Web Semantics: Science, Services and Agents on the World Wide Web 6 (2008) 266–273

Contents lists available at ScienceDirect

Web Semantics: Science, Services and Agents on the World Wide Web

journal homepage: www.elsevier.com/locate/websem

Revyu: Linking reviews and ratings into the Web of Data

Tom Heath [a,b,*], Enrico Motta [b]

[a] Talis Information Limited, Knights Court, Solihull Parkway, Birmingham B37 7YB, United Kingdom
[b] Knowledge Media Institute, The Open University, Walton Hall, Milton Keynes MK7 6AA, United Kingdom

ARTICLE INFO

Article history:
Received 15 May 2008
Received in revised form 22 August 2008
Accepted 15 September 2008
Available online 23 October 2008

Keywords:
Semantic Web
Linked Data
Reviews
Ratings
Revyu

ABSTRACT

Revyu is a live, publicly accessible reviewing and rating Web site, designed to be usable by humans whilst transparently generating machine-readable RDF metadata for the Semantic Web, based on user input. The site uses Semantic Web specifications such as RDF and SPARQL, and the latest Linked Data best practices to create a major node in a potentially Web-wide ecosystem of reviews and related data. Throughout the implementation of Revyu design decisions have been made that aim to minimize the burden on users, by maximizing the reuse of external data sources, and allowing less structured human input (in the form of Web 2.0-style tagging) from which stronger semantics can later be derived. Links to external sources such as DBpedia are exploited to create human-oriented mashups at the HTML level, whilst links are also made in RDF to ensure Revyu plays a first class role in the blossoming Web of Data. In this paper we document design decisions made during the implementation of Revyu, discuss the techniques used for linking Revyu data with external sources, and outline how data from the site is being used to infer the trustworthiness of reviewers as sources of information and recommendations.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Reviews and ratings are widely available on the Web and are one major form of 'user-generated content' that has become associated with 'Web 2.0' [25]. However, despite the availability of reviews and ratings through APIs such as the Amazon Associates Web Service [2], this data remains largely isolated in 'silos', and described in formats that hinder its integration and interlinking with data from other sources. This presents considerable barriers to the aggregation of all reviews of a particular item from across the Web, as an item reviewed in one silo cannot easily be associated with the same item reviewed elsewhere. As has been recognised by previous authors [15,16], the Semantic Web, or Web of Data, provides a technological platform with which to overcome this problem and Revyu is a significant and concrete step towards realising a solution.

Revyu is a live, publicly usable and well used reviewing and rating Web site, launched in November 2006 and available at http://revyu.com/. The site combines an approachable interface for the creation of reviews by human users with a range of APIs through which Semantic Web applications can access machine-readable data for reuse in third-party applications. The site has been developed using Semantic Web technologies and standards such as RDF [19] and SPARQL [26], and according to Linked Data principles [5] and best practices [7]. These features enable Revyu to readily consume data from external services for the creation of human-oriented 'mashups' while also seeding an ecosystem of interlinked review and rating data on the Web that is helping to bootstrap the Semantic Web as a whole. In the following sections we will describe Revyu in more detail, examine these human- and machine-oriented characteristics and discuss many of the underlying design decisions.

2. Revyu compared to conventional Web APIs

Revyu allows people to review and rate things simply by filling in a Web form. This style of interaction with the site will be familiar to those who have written reviews on sites such as Epinions [12], Amazon [1] or TripAdvisor [28]. Whilst this functionality is not especially novel, as a reviewing application Revyu improves significantly over other work in the area in a number of ways.

First, Revyu is not a data 'silo' that locks data away for 'safe keeping'. Instead, reviews and ratings created within Revyu are exposed in a reusable, machine-readable data format, RDF. This contrasts with sites that collect data from users but only republish it in HTML, thereby masking the structure created by the author of the review and no doubt still represented in the underlying database. The RDF data model represents a more flexible mechanism for publishing

* Corresponding author at: Talis Information Limited, Knights Court, Solihull Parkway, Birmingham B37 7YB, United Kingdom. Fax: +44 870 400 5001.
E-mail addresses: tom.heath@talis.com (T. Heath), e.motta@open.ac.uk (E. Motta).

1570-8268/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.websem.2008.09.003

