

# Electronic Commerce Meets the Semantic Web

Jelena Jovanovic, *University of Belgrade, Serbia*  
Ebrahim Bagheri, *Ryerson University, Canada*

**The intersection of Semantic Web technologies and business-to-consumer (B2C) e-commerce offers benefits for both online retailers and customers. The authors' framework highlights why and how the adoption of Semantic Web technologies can enhance B2C applications and platforms.**

**B**oth the spread of Internet services and the increase in their quality have immensely contributed to how electronic trade is perceived and performed around the world—a phenomenon widely known as *e-commerce*. According to Statista, worldwide business-to-consumer (B2C) e-commerce sales surpassed US\$2.3 trillion in 2014. Current statistics show that 40 percent of Internet users have made at least one e-commerce transaction, a number that amounts to close to 1 billion users ([www.statista.com/markets/413/e-commerce/](http://www.statista.com/markets/413/e-commerce/)). Advances in mobile and social commerce promise even higher penetration rates for e-commerce in the following years.

As the e-commerce domain matures and expands, many new challenges must be addressed, such as efficient customer engagement, trust management, privacy concerns, and internationalization. In this article, we look at six challenges that pertain to the efficient organization and management of e-commerce (meta)data. Several leading retailers in the e-commerce domain, such as Sears and Best Buy, have already faced these challenges and attempted to address them through community standardization efforts that focus on using Semantic Web technologies. We examine such challenges from a more comprehensive perspective and suggest how they can be addressed by exploiting the developments that

have been made at the intersection of Semantic Web technologies and e-commerce.

## E-Commerce Challenges

The first challenge (C1) is that existing product data are not suitable for automated processing. Rather than have access to structured (meta) data about the products they are offering, many e-commerce vendors receive only free text descriptions. Thus, being able to publish structured product data is conditioned on the availability of tools that support the consistent extraction of that data from free text, or the provisioning of structured data by product manufacturers. However, even when well-structured product data are available, they are often presented using Web technologies such as CSS and JavaScript, which allow for visually appealing rendering of product information, but do not preserve product data's structure and semantics.

The second challenge (C2) is that product data often lack interoperability in terms of both syntax and semantics. The lack of syntactic interoperability stems from the fact that different retailers or shops use different schemas to describe products they offer. The absence of semantic interoperability results from the use of different vocabularies for describing the values of product attributes. This makes the integration and semantic alignment of product data originating from diverse online shops inefficient, and limits the potential of published product data (for instance, it impedes efficient aggregate search over product data).<sup>1</sup>

The third challenge (C3) is insufficient use of unique product identifiers. Efficient product data integration requires product identity resolution—that is, identifying whether two or more product descriptions, found on different shopping sites, refer to the same product. This challenge is partially addressed by increasing the use of unique product identifiers enforced by some vertical search engines (such as Google Shopping) and major online retailers (Amazon and eBay, for example).<sup>2</sup> Still, more widespread adoption of such identifiers or further improvements of techniques for automated product matching<sup>3,4</sup> are needed to fully address this challenge.

The fourth challenge (C4) is the heterogeneity of product category taxonomies. The problem of semantic interoperability is further exacerbated by the diversity of product category schemes or taxonomies used by different shopping sites. This

heterogeneity implies that to aggregate product data from disparate sources, you need an efficient mapping of product taxonomies. This is a real challenge because the semantics of concepts or categories in a taxonomy are not directly accessible to computers, but rather have to be inferred (for example, using their context in the concepts hierarchy) before any mapping can be done.<sup>5</sup>

The fifth challenge (C5) is incomplete, inconsistent, or outdated product descriptions. There is often a discrepancy between the richness and variety of product features data offered by manufacturers (known as *product master data*) and product features available to and exposed by online retailers. As a consequence, product descriptions in online shops tend to be incomplete, inconsistent, or outdated.<sup>6</sup>

The final challenge (C6) is the weakness of current product recommender systems. A great majority of recommender systems in the e-commerce domain are based on collaborative filtering, and thus rely highly on user ratings. This

**There is often a discrepancy between the richness of product features data offered by manufacturers and product features available to online retailers.**

makes them susceptible to the cold start and to data sparsity problems.<sup>7</sup>

## Advancing the State of the Practice

One major advantage of Semantic Web technologies is an increase in the quality of product data—that is, diversity, completeness, and accuracy. This can be achieved by

- gathering and semantically aligning product data from different online retailers,
- exposing product master data on the Web as Linked (Open) Data,<sup>8</sup> or
- enriching product data with other kinds of data relevant for making purchases—for example, users' intent and location data.

