GDR: A Language for Flexible Interlinking, Integrating, and Enriching User Data

Erwin Leonardi[1], Jan Hidders[1], Geert-Jan Houben[1],
Fabian Abel[2], Eelco Herder[2]

[1] Delft University of Technology, PO Box 5031, 2600 GA Delft, the Netherlands
{e.leonardi, a.j.h.hidders, g.j.p.m.houben}@tudelft.nl
[2] L3S Research Center, Appelstrasse 9a, 30167 Hannover, Germany
{abel, herder}@l3s.de

Abstract. The Grapple Derivation Rule (GDR) language is a rule language that is used to interlink, integrate, and enrich user data from multiple distributed user data repositories as well as data published as Linked Data on the Web (such as DBpedia and GeoNames). To demonstrate this, we have applied GDR in the Grapple User Modeling Framework (GUMF) – a framework that facilitates the brokerage of user profile information and user model representations for adaptive systems – to leverage user data for the purpose of personalization in e-learning systems.

1 Introduction

Adaptation and personalization become important features offered by today's Web applications and services. To be able to provide such personalized and adapted contents and services, these applications explicitly or implicitly collect data about their users and their behavior, and thus build up user profiles. Furthermore, these Web applications are becoming increasingly connected. This creates the interesting challenge of performing user modeling and personalization across application boundaries. It requires approaches allowing various Web applications to exchange, reuse, interlink, and integrate user data. We also observe that there is a growing effort to make data interlinked and freely available and accessible on the Web following the principles of Linked Data[1]. This data is typically published as RDF (c.f. http://www.w3.org/RDF) and accessible through a SPARQL (c.f. http://www.w3.org/TR/rdf-sparql-query) endpoint. This effort opens opportunities to unlock a huge potential of background data, to complement and enhance the user data. By *reusing* this interlinked data (such as DBpedia[2] and GeoNames[3]), various relationships between user data can now be derived and discovered, and thus make user data more meaningful and richer. This opens opportunities for Web applications to provide

[1] http://www.w3.org/DesignIssues/LinkedData .html
[2] http://dbpedia.org/
[3] http://geonames.org/

better adaptation and personalization to their users as they have more knowledge about the users.

In this paper we demonstrate the Grapple Derivation Rule (GDR) language [1] that allows for the integration of user data from multiple logically distributed user data repositories and background data published as Linked Data on the Web. The important features that GDR has are the following: (1) It provides the basis for reasoning over distributed user data. (2) It offers a view-based query mechanism to the applications. (3) It allows the user data to be extended and enriched using other user data repositories and Linked Data available on the Web. (4) It gives a flexible way of integrating, interlinking, and enriching user data. To this end, we have integrated GDR into GUMF [2] in the context of the GRAPPLE project[4] to enable the client applications to create flexible, rule-based plug-ins for leveraging user data that go beyond the features of RuleML [3] or SWRL [4].

2 GDR

In human-readable syntax, a GDR rule has the form: $a \Rightarrow c$, where a and c are the antecedent and consequent of the rule, respectively, and where a is a conjunction of premises written as $p_1 \wedge ... \wedge p_n$. The premises of a GDR rule are classified into two types: *dataspace premises* and *external source premises*. A dataspace premise describes conditions over a Grapple dataspace [2], e.g. a user data repository, in the form of a pattern-based Grapple Query [2]. This will provide higher-level abstraction of the condition definition in the premise. An external source premise specifies conditions in the form of basic graph patterns over an external data source, e.g. background knowledge, accessible through its SPARQL endpoint. As each Grapple dataspace has its own SPARQL endpoint, it is also possible to define an external source premise over a Grapple dataspace. The consequent describes knowledge modeled as Grapple statements that will be derived if all the premises hold.

