d'un domaine » (CIGREF, 2000). En synthèse, nous avons modélisé dans le schéma de la figure 1 le processus BPMN du passage d'une donnée à une connaissance en identifiant les concepts clés.

Cette modélisation montre les étapes dans la production d'une ressource numérique. L'identification et la description de l'usage permettent de transformer un document en une ressource. Une activité est la mise en situation d'une ressource dans un environnement avec des acteurs ayant des rôles. Ainsi, dans la chaîne de la figure 1, notre travail consiste à :

- caractériser la pertinence d'une ressource numérique pour son utilisateur
- proposer une architecture pour une gestion consolidée des ressources numériques.

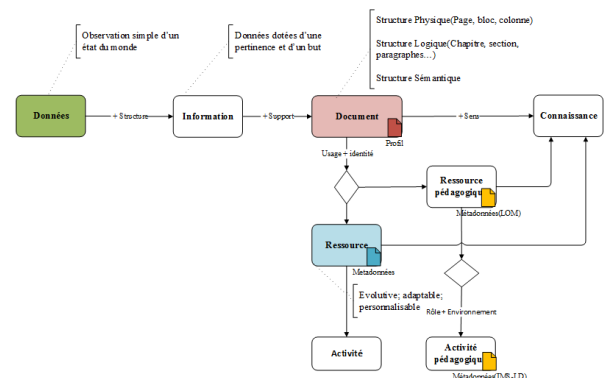


Figure 1: Document numérique, Ressource numérique et activité numérique

Dans la première partie, nous proposons une modélisation de l'utilisateur et des ces ressources dans l'environnement. Ensuite, nous mesurons la distance entre les informations du profil et les métadonnées de la ressource avec le TF-IDF, pour caractériser la pertinence des ressources numériques par rapport au profil de l'utilisateur. Dans la deuxième partie, nous proposons une architecture à travers une étude de cas pour la gestion des ressources numériques. Notre contribution concerne principalement l'usage des informations du profil de l'utilisateur et des métadonnées de la ressource pour déterminer sa pertinence. Nos résultats sont présentés sous forme d'une expérimentation dont l'objectif est d'assister l'utilisateur (chercheur) dans la gestion consolidée de ses ressources numériques.

Processus d'Identification de la Pertinence des Ressources Numériques

La croissance exponentielle des contenus numériques oblige les utilisateurs des systèmes d'informations à faire un tri et une sélection des contenus les plus pertinents et les mieux adaptés à leurs besoins d'usage. Dans ce contexte, nous proposons une approche permettant de déterminer automatiquement la

pertinence d'une ressource numérique. Ces ressources numériques sont utilisées dans des environnements numériques. Nous employons le terme ressource numérique qui caractérise mieux l'usage dans un système. Pour (Roose, 2008), « Tout système d'information conçu dans une logique de service devrait supporter principalement des ressources numériques pour respecter la logique d'usage voulu par le système. Dans cette logique d'usage, « l'utilisateur numérique » est un acteur essentiel du paysage de l'information numérique, il agit, interagit avec d'autres, reprenant, transformant, produisant de l'information. » Ainsi, commencer par étudier l'utilisateur et ses interactions avec la ou les ressource(s) numérique(s) est important, c'est à ce niveau que nous pouvons caractériser la gestion consolidée des ressources numériques, en caractérisant l'utilisateur à travers son profil, la ressource et les interactions entre l'utilisateur et ces ressources dans le contexte de l'environnement. Le but de cette caractérisation est de déterminer et d'optimiser la cohérence et la pertinence, en un mot la qualité des interactions.

Profil de l'Utilisateur –Extension IMS-LIP

Dans nos travaux, nous avons proposé une modélisation de l'utilisateur dans notre système avec une extension du modèle IMS-LIP(Sawadogo, Champagnat, & Estrailier, 2014).

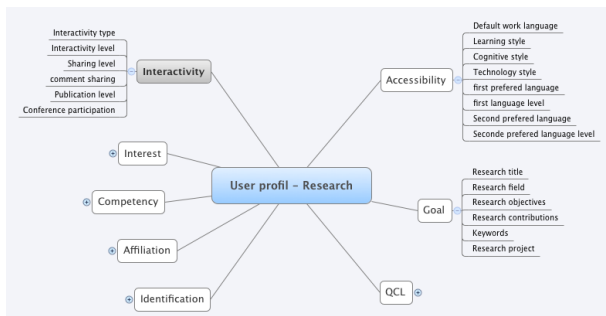


Figure 2 : user profil - IMS-LIP extension

Nous avons proposé un modèle (Figure 2) de l'utilisateur dans notre système et nous avons présenté des techniques pour remplir les informations de l'utilisateur. La suite de ce travail est d'identifier les éléments du profil et de la ressource nécessaire pour le calcul de la pertinence.

