User Profile Modelling for Digital Resource Management Systems

Daouda Sawadogo, Ronan Champagnat, and Pascal Estraillier

L3i Laboratory, University of La Rochelle, France
{daouda.sawadogo,ronan.champagnat,pascal.estraillier}@univ-lr.fr
http://l3i.univ-lr.fr

Abstract. In this paper, we focus on user assistance in an interactive and adaptive system. The increase in production of digital data, these last two years has raised several issues regarding the management of heterogeneous and multiple source data in the user's environment. One distinctive feature of user assistance is the user model that represents essential information about each user. We propose a modeling of a scientific user, who is a researcher, in a personal resource management system. Our methodology is based on the IMS-LIP standard extension and the user's trace management. Our work can assist users in the consolidated management of their resources and their environment, based on the user's profile. Experimental results obtained by the emperical evaluation in our laboratory are presented.

Keywords: User modelling, adaptation, profile management, trace-based system, research resource management system

1 Introduction

Currently, user profiles have a very important role in all digital environments [1] [2]. The use of profiles is one of the means that enables systems to adapt to the specificities of its users in the digital environments. Every information system, created as a service, ought to support mainly the digital resources in order to respect the logic of usage required by the system. In this logic of usage, the digital user is an essential actor in the landscape of digital information, he acts and interacts with others; making use of, transforming and producing information. Each individual composes their own digital identity, thus generating the e-reputation that characterises them and for which they have to take care [3]. Our general objective was to design an application that would provide a consolidated management of the users' digital resources and environment. We opted for an application that would be used by researchers as several studies had shown that the problem of researcher profiling poses many challenging issues including : the automatic extraction of researcher profiles from distributed sources (homepages, indexing authorities, etc.) [4], the consistency and completeness of information and the resolution of ambiguity [5]. These issues are further aggravated by the explosion of information of research artifacts with the research community. Researchers'

Facets of user assistance	Characteristics to act on this facet
Resources management	Characterises digital resources in order to calculate the relevance of these resources for users in the systems
Process management	Characterises resources management rules and user's process management interaction to adapt the usage
Collaboration management	Criterions for building a user trust community [7], resource sharing and recommendation

Table 1. Facets of assistance in PRISE

information is mostly scattered and is represented in a syntactic manner [4] (not meaningfully described) thus minimizing the interoperability between heterogeneous knowledge and information sources [6]. In our application we first decided to design a consistent researcher profile based on IMS-LIP extension and a trace-based system which can be used to complete the researcher profile. Indeed, the user's profiles provide the system with pertinent information to assist the user in aspects defined in Table1.
Our contribution concerns the first facet of this table, i.e. resources management. In this paper we proposed :

- an extension of the IMS-LIP model for research;
- criterions to characterise the relevance of a digital resource according to the user profile.

2 State of the art

The user model is a representation of information about an individual user that is essential for an adaptive system to provide the adaptation effect. i.e. to behave differently for different users[2]. The researcher is like a learner in the system, who uses resources to acquire knowledge and produce scientific results.

Usually, five most popular and useful features are found when viewing the user as an individual: the user's knowledge, interests, goals, background and individual traits. In [2] The authors discuss modelling the context of a user's work. In our work we wanted to modelize the user in order to assist him in his resource management process and to **characterise his interaction in the environment**.
Some existing learner models in the literature are as follows :

1. IEEE PAPI-Learner [8] (Public and Private Information-Learner specification) is a standard developed within the IEE P1484.2 Learner Model Working Group. Its objective is to specify the semantics and the syntax of a Learner Model, which characterises a learner and his knowledge. It includes elements such as knowledge, skills, abilities, learning styles, records, and personal information. PAPI Learner was initially developed for learning technology

applications but can easily be extended to other types of human related information such as medical and financial applications. PAPI is one of the first standards which provides a framework that organizes learner data. There is lot of learner data that this standard does not take into account [9] and which can be exchangeable between various e-learning systems. This explains why this proposal has been extended by IMS in its new standard IMS-LIP[10].
2. IMS-LIP[11] is based on a data model that describes those characteristics of a learner that are needed for the general purposes of: Recording and managing learning-related history, goals and accomplishments; Engaging a learner in a learning experience; Discovering learning opportunities for learners. The specification supports the exchange of learner information among learning management systems, human resource systems, student information systems, enterprise e-learning systems, knowledge management systems and other systems used in the learning process. We note that IMS LIP provided valuable extensions compared to the PAPI model, but it does not meet all the systems' needs in terms of user data completeness and management, which explains its adaptation within application profiles.
3. SERPOLET [10] offers solutions to the issue of learner data interoperability between e-learning systems. This model is based on IMS-LIP with extension to the need of learner data interoparability.