High-quality product data coupled with semantic technologies for data querying, analysis,

Table 1. Semantic Web standards and technologies juxtaposed with benefits.

Observable benefits vs. enabling standards/technologies		Specifications for embedding semantic markup (L1a)*	Standard data-exchange formats (L1b)	Vocabularies for describing products, offers, and stores (L2a)
Online retailers	Rich Snippets lead to higher click-through rates	Functional (essential) feature	N/A	Functional (essential) feature
	Increased visibility in vertical search engine results pages (SERPs)	Quality attribute (keeping data current)	Functional (essential) feature	Functional (essential) feature
	Increased return on advertising spending	N/A	Functional (essential) feature	Functional (essential) feature
	Seamless, just-in-time introduction of occasion-specific product categories	Quality attribute (keeping data current)	Functional (essential) feature	Functional (essential) feature
Online shoppers	Search of niche and long-tail products	Quality attribute (keeping data current)	Functional (essential) feature	Functional (essential) feature
	Faceted product search at Web scale	Quality attribute (keeping data current)	Functional (essential) feature	Functional (essential) feature
	Better, more personalized product recommendation	Quality attribute (keeping data current)	Functional (essential) feature	Functional (essential) feature

\* L1–L5 refer to layers of the technology stack as shown in Figure 1.

and reasoning can lead to a number of “tangible” benefits for both online retailers and customers. Benefits for retailers are as follows:

- Rich Snippets tend to lead to higher click-through rates (CTR). Products that are published on the Web with embedded structured data will appear in rich snippet format on Google and Bing search results pages. The experience of Best Buy and some other online retailers indicate that such Rich Snippets lead to a non-negligible increase in CTR.<sup>9</sup>
- Visibility increases in vertical search engine results pages. For instance, Google Shopping, one of the major vertical search engines in the retail domain, gives more visibility to products that are associated with rich structured data.<sup>2</sup>
- The return on advertising spending increases. For instance, one of the policies that Google Shopping applies to motivate retailers to provide high-quality product data consists of lowering the price for product ads for those retailers who supply such data.
- High-quality product data enables the seamless, just-in-time introduction of occasion-specific product categories. When product data are well structured and semantically rich, it is easy to de-

fine rules for selecting products based on a given set of requirements—for instance, particular kinds of products to be put on sale for a limited time period or products to be offered during the Christmas season as stocking stuffers (<http://jaymmymyers.tumblr.com/post/69512519550/spar-ql-these-are-the-numbers-were-looking-for>).

High-quality product data can also provide the following benefits to online customers or shoppers:

- It enables search of niche and long-tail products. Rich and accurate product descriptions enable users to do more sophisticated searches, including search for products with specific features. This also allows for efficient search of long-tail products.
- Faceted product search is enabled at Web scale. Although the use of facets (that is, filters) is quite handy for reducing shoppers’ information overload,<sup>10</sup> faceted product search is currently limited to the level of individual shopping sites and vertical search engines. The use of strong product identifiers or product-matching techniques (as we discuss later) allows for the integration of product data from disparate sources, and thus leads to faceted product search at Web scale.<sup>3</sup>

Vocabularies for describing users and their shopping history (L2b)	Strong product identifiers (L3)	Product ontologies/catalogs (L4)	Semantic technologies for storage, search, and product data manipulation (L5)
N/A	Quality attribute (product data integration)	Quality attribute (rich product display)	Quality attribute (flexibility, adaptability)
N/A	Functional (essential) feature	Measurable performance improvement	Quality attribute (flexibility, adaptability)
N/A	Functional (essential) feature	Measurable performance improvement	N/A
N/A	Quality attribute (product data integration)	Quality attribute (highly customized product offers)	Quality attribute (flexibility, adaptability)
Quality attribute (personalization)	Measurable performance improvement	Measurable performance improvement	Quality attribute (flexibility, adaptability)
Quality attribute (personalization)	Measurable performance improvement	Measurable performance improvement	Quality attribute (flexibility, adaptability)
Functional (essential) feature	Measurable performance improvement	Measurable performance improvement	Quality attribute (flexibility, adaptability)

- Better, more personalized product recommendations are possible. Some systems<sup>11</sup> rely on users' product ownership data along with semantically rich descriptions of product relationships to offer personalized recommendations; for instance, such systems could recommend accessories or spare parts for products owned by the user.