La figure 3 présente un modèle très simple, nous permettant d'appliquer notre approche pour déterminer la pertinence d'une ressource numérique.

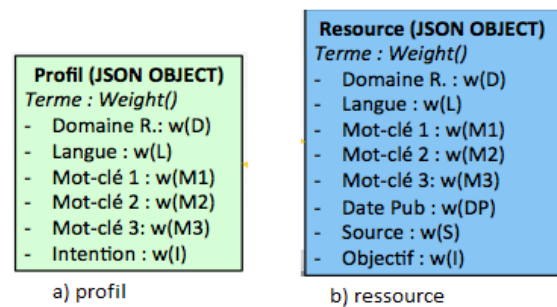


Figure 3: Exemple critère du profil et de la ressource

Ressource Numérique - Profil d'application du LOM

Une ressource est représentée par des métadonnées, nous avons choisi le LOM pour plusieurs raisons, d'abord c'est un standard dans le domaine de l'éducation, et ce modèle est facilement extensible pour les contraintes de notre expérimentation. Nous avons modélisé un profil d'application LOM pour le domaine de la recherche.

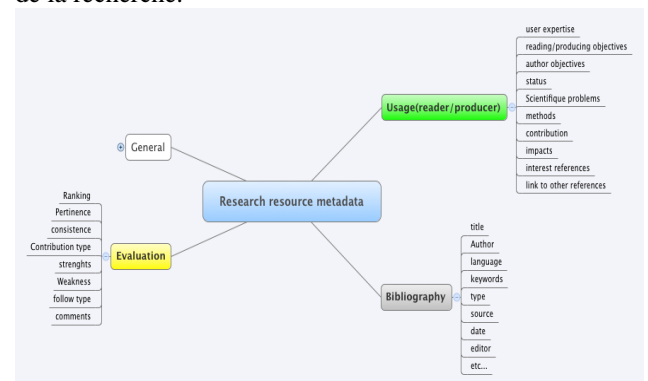


Figure 4: LOM pour la recherche

La figure 3 présente notre modèle de métadonnée pour les ressources. Nous avons choisis des critères (Figure 3b) de ce modèle pour déterminer la pertinence de chaque ressource.

3. Calcul de la pertinence – Ressource et profil

Une des étapes du processus de consolidation de la figure 4 consiste à évaluer la cohérence et la pertinence d'une ressource en fonction du profil de l'utilisateur. Pour la proposition de la méthode de calcul de la pertinence, nous nous sommes intéressés à deux acteurs du tableau 1.

Le premier acteur est l'utilisateur qui a besoin d'utiliser des ressources pertinentes pour la réalisation des situations telles que définies au sein de notre équipe (Trillaud, Pham, Rabah, Estrailier, & Malki, 2012).

Le second acteur est le compagnon qui est chargé de calculer la pertinence des ressources en fonction de l'utilisateur. Nous utilisons l'approche MMPE (Modèle Multicritère basé sur la découverte des Préférences à partir d'Évaluation) (Adomavicius, Manouselis, & Kwon, 2011) pour l'évaluation des critères et la méthode TF-IDF (Wu, Luk, Wong, & Kwok, 2008) pour la mesure de la pertinence.

Notre méthode fonctionne comme suit :

- l'utilisateur évalue les différents critères du profil;
- chaque critère est pondéré dans le système ;
- l'utilisateur évalue les différents critères du modèle de ressource ;
- chaque critère est également pondéré ;
- nous appliquons dans le système la méthode de mesure de distance TF-IDF sur les deux modèles afin de déterminer la pertinence de la ressource pour le profil de l'utilisateur.

Dans (Bethard, Butcher, & Martin, 2009), les auteurs proposent une technique de caractérisation de la qualité des ressources par une méthode d'apprentissage. Notre proposition se base sur le profil de l'utilisateur et vient compléter les travaux qui existent.

Processus de gestion consolidée

Pour valider notre approche, nous avons proposé un processus de gestion consolidée dans le système de notre étude de cas. La figure 4 présente ce processus.