From these propositions, we chose to extend the IMS-LIP standard for many reasons. Firstly, IMS-LIP is the interoperability standard chosen by CEN/ISSS [12]. Secondly, it defines a user data model as a set of 11 categories to be imported or exported between systems. The IMS-LIP extension in Fig. 1 that we defined provides relevant information about the researcher and his activities. The trace-based system maintains a consistent profile and makes it more complete.

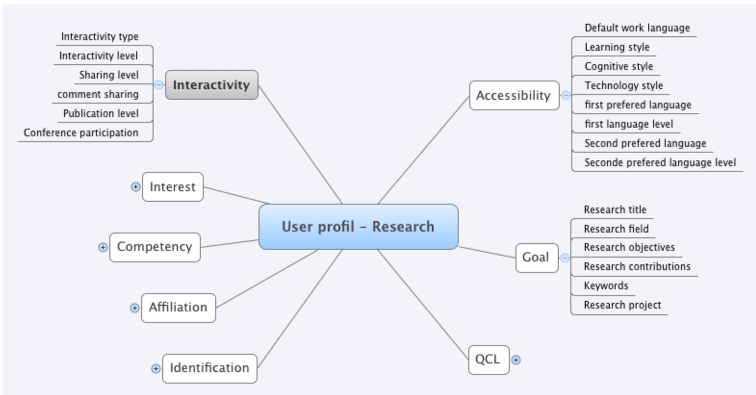

Fig. 1. IMS-LIP extension application profile for researcher

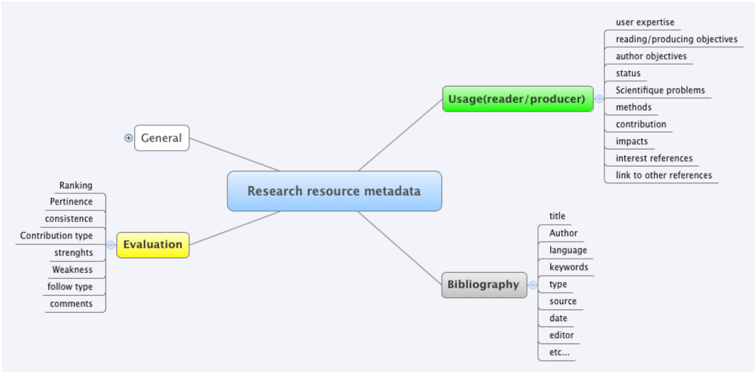

Fig. 2. LOM application profil for researcher resource

We need to understand the interactivity of the user in order to assist him. In [13] the authors proposed an ontology-Based User modelling for Knowledge Management Systems using IMS-LIP. They defined the behaviour category that extended the IMS-LIP model. In our work, we consider that the behaviour is the dynamic information of profile, so we use interactivity to capture the model of user interaction with the resources which can characterise the user model.

3 Our methods

It should be remembered that the objective of the profile is to provide relevant information about the user and provide mechanisms to assist a user in the consolidated management of his resources and environment. Firstly, we proposed an extension of the IMS-LIP model in Fig. 1 to take into account the user's interaction. In digital resources management system it is very important to characterise the user's interactivity type and interactivity level to adapt the digital resource usage process.
Secondly, our methods involved determining what criterions are necessary to characterise the user's digital resources relevance and how to use it the latter. We identified two levels of use of a user's profile:

- the static level: we can directly use the informations, of the user, stored in the database.
- the dynamic level: we can calculate the information of the user from information stored and his context in the system.

The first level requires less effort in the semantic analysis. The informations that we deemed relevant for the profile are : research domain, research field, user's prefered language, user's technology style, qualification level, keywords and research interests
The second level requires more effort in the semantic analysis. The informations

needed are : the user intention [14] [15], the user objective the user interactivity type, the user interactivity level and the scientific problems. These informations are context-aware, so they require a real-time analysis and they characterise the user's information needs. We have chosen these criterions because they are importants for a researcher looking for informations and in the objectives of our study. Our contribution consisted in combining these two types of data on the user, in order to determine the relevance of the resource and to adapt it to the usage needs.
We characterised the resource using all the metadata associated with it. The Fig. 2 shows the main elements of the research resource metadata in our environment. Despite the fact that quite often the metadata are not the main elements that characterise a document. In the context of this work we put forth the hypothesis that the metadata provide a good characterisation of the resource. In our experimental environment 3, we characterised these metadata as mainly: bibliography data [16], usage (reader/producer) data and evaluation data. After that we compared the matching of the profile information 4 with the resource metadata. We experiment as well the use of TF-IDF [17] to calculate these levels of relevancy. However this aspect is not yet ready enough to be presented.