Semantic Web standards and technologies allow for systematically addressing the challenges associated with product data management and use (C1–C6), for the benefit of both online retailers and customers.

### Enabling Standards and Technologies

Table 1 presents standards and technologies that allow for fulfilling the benefits outlined in the previous section. To denote the relationship between a certain technology and an associated benefit, we rely on the following software engineering terminology:

- *Functional (essential) feature* denotes that a feature or technology is essential for the provision of the stated benefit.
- *Measurable performance improvement* indicates that the given feature or technology can quantitatively augment the core benefit. These are

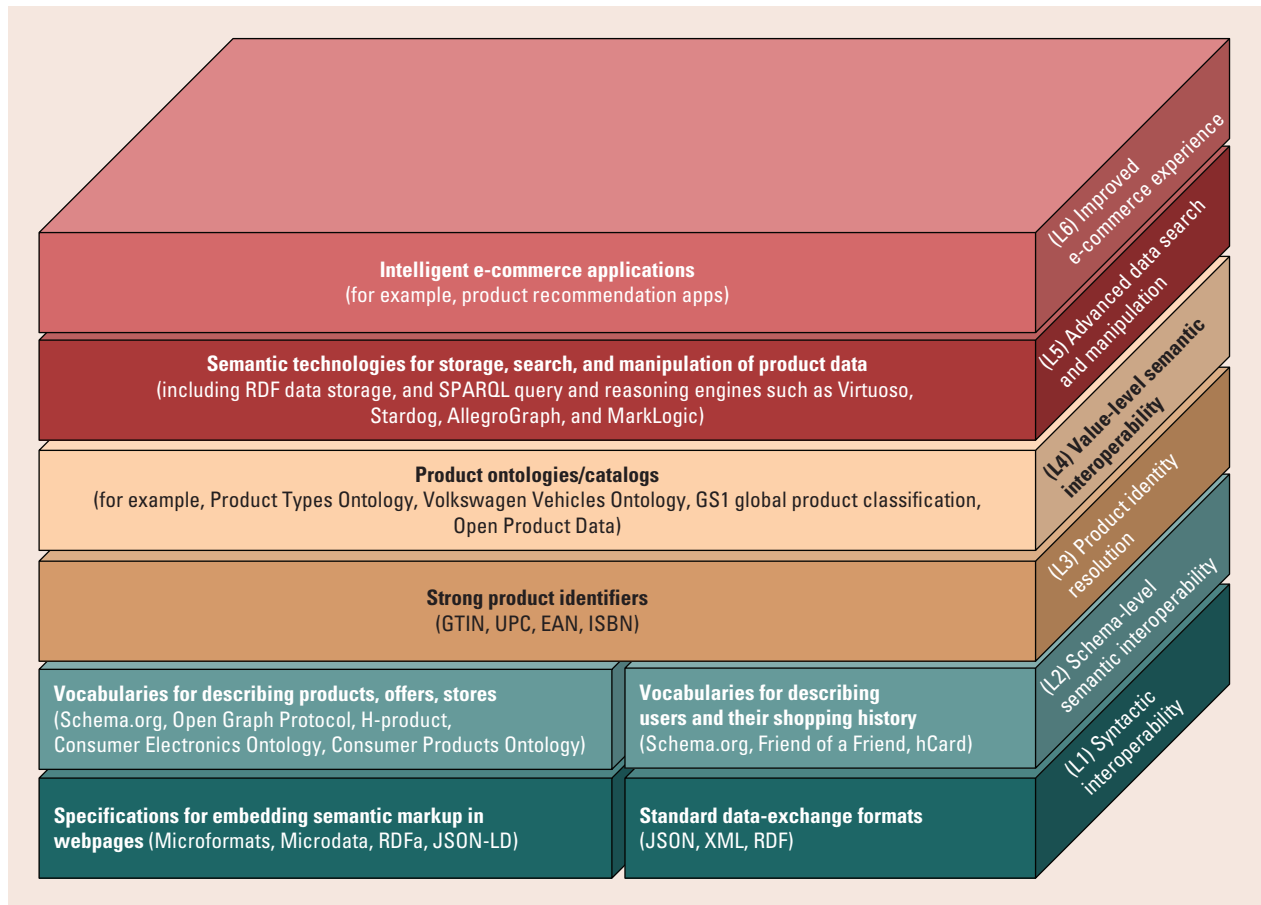
cases in which the adoption of the respective technology is not mandatory, but if adopted, it would lead to a measurable increase in the respective benefit.