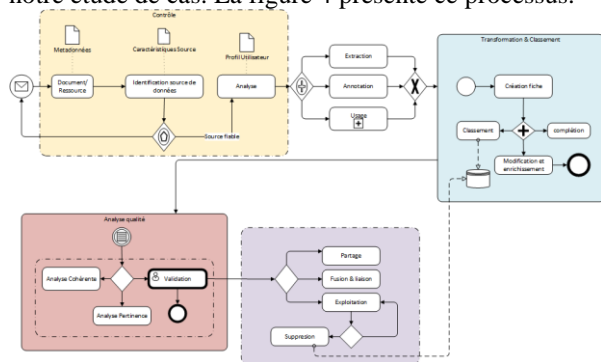


Figure 4: Processus BPMN de consolidation de ressources numériques

Le tableau 1 nous présente les acteurs avec les rôles dans le processus de la figure 4

Acteurs	Rôles	Fonctionnalités
Auteur	Producteur de document numérique	- Création de documents numériques
Utilisateur	Usager de l'information numérique	- Utilisation de ressource numérique - Capitalisation de connaissance - Consolidation de ressource numérique
Compagnon (système)	Assistance dans l'organisation de l'environnement de l'utilisateur.	- Assistance et automatisation des tâches - Gestion de pertinence dans le système - Découverte automatique de source d'information - Substitut (Modèle) de l'utilisateur
Communauté	Validateur d'information numérique	- validation de la qualité de l'information

		numérique - partage et recommandation d'information numérique
--	--	--

Tableau 1 : Acteurs et rôles dans le processus de consolidation

Dans Tableau 1, les différentes fonctionnalités des acteurs correspondent aux principales fonctionnalités offertes par les solutions de gestion de contenus telles que décrites dans (Interactif, 2003). L'ensemble du processus ainsi que les fonctionnalités correspondent au modèle du cycle de gestion de connaissance tel que défini par M. Gardoni et al. Dans ce modèle, nous avons ajouté un compagnon qui assiste l'utilisateur dans la gestion et principalement pour déterminer la pertinence de ses ressources. Dans notre étude de cas, nous avons défini quelques fonctionnalités du compagnon.

Etude de cas

Nous travaillons sur l'application de nos résultats dans un environnement numérique dédié à la recherche. Cet environnement permet particulièrement aux chercheurs de collecter, de gérer, de produire, de partager et de capitaliser des ressources numériques pendant la réalisation de leurs travaux. Le chercheur est un utilisateur confronté aux problèmes de masse de données hétérogènes et multi-sources. Les fonctionnalités que nous proposons permettent d'expérimenter nos résultats sur la pertinence des ressources et l'architecture logicielle adaptée pour supporter ces ressources. Les outils tels que Mendeley¹, Zotero² sont des outils bibliographiques et d'organisation de recherche. Le cahier de laboratoire (Anabelle et al., 2011) sert pour le suivi des expériences de recherches et leurs pérennités. Ces outils ont des limites quant à la caractérisation de la pertinence des ressources qui y sont gérées. L'objectif de notre proposition est de fournir une solution à ces outils pour aider l'utilisateur à avoir des ressources pertinentes. C'est dans cette perspective que nous proposons d'étendre les fonctionnalités de ces outils à travers notre application PRISE (PeRsonal Interactive research Smart Environment). Les fonctionnalités principales de notre application sont :

- système de gestion de ressources numériques : il permet le partage, la validation collaborative, la recommandation de ressources et surtout la gestion des ressources numériques dédiées à la recherche.
- système de consolidation : il permet de vérifier la cohérence et la pertinence des ressources utilisées dans l'environnement de l'utilisateur.
- assistant intelligent (Compagnon) : son rôle est de piloter, d'automatiser les tâches de gestion

¹ <http://www.mendeley.com>

² <http://www.zotero.org>

des ressources numériques.

Dans l'environnement à chaque ressource est associée une fiche de lecture. Chaque fiche de lecture représente une connaissance acquise à travers l'usage d'une ressource. Le compagnon peut contrôler et automatiser les tâches de création et de complétion des fiches de lecture suivant le profil de l'utilisateur. Elle nous permet également de persister des données structurées dans la base de données du système. La figure 5 représente l'architecture de notre étude de cas.