Our methodology for modelling the research user and managing his profile consists in explaining how the data in the user's profiles is collected, how it is fomalized and how it is used to satisfy system assistance needs.

There are two possibilities for collecting the necessary user data [18] :

1. Explicit collection of the data: users' preferences are found explicitly, by asking them to submit the necessary information manually before any personalisation is provided. Explicitly entered profile information is considered to be high quality, but users generally dislike having to spend time and effort submitting data to a system, especially when the benefits may not be immediately obvious. This can make the explicit collection of sufficient profile data difficult[18]. This type of collection is achieved through the fields that we propose in the IMS LIP model extension for the researcher.
2. Implicit collection of data: users' preferences are inferred from their normal interactions with the system. The advantage of collecting profile data this way is that the user is relieved of the burden of having to supply and keep up-to-date the necessary information. But the implicit measures of interest are generally thought to be lower quality than explicitly gathered preferences [19].

The purpose of trace-based management system is to identify the trace elements necessary to propose a relevant user profile and system for the automatic update of user's data.

We considered the activity category in the profile because it is data that we mainly use in our research. There are traces in the past of the users. Furthermore, there are many interactions in an interactive environment and during its execution, actors can generate traces [20]. We combined the user profile and the trace-based system to construct our model. These traces were collected from the interaction of the user with the environment. The users' traces were used by carrying out three steps consecutively. Firstly, a function was defined to collect all traces that were produced by the interaction between the user and the environment. We propose then a step of traces' filtering. This step aims to filter all the traces in order to get the pertinent traces. Finally, we present the filtered traces and integrate them in the user's profile.

4 Experimentation and discussion

To experiment our proposal we created an environment that allows a researcher to manage their digital resources. The system offers a set of tools that make it more effective in the production of scientific results. Our research environment PRISE (PeRsonal Interactive research Smart Environment) includes several tools : digital resources management (Fig. 3), social network including our model of the researcher's profile (Fig. 4) and events management.

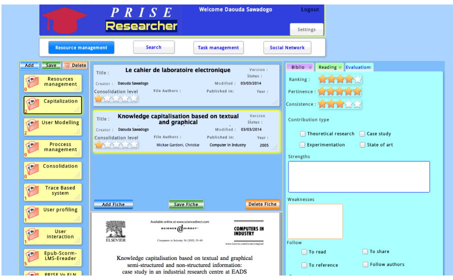

Fig. 3. Digital Resource management in PRISE

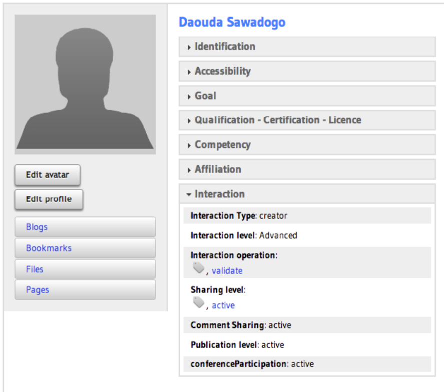

Fig. 4. example of IMS-LIP extension profile for researcher

Through our experimentation in PRISE, we used the elements of the user's profile necessary to achieve the first facets proposed in the Table 1.
Table 2 shows these profile criterions in order to characterise the resource management relevance in the system. These criterions were used to assist the user resources management process, resources relevant for the user and resource recommendation criterions based on the user profile.

Our experimental system implemented our LOM [21] application profile for research. The resource metadata was stored in a NoSQL database using JSON API to manipulate it. NoSQL has the quality required to be a document store database. The user profile model was stored in SQL database for reuse needs and we used REST web services to retrieve the profile information.

Facets of user assistance	Item criterions used in the profil
Resources management	Prefered language; Language level; Research field; Keywords; Research field level; Qualification level; Research interests; interactivity type; interactivity level

Table 2. User profil criterions for user assistance in PRISE

We have discussed our work in terms of the research described in [22] [23]. These authors made an important contribution by automatically characterizing the resource quality using the machine learning. Our contribution completes these works using the user profile and resource metadata. We also found that our user model was more complete and provided relevant information about the researcher in the system. The models we compared, in the main lack information on the objectives of the researcher, their scientific problems, their preferences, as well as entire elements on their activities as well. Our model could also serve as a reference for the social research networks like Mendeley, Academia and ResearchGate etc.