- *Quality attribute (...)* denotes that a feature or technology, if present, would qualitatively augment the corresponding core benefit, or would augment the system's quality. This descriptor is used to mark cases in which the core functionality (that is, the benefit) could be provided without the use of the respective technology, but if the technology is used, it could further augment the stated benefit by adding some qualitative difference (such as product search personalization).

As Figure 1 illustrates, standards and technologies for advancing e-commerce form a technology stack in which each layer (L1–L5) builds on the previous one, thus forming a structure that eventually leads to intelligent e-commerce applications (L6). In the following, we briefly review each layer of this stack.

### Essential Specifications and Standards

The bottom-most layer (L1) comprises specifications and standards that are essential for the au-



**Figure 1.** Technology stack for semantically rich e-commerce applications. A list of links to all the specifications, technologies, vocabularies, and systems mentioned in the figure is available at [http://ls3.rnet.ryerson.ca/?page\\_id=685](http://ls3.rnet.ryerson.ca/?page_id=685).

tomated processing of product data published on the Web (challenge C1). These include

- standard formats for data exchange on the Web, namely XML, JavaScript Object Notation (JSON), and Resource Description Framework (RDF); and
- specifications for embedding semantic markup in webpages, including Microformats, RDFa, Microdata, and JSON for Linked Data (JSON-LD).

The latest Common Crawl corpus (Winter 2014; <http://commoncrawl.org>) consists of 2.01 billion HTML pages collected from more than 15.68 million pay-level domains (PLDs). An analysis of this corpus shows that 30 percent of pages and 17 percent of PLDs covered by the crawl use at least one of the three standard markup formats (Microdata, RDFa, and Microformats).<sup>12</sup> Microformats are the most represented in the corpus, with a presence in 65.14 percent of all

the examined PLDs that use any markup format. RDFa contributes 14.32 percent, and Microdata covers 20.54 percent of all the crawled PLDs that contain structured data. Notably, Microdata has achieved significant growth as compared to 2012<sup>13</sup>; its adoption, as measured at the PLD level, increased six-fold. This is an expected finding, given that Microdata is promoted by the major search engines and has the best tool support (see Table 2). Microformats likely dominate at present because these data formats were the first to appear and are recognized by all the major search engines (<http://microformats.org/wiki/search-engines>).

Note that some of the stated benefits of enriching e-shops with semantic markup (Table 1) can also be achieved by providing rich structured data in the product feeds used by search engines. However, embedding semantic markup in webpages is a more viable option because it does not require the creation of a different product feed for

**Table 2. An overview of tools aimed at facilitating the use of Semantic Web technologies.**

Type of tool support	Example tools
Content management systems (CMSs) offering native support for embedding structured data in webpages (L1a and L2a)	Drupal ( <a href="https://www.drupal.org">https://www.drupal.org</a> ) Webnodes ( <a href="http://www.webnodes.com">www.webnodes.com</a> )
Extensions (add-ons) of major CMSs and shopping platforms for embedding structured data in webpages (L1a and L2a)	All in One Schema.org Rich Snippets extension for WordPress ( <a href="https://wordpress.org/plugins/all-in-one-schemaorg-rich-snippets/">https://wordpress.org/plugins/all-in-one-schemaorg-rich-snippets/</a> ) Rich Snippets add-on for Yahoo Stores ( <a href="http://www.ytimes.com/rich-snippets-for-yahoo-stores.html">www.ytimes.com/rich-snippets-for-yahoo-stores.html</a> ) Numerous extensions for Joomla ( <a href="http://www.microdataforjoomla.com">www.microdataforjoomla.com</a> )
Tools for webmasters and website owners aimed at supporting the tasks of embedding and managing structured data (L1a and L2a)	Data Highlighter ( <a href="https://support.google.com/webmasters/answer/2774358?hl=en">https://support.google.com/webmasters/answer/2774358?hl=en</a> ) Structured Data Testing Tool ( <a href="https://search.google.com/structured-data/testing-tool">https://search.google.com/structured-data/testing-tool</a> ) Bing Markup Validator ( <a href="http://www.bing.com/toolbox/markup-validator">www.bing.com/toolbox/markup-validator</a> ) Structured Data Dashboard ( <a href="https://support.google.com/webmasters/answer/2650907?hl=en">https://support.google.com/webmasters/answer/2650907?hl=en</a> )
Tools for Web developers who want to leverage the potential of semantic technologies but are not familiar with RDF, SPARQL, product vocabularies, and related technologies (L1b, L2, L5)	Elda: Java-based implementation of linked data API ( <a href="https://github.com/epimorphics/elda">https://github.com/epimorphics/elda</a> ) Specialized software libraries (such as GR4PHP <sup>6</sup> ) for working with structured and semantically rich product data
Tools for the development and maintenance of domain-specific (such as product) ontologies (L2 and L4)	Semaphore Ontology Manager ( <a href="http://goo.gl/SYCI4o">http://goo.gl/SYCI4o</a> ) TopBraid taxonomy and ontology management tools ( <a href="http://goo.gl/brmZQn">http://goo.gl/brmZQn</a> ) PoolParty taxonomy and thesaurus management tools ( <a href="http://www.poolparty.biz/portfolio-item/thesaurus-management/">www.poolparty.biz/portfolio-item/thesaurus-management/</a> )
Tools for publishing product master data as linked data (L3 and L4)	BMEcat2goodrelations mapping tool ( <a href="https://code.google.com/p/bmecat2goodrelations/">https://code.google.com/p/bmecat2goodrelations/</a> )