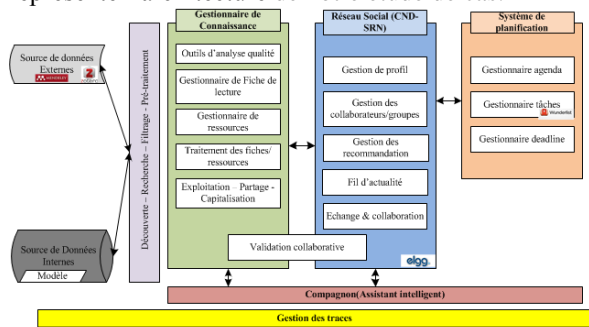


Figure 5 : Architecture générale de l'environnement PRISE

Dans cette architecture, nous avons intégré le modèle de profil de l'utilisateur de la figure 2 ainsi que le modèle de ressource numérique de la figure 4.

Pour caractériser la pertinence des ressources, le système :

- Gère les identités des utilisateurs, chaque profil est complété à travers deux méthodes : Complétion manuelle par l'utilisateur, et complétion à l'aide des traces du système.
- Vérifie la pertinence des ressources ;
- Gère et complète les métadonnées des ressources :
 - Biblio. : métadonnée pour la documentation, extraction à l'aide des méthodes existantes ;
 - usage : usage de la ressource, nécessite un niveau de sémantique plus fort ;
 - évaluation : évalue la qualité de la ressource.

Expérimentation

Pour l'expérimentation de notre proposition, nous avons créé un environnement basé sur notre étude de cas permettant aux chercheurs de gérer leurs ressources numériques. L'environnement PRISE³ est composé :

- d'un système de gestion de ressources numériques avec des fiches de lecture
- un système de gestion de réseau social avec notre modèle de profil du chercheur ;
- un système de gestion d'agenda et de tâches.

La figure 6 présente l'architecture de notre système avec principalement deux grands composants, le gestionnaire de ressource numérique et le réseau social. Chaque utilisateur est identifié par son profil dans le réseau social.

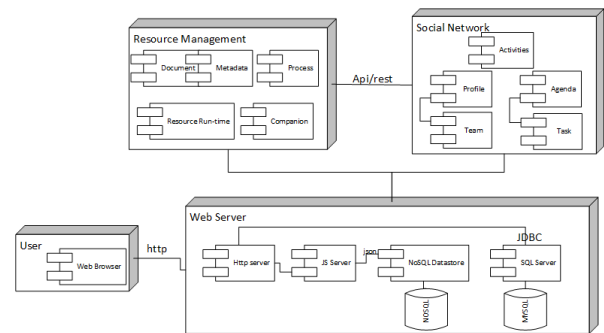


Figure 6 : Architecture PRISE



Figure 7: Environnement PRISE - Système de gestion de ressources numériques

À travers cette expérimentation, nous avons identifié quelques éléments (Langue de préférence, niveau de langue, champs de recherche, mots-clés, type et niveau d'interactivité...) du profil permettant d'assister l'utilisateur dans la gestion consolidée. Nous mettons dans l'environnement des processus de contrôles sur la qualité des ressources en fonction des critères choisis et des mécanismes de mesure de la pertinence.

Contributions

Notre contribution, concerne la méthode de calcul de pertinence d'une ressource numérique en utilisant les informations du profil de l'utilisateur. Ces travaux sont toujours en cours et nous envisageons d'étendre notre proposition sur les situations d'un scénario.

Dans cet article, nous avons considéré le profil de l'utilisateur et les métadonnées des ressources numériques. Nous savons que ces ressources sont souvent utilisées dans un contexte précis et pour une situation pédagogique précise. Donc nous étendons notre approche pour prendre en considération le profil, la ressource et la situation d'apprentissage concernée.

Conclusion et perspectives

Nous avons présenté dans cet article une nouvelle approche pour caractériser la pertinence des ressources numériques dans une architecture d'un système interactif. L'intérêt de ce travail est de proposer un système permettant d'assister un utilisateur dans la gestion consolidée de ses ressources numériques. Nous avons proposé une étude de cas pour valider cette

³ PeRsonal Interactive research Smart Environment

méthode dont les résultats sont en cours d'expérimentation.