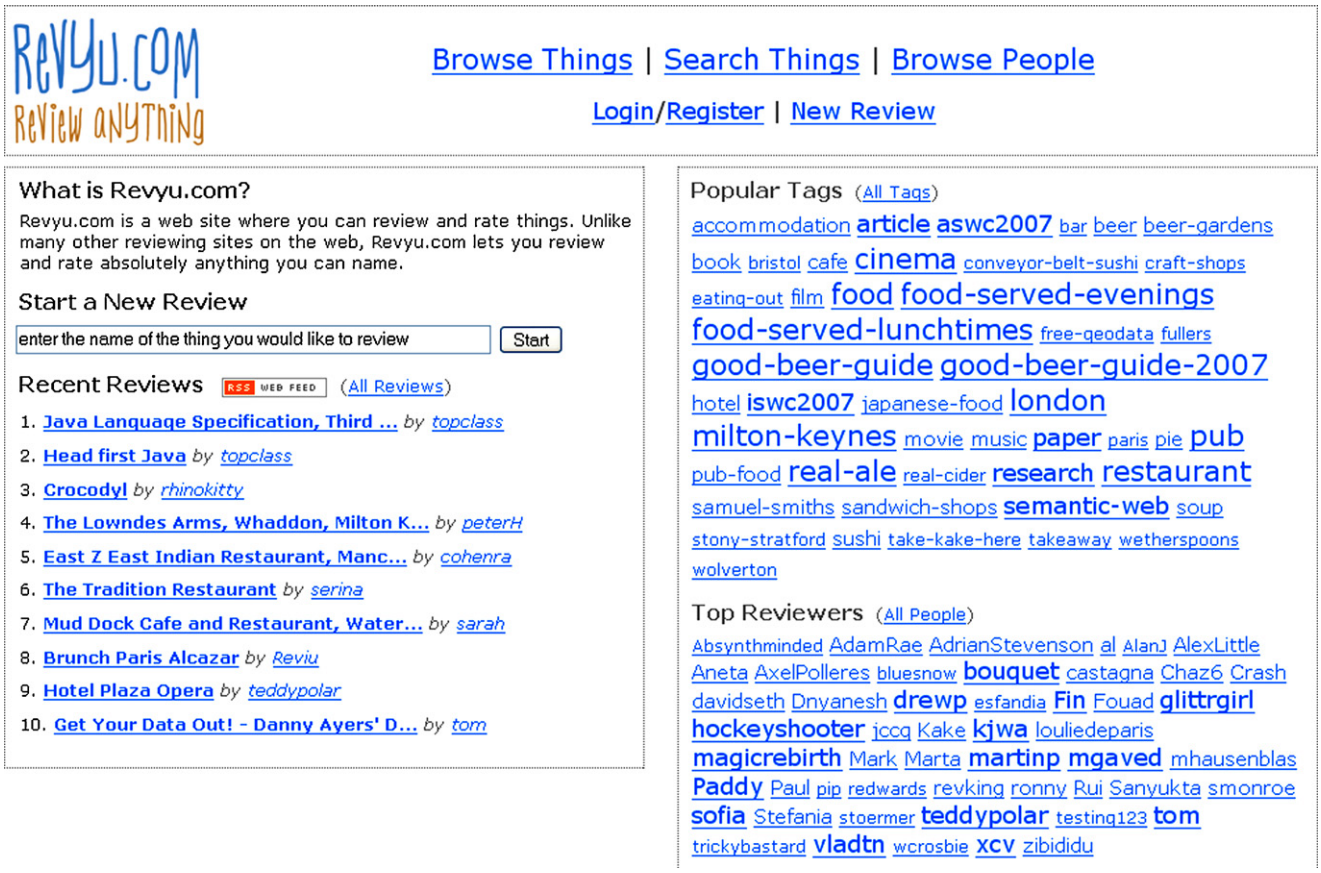

Fig. 1. The Revyu.com home page.

structured data than approaches such as *microformats*,[1] which are not supported by a common underlying data model and consequently present greater challenges for those wishing to parse and consume published data.

Secondly, the data access mechanisms provided by Revyu improve upon the APIs of sites such as Amazon. Publishing data in RDF allows for easier merging of data from disparate sources, as heterogeneous data can be combined in one document without the document as a whole needing to conform to a single schema. Golbeck's *FilmTrust* [15] is noteworthy as one of the first applications to make review data available on the Web in RDF. However it is limited to the domain of films and does not provide a query interface to the underlying data via languages such as SPARQL. The Revyu SPARQL endpoint provides developers with greater flexibility in querying the underlying data set than is generally possible with conventional Web APIs.

Thirdly, Revyu takes a Linked Data [5] approach to publishing reviews and ratings on the Web. The technical aspects of Revyu as a Linked Data application will be described in detail in later sections—at this stage it is sufficient to outline the benefits of the approach:

1. All entities within the Revyu site are addressable over the Web, allowing these to be referenced from other online data sets. In the context of Revyu, third parties may use this capability to indicate (dis-)agreement with a particular review, or where there may be a conflict of interest that compromises the credibility of a particular reviewer (as reported in [31]) this can be highlighted. Crucially, this additional information can be published elsewhere on the Web in RDF, simply referencing the appropriate items on Revyu. This avoids the creation of a single silo in which all information must be located.
2. Duplication of data across data providers is reduced. There is less need for Revyu to maintain local copies of data about reviewed items, and incur the associated data management overhead, as RDF statements can be used to connect review data with richer item descriptions in other locations on the Web. This has the effect of encouraging data to remain published and managed by the authoritative source.
3. Data integration can be performed once, and reused many times. A data publisher need only expend the effort once to link entities in their data with those in external data sets, and publish these links as RDF statements for consumption by third parties. This removes the need for each data consumer to perform their own data integration, the results of which may be locked within application code used to create 'mashups' and therefore not available for reuse by others.

Lastly, Revyu takes an open world perspective on the reviewing process by not constraining users to reviewing items from a fixed and pre-selected database. Anything a user can name can be reviewed, and consequently reviewers are not restricted to reviews and ratings in one domain. By supplying links related to the item they are reviewing, users enable disambiguation and linking of reviewed items through inverse functional properties such as `foaf:homepage`.

[1] http://microformats.org/.

Download English Version:

https://daneshyari.com/en/article/557906

Download Persian Version:

https://daneshyari.com/article/557906

Daneshyari.com