each online shop or product search engine.<sup>3</sup> Even Google Shopping is encouraging the use of semantic markup (that is, Microdata) so that its crawlers can pull the data directly from retailers' websites and thus ensure that the product information displayed on Google Shopping is up-to-date.

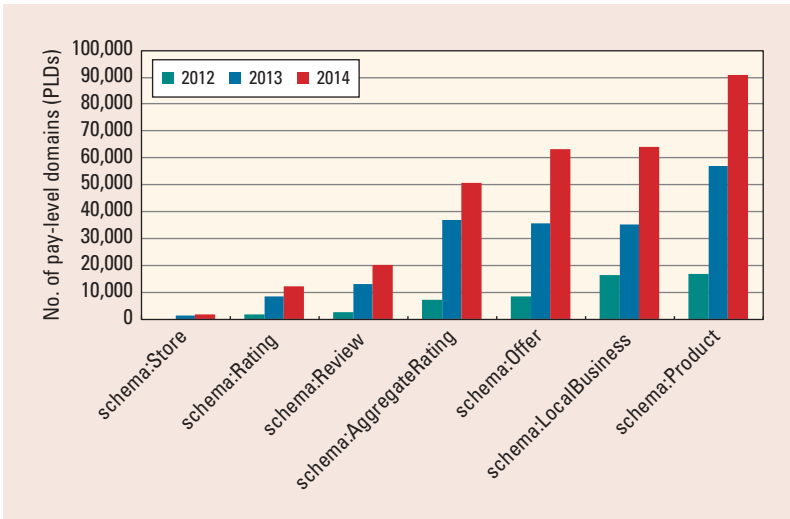
### Vocabularies

The second layer in the technology stack (L2) comprises vocabularies for describing products, offers, and stores. These vocabularies allow for establishing syntactic and semantic interoperability of product data (challenge C2). The most widely used vocabularies include Schema.org and the Open Graph Protocol (OGP). Whereas OGP allows for a very basic product description, Schema.org offers a detailed vocabulary for the description of various kinds of products, offers, and retailers (see <http://schema.org/Product>, <http://schema.org/Offer>, and <http://schema.org/LocalBusiness>, respectively). Figures 2 and 3 depict a positive trend in the adoption of Schema.org classes and properties related to e-commerce.

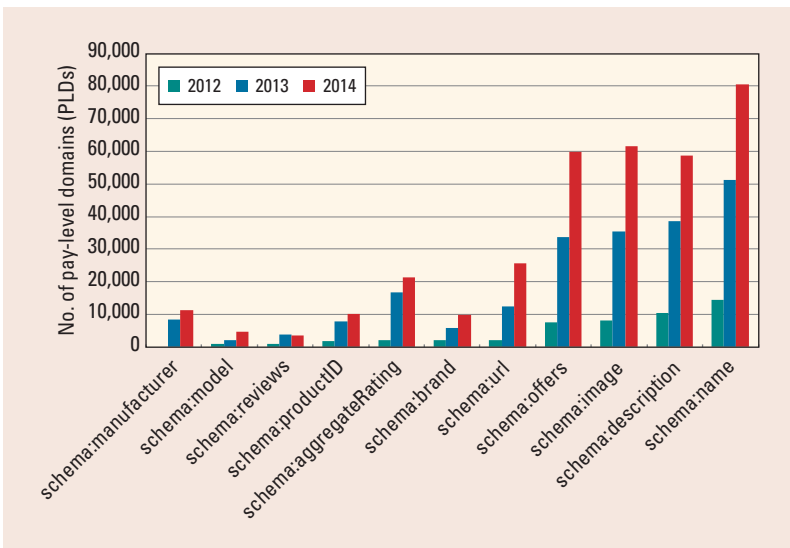
The charts are based on the data published by the Web Data Commons project.<sup>12,13</sup> They present the adoption level as the number of PLDs with entities of the Schema.org classes related to e-commerce (Figure 2), and properties associated with entities of either `schema:Product` or `schema:IndividualProduct` classes (Figure 3). To be able to compare the adoption levels over time—namely, in the three consecutive years for which data are available (2012, 2013, and 2014)—we focused only on (e-commerce related) classes and properties that have been in use throughout the examined time period.