Let us consider a simple example of using GDR to enrich user data. Suppose an adaptive application wants to select the language in which it presents pages to its user based on the country where the user lives. Unfortunately, the application has only information about the city where the user lives in her user profile. To address this, the administrator of the application defines a GDR rule as depicted in Figure 1 over the application's own data space and background knowledge. This rule specifies that if the name of the city where a user lives is *cityName* and the city is located in a country with name *countryName*, then the name of the country where the user lives is *countryName*. The relation between city and country can be discovered by exploiting knowledge available in an external data source, namely, GeoNames. The rule not only reasons over the user data that the application has, but also extends and enriches it with knowledge from external data sources. It also provides a view-based query mechanism to the application, as the user's country information is dynamically computed and available from another external data source. With this rule language, applications can easily integrate user data from other applications they have access to,

[4] http://www.grapple-project.org/

or add external data sources, by modifying the rules without changing the applications. Thus, it provides flexibility to the administrators of adaptive applications.

```
01 <gdr:rule xmlns:gc="http://www.grapple-project.org/grapple-core/"
02      xmlns:gdr="http://www.grapple-project.org/grapple-derivation-rule/"
03      xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
04      name="Get User's Country"  ... >
05  <gdr:premise dataspace="http://www.grapple-project.org/gumf/ds/1">
06     <gc:subject>?user</gc:subject>
07     <gc:predicate rdf:resource="http://profile.org/city" />
08     <gc:object>?cityName</gc:object>
09  </gdr:premise>
10  <gdr:premise sourceURI="http://geonames.org/" ... >
11     <gdr:pattern>?city geo:name ?cityName</gdr:pattern>
12     <gdr:pattern>?city geo:featureClass geo:P</gdr:pattern>
13     <gdr:pattern>?city geo:inCountry ?countryURI</gdr:pattern>
14     <gdr:pattern>?country geo:featureClass geo:A</gdr:pattern>
15     <gdr:pattern>?country geo:inCountry ?countryURI</gdr:pattern>
16     <gdr:pattern>?country geo:name ?countryName</gdr:pattern>
17  </gdr:premise>
18  <gdr:consequent dataspace="http://www.grapple-project.org/gumf/ds/1">
19     <gc:subject>?user</gc:subject>
20     <gc:predicate rdf:resource="http://profile.org/country" />
21     <gc:object>?countryName</gc:object>
22  </gdr:consequent>
23 </gdr:rule>
```

Fig. 1 An Example of a GDR Rule in XML Syntax (partial view)

3 Demonstration Overview

Our demonstration showcases the functionality of GDR in integrating, interlinking, and enriching user data. For this, we use the GDR implementation that has been integrated into GUMF, as part of the research performed in the EU project GRAPPLE. In the demonstration, we will: (1) show how GDR is used for the integration between distributed user data, (2) demonstrate the enrichment of user data with data published as Linked Data on the Web, (3) showcase the view-based query mechanism offered by GDR, and (4) demonstrate the flexibility that GDR offers in integrating, interlinking, and enriching user data.

Acknowledgments. This work was partially supported by the EU FP7 project GRAPPLE (Generic Responsive Adaptive Personalized Learning Environment)

References

1. Leonardi, E., Abel, F., Heckmann, D., Herder, E., Hidders, J., Houben, G.J.: A Flexible Rule-Based Method for Interlinking, Integrating, and Enriching User Data. In the Proc. of ICWE 2010, Vienna, Austria, July, 2010.
2. Abel, F., Heckmann, D., Herder, E., Hidders, J., Houben, G.J., Krause, D., Leonardi, E., van der Sluijs, K.: A Framework for Flexible User Profile Mashups. In Proc. of APWEB 2.0 2009 Workshop in conjunction with UMAP 2009, Trento, Italy, June, 2009.
3. Rule Markup Language Initiative. Rule Markup Language (RuleML). http://ruleml.org/
4. Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., Dean, M.: SWRL: A Semantic Web Rule Language - Combining OWL and RuleML. W3C Member Submission, 2004.