What If Annotations Were Reusable: A Preliminary Discussion

Nikos Manouselis[1] and Riina Vuorikari[2]

[1] Greek Research & Technology Network
56 Messogion Str.
Athens, Greece
`nikosm@ieee.org`
[2] European Schoolnet, Rue de Treves, 61,
1040 Brussels, Belgium
`Riina.Vuorikari@eun.org`

Abstract. This paper discusses the rationale for the representation of user feedback in a structured and reusable format so that it can be reused by different recommender systems. We emphasize how information about the context can be included in such a representation. This work-in-progress takes place in the context of two large European initiatives that set up collections of digital educational resources in distributed repositories to serve the needs of different user communities, and to collect user feedback such as ratings, bookmarks and tags related to the resources. The overall aim is to facilitate the exchange and reuse of their data sets in order to support recommendation of appropriate resources to the end users.

Keywords: Learning resource metadata, annotations, interoperability, reusability, data set.

1 Introduction

A representation of implicit or explicit feedback from the users regarding the candidate items is required by a recommender system to produce a recommendation. This feedback can be in several forms. For example, in the case of collaborative filtering systems it can be ratings or votes (i.e. if an item has been purchased, viewed or bookmarked). In the case of content-based recommenders, it can be product reviews or simple tags (keywords) that users provide for items. Additional information is also required such a unique way to identify who provides this feedback (user ID) and upon which item (item ID). The user-rating matrix [1] used in collaborative filtering systems is a well-known example.

User feedback representations are stored as data sets that can be used in a number of ways. A very popular use is in the context of evaluation experiments. Published data sets, such as the MovieLeans and EachMovie ones, are very often used in experiment testing the performance of new algorithms [2]. Less often, such experiments are based on usage data sets from particular applications (e.g. to support the needs of an existing e-commerce site). This means that in many evaluation experiments, testing takes place using data sets that are different than the ones of the potential application context. We

believe that a structured way to represent the different types of user feedback in a reusable and interoperable format, while also maintaining information about the context in which this user feedback has been collected, could prove of particular value.

This would be particularly useful in scenarios where different online environment want to exchange the feedback that users have provided upon the same items, so that they can enhance the available data for their recommender systems. We deal with such situation in two European initiatives that both collect user feedback on digital educational resources: the Metadata Ecology for Learning and Teaching (MELT, http://info.melt-project.eu) and the Organic.Edunet (http://www.organic-edunet.eu) eContentplus projects.

In this paper we describe our first steps towards deciding upon a common representation format so that the two projects can collect, store, and exchange data sets containing ratings, bookmarks and tags by their users. It is continuation for the work presented at RecSys07 Doctoral Symposium [3].

2   Problem Description

The main questions discussed in this paper are: (a) can we find a way to represent and store data sets with user feedback in a structured, interoperable and reusable format?; and (b) is it safe to assume that some given data set with user feedback, which is stored in such a structured, interoperable and reusable way, can be used in other context, such as a new recommender system of a different application context?

Figure 1 illustrates one of the main problems to be addressed in a rather simplistic way. In this figure, the same movie can be found as an item (e.g. for a review, purchase, viewing, downloading) in three different application contexts: one movie

Fig. 1. Ratings on the same items from different contexts