Dans la suite de nos travaux, nous proposons des approches permettant de vérifier la cohérence d'une ressource dans la situation d'apprentissage avec la même approche. Puis proposer un modèle de ressource numérique orienté vers la pertinence dans l'usage. Il existe plusieurs méthodes de conception de ressources numériques. Dans le monde de l'éducation, nous avons le LOM, qui est orienté vers l'indexation, le SCORM, qui est orienté vers l'interopérabilité et les OER, qui sont orientés vers la réutilisation. Mais nous voulons un modèle orienté vers l'adaptabilité.

Notre travail vise à offrir à chaque utilisateur la meilleure ressource adaptée à son profil et à sa situation d'apprentissage. Cela permet aux utilisateurs (apprenant, formateur, concepteur de situation) des environnements d'apprentissage d'avoir une meilleure expérience d'usage des ressources numériques pour mieux atteindre les objectifs dans leurs scénarios.

Références

- Adomavicius, G., Manouselis, N., & Kwon, Y. (2011). Multi-Criteria Recommender Systems. In *Recommender Systems Handbook* (pp. 769–803). Springer US.
- Anabelle, A., Céline, B., Christophe, B., Alexandre De, B., Emmanuelle, D.-M., Eric, G., ... Céline, R. (2011). Le cahier de laboratoire électronique (CLE). *STP Pharma Pratiques*, 21, 1–28.
- Bachimont, B. (2004). *Arts et sciences du numérique: Ingénierie des connaissances et critique de la raison computationnelle. Mémoire d'Habilitation à Diriger des Recherches*, Université De Technologie De Compiègne.
- Bethard, S., Butcher, K., & Martin, J. H. (2009). Automatically Characterizing Resource Quality for Educational Digital Libraries, 221–230.
- CIGREF. (2000). *Gérer les connaissances : Défis, enjeux e conduite de projet*. Davenport, T. H. (1997). *Information Ecology: Mastering the Information and Knowledge Environment* (1st ed.). Oxford University Press, {USA}.
- Gantz, J., & Reinsel, D. (2012). The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the Future, 2007*, 1–16.
- Interactif, B. (2003). EtudeLCMS. EDUSCOL.
- Jérôme, A., Nicolas, B., Laure Berti, E., Sylvie, G., Mathieu, R., & Fatiha, S. (2011). Qualité des données et des connaissances. In QDC (Ed.), *QDC - EvalECD* (p. 87). Brest, France: QDC.
- Lainé-Cruzel, S. (2004). Documents, ressources, données: les avatars de l'information numérique. *Information-Interaction-Intelligence*, 4, 105–120.
- Mombrun, Y., & Pauchet, A. (2010). Collecte, analyse et évaluation d'informations en sources ouvertes. *Atelier COTA Des 21es Journées Francophones d'Ingénierie Des Connaissances, Nîmes : France (2010)*.
- Pedauque, R. T. (2006). *Le Document à la lumière du numérique*. {C&F} éditions.
- Pernin, J. (2003). Objets pédagogiques: unités d'apprentissage, activités ou ressources. *Revue "Sciences et Techniques Educatives", Hors Series*, 179–210.
- PFIA, A. (2013). SOS-DLWD 2013 Des sources Ouvertes au Web de Données. In A. P. 2013 (Ed.), *Atelier SOS-DLWD'2013 : des Sources Ouvertes au Web de Données* (p. 116). Atelier PFIA 2013.
- Roose, P. (2008). *De la réutilisation à l'adaptabilité. Mémoire d'Habilitation à Diriger des Recherches*, Université de Pau et des Pays de l'Adour. Saint-Réquier, A. (2012).
- Sélection adaptative de Services de Recherche d'Information web par analyse du besoin et du comportement de l'utilisateur. *CORIA*, 395–400.
- Sawadogo, D., Champagnat, R., & Estraillier, P. (2014). User profile modelling for researcher digital resource management systems. *Proceedings of the International Workshop on Personalization Approaches in Learning Environments - à paraître*.
- Theodorakis, M. (2001). *Contextualization : An Abstraction Mechanism for Information Modeling*. Doctoral Dissertation Department of Computer Science University of Crete.
- Trillaud, F., Pham, P. T., Rabah, M., Estraillier, P., & Malki, J. (2012). Situation-Based Scenarios For E-Learning. In *IADIS e-learning 2012*. Lisbon, Portugal.
- Wu, H. C., Luk, R. W. P., Wong, K. F., & Kwok, K. L. (2008). Interpreting TF-IDF term weights as making relevance decisions. *ACM Transactions on Information Systems*, 26(3), 1–37.