5 Conclusion and future works

In this paper we presented work in progress on the researcher profile modelling in a personal research system. We have identified and used the relevant characteristics to adapt and assist the researcher with his digital resources, to achieve relevant management of his resources. The main contribution was in consistent researcher profile modelling based on IMS-LIP extension and the approach of trace-based management to fill the user profile model use this model to assist user's in research resources management systems.
Future work will firstly consist in improving the dynamic usage of the user profile information and the TF-IDF method to calculate the resource relevance. Secondly, it will provide some mechanisms for a dynamic usage of user's resource process based on the user profile, in order to maintain his environment's consistency and assist him in the consolidated management of his resources.

References

1. Brusilovsky, P.: Methods and techniques of adaptive hypermedia. User Model. User-Adapt. Interact. **6**(2-3) (1996) 87–129
2. Brusilovsky, P., Millán, E.: User models for adaptive hypermedia and adaptive educational systems. In: The Adaptive Web. (2007) 3–53
3. Roose, P.: From reuse to adaptability. Hdr, University of Pau and Adour Country (2008)
4. Oberle, D., Spyns, P.: The knowledge portal "ontoweb". In: Handbook on Ontologies. Springer (2004) 499–516

5. Tang, J., Zhang, D., Yao, L.: Social network extraction of academic researchers. In: ICDM. (2007) 292–301
6. Adnan, S., Tahir, A., Basharat, A., de Cesare, S.: Semantic agent oriented architecture for researcher profiling and association (semora). In: Web Intelligence/IAT Workshops. (2009) 559–562
7. Nepal, S., Sherchan, W., Paris, C.: Building trust communities using social trust. In: Advances in User Modeling. Springer (2012) 243–255
8. IEEE LTSC: Ltsc public and private information for learners, http://www.cen-ltso.net/main.aspx?put=230
9. Paramythis, A., Loidl-Reisinger, S.: Adaptive learning environments and e-learning standards. In: Second European Conference on e-Learning. (2003) 369–379
10. Oubahssi, L., Grandbastien, M.: From learner information packages to student models: Which continuum? In: Intelligent Tutoring Systems, Springer (2006) 288–297
11. IMS GLC: IMS Learner Information Packaging Information Model Specification
12. Ehlers, U.D., Pawlowski, J.M.: Handbook on quality and standardisation in e-learning [electronic resource]. Springer (2006)
13. Razmerita, L., Angehrn, A.A., Maedche, A.: Ontology-based user modeling for knowledge management systems. In: User Modeling. (2003) 213–217
14. Chen, Y.C., Shang, R.A., Li, M.J.: The effects of perceived relevance of travel blogs' content on the behavioral intention to visit a tourist destination. Computers in Human Behavior **30** (2014) 787–799
15. Djamasbi, S., Strong, D.M.: The effect of positive mood on intention to use computerized decision aids. Information & Management **45**(1) (2008) 43–51
16. Lipinski, M., Yao, K., Breitinger, C., Beel, J., Gipp, B.: Evaluation of header metadata extraction approaches and tools for scientific pdf documents. In: JCDL. (2013) 385–386
17. Wu, H.C., Luk, R.W.P., Wong, K.F., Kwok, K.L.: Interpreting tf-idf term weights as making relevance decisions. ACM Transactions on Information Systems (TOIS) **26**(3) (2008) 13
18. Keenoy, K., de Freitas, S., Levene, M., Jones, A., Brasher, A., Waycott, J., Kaszas, P., Turcsanyi-Szabo, M., Montandon, L.: Personalised trails and learner profiling within e-learning environments. Personalised and Collaborative Trails of Digital and Non-Digital Learning Objects **22**(4.1) (2004) 1–43
19. Nichols, D.M.: Implicit rating and filtering. In: Proceedings of the Fifth DELOS Workshop on Filtering and Collaborative Filtering, ERCIM (1997) 31–36
20. Laflaquière, J., Settouti, L.S., Prié, Y., Mille, A.: Trace-based framework for experience management and engineering. In: Proceedings of the 10th International Conference on Knowledge-Based Intelligent Information and Engineering Systems - Volume Part I. KES'06, Springer-Verlag (2006) 1171–1178
21. Barker, P.: What is ieee learning object metadata/ims learning resource metadata? Cetis standards briefings series (2005)
22. Bethard, S., Wetzer, P., Butcher, K., Martin, J.H., Sumner, T.: Automatically characterizing resource quality for educational digital libraries. In: Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries. JCDL '09, ACM (2009) 221–230
23. Wetzler, P., Bethard, S., Leary, H., Butcher, K., Bahreini, S.D., Zhao, J., Martin, J.H., Sumner, T.: Characterizing and predicting the multifaceted nature of quality in educational web resources. ACM Trans. Interact. Intell. Syst. **3**(3) (October 2013) 15:1–15:25