There are also many product and service-specific vocabularies, some of which are listed in Figure 1 (L2). In addition, one study<sup>14</sup> has demonstrated a semiautomated approach for generating product-specific vocabularies by extending the GoodRelations vocabulary with product-specific knowledge originating from Freebase ([www.freebase.com](http://www.freebase.com)).

Vocabularies for describing users and their shopping histories (L2) are essential for the



**Figure 2.** Adoption trends in Schema.org classes for adding semantic markup to e-commerce websites. The data originate from Web Data Commons reports<sup>12</sup> and present the number of pay-level domains (y-axis) with entities of e-commerce-related Schema.org classes (x-axis).



**Figure 3.** Adoption trends in Schema.org properties for embedding semantic product descriptions in e-commerce websites. The data come from Web Data Commons reports<sup>12</sup> and present the number of pay-level domains (y-axis) that use particular Schema.org property for product descriptions (x-axis).

provision of personalized product recommendations (challenge C6), and highly desirable if product search is to be improved through personalization. Schema.org offers a comprehensive vocabulary for describing people; it also allows for representing items that a person owns or has owned, and enables the tracking of ownership details.<sup>11</sup> However, technical support

for user data representation, storage, and exchange has to be complemented with effective strategies for protecting users' privacy while still providing online retailers with sufficient information for personalization purposes. Retailers might overcome this challenge by being transparent about the intended use of the acquired user data, and thus develop trust and a positive reputation among their customers.<sup>11</sup>

### Strong Product Identifiers

Strong product identifiers (L3) are required for overcoming the challenge of how to uniquely identify products across different e-commerce websites and the Web in general (challenge C3). For some of the outlined benefits (Table 1), strong product identifiers are essential features. In other cases, if used, they tend to lead to measurable performance improvements by reducing, if not fully eliminating, incorrect product matching. In addition, if product identifiers are published in accordance with Linked Data principles,<sup>8</sup> they can further facilitate product search, discovery, and integration.<sup>1</sup> However, the aforementioned analysis of the 2014 Common Crawl corpus shows that globally unique product identifiers are largely missing at present<sup>12</sup>—namely, only 1.1 percent of products described using Schema.org have a Global Trade Item Number, and 1.05 percent have a manufacturer product number. To compensate for this lack of strong product identifiers, today's intelligent e-commerce applications leverage different identity resolution methods

that often rely on text analysis and string similarity metrics,<sup>3</sup> machine learning techniques,<sup>15</sup> or evolutionary algorithms.<sup>4</sup>

### Product Ontologies or Catalogs

The use of product ontologies—that is, semantically rich and machine-processable product catalogs (L4)—allows for consistent and unambiguous

product descriptions, thus addressing challenges C2 and C4, and leading to high-quality product data. This in turn increases the chance that major search engines will display the given product in the rich snippet format (<http://goo.gl/rxAjcm>), and also positively influences product visibility in vertical SERPs. In addition, semantically enriched product descriptions greatly facilitate the provision of advanced customer services such as custom-made offers or product recommendations based on product-specific features, or faceted search based on the features specific to a given product category (Table 1). However, this layer is still underdeveloped given that product ontologies that formally present and semantically interlink products within a catalog are still largely missing. A recent GS1 initiative broadly aimed at advancing the state of the practice in e-commerce promises to significantly contribute to the development of this technology layer. The foundation for this development is Global Product Classification (GPC)—GS1's standard product classification scheme, which is already used by thousands of companies worldwide. Its wide adoption, regular updates, and public availability make GPC a highly suitable basis for the development of a product ontology. This is exactly the direction that GS1 is currently considering for further GPC development.<sup>16</sup>

GS1 Source or Trusted Source of Data is a GS1 initiative focused on establishing a single, trusted source of product information controlled by brand owners, and thus on overcoming the problem of inconsistent product presentation across the Web in terms of both product features and their values (challenge C5). This initiative is currently examining the adoption of Semantic Web technologies.<sup>16,17</sup>

### **Technologies for Data Storage, Search, and Manipulation**

The fifth layer, semantic technologies for storage, search, and manipulation of product data (L5) includes RDF-based technologies (RDF Schema and the Web Ontology Language) for explicit representation of data semantics, which in turn allows for sophisticated search and automated reasoning over the available data. Search is powered by the SPARQL data query and manipulation language, which allows for advanced, semantic search of stored data. Today's RDF

triple stores, some of which are shown in Figure 1 (L5), support enterprise-level data integration<sup>18</sup> and sophisticated cross-department or sector data search and reasoning.<sup>19</sup> One of the main advantages of these data stores over traditional relational databases is in the flexibility of the underlying data model, which allows for rapid adaptation of the data store to continuously changing market demands.<sup>19</sup> Although this flexibility is also a feature of NoSQL databases, RDF triple stores offer further advantages above and beyond both relational and NoSQL databases, due to the fact that the semantics of data items and their relationships are made explicit in the data model. This allows sophisticated searches that leverage not only relationships stored in the database, but also those that can be automatically inferred from existing ones. Such features have a strong potential for overcoming some of the challenges faced by today's B2C e-commerce applications (challenges C4, C5, and C6).

**One of the main advantages of data stores over traditional relational databases is in the flexibility of the underlying data model.**

### **Intelligent E-Commerce Applications**

At the top of the technology stack (L6) are intelligent e-commerce applications that use technologies from the lower layers (L1–L5) to offer retailers or customers the previously described benefits. These applications might feature, for instance, faceted search at Web scale,<sup>3</sup> search of long-tail<sup>9</sup> or highly customizable products,<sup>20</sup> or personalized product recommendations.<sup>11</sup> Applications of this type have only recently started to emerge, which is expected considering the novelty of the underlying technologies.

### **Available Tool Support**

Table 2 gives an overview of tools that could be used to seamlessly leverage the benefits offered by Semantic Web technologies for e-commerce platforms and websites. By hiding the intricate details of the proposed technology stack (Figure 1), the presented tools facilitate the



adoption of these novel technologies. Note that Table 2 does not aim to provide an exhaustive list of all the available tools, but rather to exemplify each tool type with some well-known tools of that type.

**O**ur proposed framework could be highly beneficial to decision makers in the e-commerce domain because it can inform and facilitate their decisions regarding the adoption of new technologies, and advance their businesses through the realization of the benefits that these technologies can bring about.

In particular, our framework suggests that the technologies qualified as a “functional (essential) feature” in Table 1 should be the first to be considered and adopted because they are required for all the envisioned advantages for both retailers and shoppers. Specifically, standard data exchange formats (L1b) and vocabularies for describing products, offers, and stores (L2a) seem to be the most significant for achieving the considered benefits and should be adopted first. These technologies have well-developed and mature tool support (Table 2), which can facilitate and speed up the adoption process. The next ones to be considered for adoption include strong product identifiers (L3), specifications for embedding semantic markup in webpages (L1a), and product ontologies and catalogs (L4). Although adoption is not urgent for the other layers, they could be influential for the development of advanced e-commerce applications. ■

## References

1. S. Bouzidi et al., “Product Information Retrieval on the Web: An Empirical Study,” *Proc. 8th Int’l Conf. Knowledge Management in Organizations*, 2013, pp. 439–450.
2. B. Starr, “How Online Retailers Can Leverage Unique Identifiers & Structured Data,” *Search Engine Land*, 5 Dec. 2013; <http://searchengineland.com/unzip-the-secrets-behind-leveraging-unique-identifiers-structured-data-in-e-commerce-178692>.
3. D. Vandic, J.W. van Dam, and F. Frasincar, “Faceted Product Search Powered by the Semantic Web,” *Decision Support Systems*, vol. 53, no. 3, 2012, pp. 425–437.
4. P. Petrovski, V. Bryl, and C. Bizer, “Integrating Product Data from Websites Offering Microdata Markup,” *Proc. Companion Publication of the 23rd World Wide Web Conf.*, 2014, pp. 1299–1304.
5. S.S. Aanen, D. Vandic, and F. Frasincar, “Automated Product Taxonomy Mapping in an E-commerce Environment,” *Expert Systems with Application*, vol. 42, no. 3, 2015, pp. 1298–1313.
6. A. Stolz, B. Rodriguez-Castro, and Martin Hepp, “Using BMEcat Catalogs as a Lever for Product Master Data on the Semantic Web,” *Proc. 10th Extended Semantic Web Conf.*, 2013, pp. 623–638.
7. M.D. Ekstrand, J.T. Riedl, and J.A. Konstan, “Collaborative Filtering Recommender Systems,” *Foundations and Trends Human-Computer Interaction*, vol. 4, no. 2, 2011, pp. 81–173.
8. T. Heath and C. Bizer, *Linked Data: Evolving the Web into a Global Data Space*, Synthesis Lectures on the Semantic Web: Theory and Technology, Morgan & Claypool, 2011.
9. D. Sheldon, “RDFa: The Inside Story from Best Buy: An Interview with Jay Myers,” *Search News Central*, 7 Feb. 2011; <http://searchnewscentral.com/20110207129/Technical/rdfa-the-inside-story-from-best-buy.html>.
10. Y.C. Chen, R.A. Shang, and C.Y. Kao, “The Effects of Information Overload on Consumers’ Subjective State towards Buying Decision in the Internet Shopping Environment,” *Electronic Commerce Research & Application*, vol. 8, no. 1, 2009, pp. 48–58.
11. L. Török and M. Hepp, “Towards Portable Shopping Histories: Using GoodRelations to Expose Ownership Information to E-Commerce Sites,” *Proc. 11th Extended Semantic Web Conf.*, 2014, pp. 691–705.
12. C. Bizer, R. Meusel, and A. Primpeli, “Web Data Commons—RDFa, Microdata, and Microformat Data Sets,” extraction results from the Dec. 2014 Common Crawl corpus, Apr. 2015; <http://webdatacommons.org/structureddata/#toc3>.
13. R. Meusel, P. Petrovski, and C. Bizer, “The Web-DataCommons Microdata, RDFa, and Microformat Dataset Series,” *Proc. 13th Int’l Semantic Web Conf.*, 2014, pp. 277–292.
14. M. Dudáš et al., “Exploiting Freebase to Obtain GoodRelations-Based Product Ontologies,” *Proc. 15th Int’l Conf. E-Commerce and Web Technologies*, 2014, pp. 34–45.
15. A. Kannan et al., “Matching Unstructured Offers to Structured Product Descriptions,” *Proc. 17th ACM SIGKDD Int’l Conf. Knowledge Discovery and Data Mining (KDD)*, 2011, pp. 404–412.
16. A. Osborne, “What Linked Data and the Semantic Web Mean for Search,” *Essential E-Commerce*, 28 July 2014; [www.essentialretail.com/news/ecommerce/article/53d61aade84d-comment-what-linked-data-and-the-semantic-web-mean-for-search](http://www.essentialretail.com/news/ecommerce/article/53d61aade84d-comment-what-linked-data-and-the-semantic-web-mean-for-search).

17. J. Zaino, "GS1 Explores How Its Systems and Standards Will Fit into the Semantic Web," *Dataversity*, 23 Dec. 2013; [http://semanticweb.com/gs1-explores-systems-standards-will-fit-semantic-web\\_b41164](http://semanticweb.com/gs1-explores-systems-standards-will-fit-semantic-web_b41164).
18. M. Graube et al., "Linked Data as Integrating Technology for Industrial Data," *Int'l J. Distributed Systems and Technologies*, vol. 3, no. 3, 2012, pp. 40–52.
19. J. Breslin et al., "Semantic Web Computing in Industry," *Computers in Industry*, vol. 61, no. 8, 2010, pp. 729–741.
20. F.P. Servant and E. Chevalier, "Describing Customizable Products on the Web of Data," *Proc. World Wide Web Conf. Workshop on Linked Data on the Web*, 2013; <http://ceur-ws.org/Vol-996/papers/ldow2013-paper11.pdf>.

*Jovanovic received a PhD in information systems and technologies from the University of Belgrade. Contact her at [jeljov@fon.bg.ac.rs](mailto:jeljov@fon.bg.ac.rs); <http://jelenajovanovic.net>.*

*Ebrahim Bagheri is an associate professor at Ryerson University, Canada, where he leads the Laboratory for Systems, Software, and Semantics (LS3). His research interests include large-scale software reuse and Semantic Web technologies. Bagheri received a PhD in computer science from the University of New Brunswick, Canada. Contact him at [bagheri@ryerson.ca](mailto:bagheri@ryerson.ca).*

*Jelena Jovanovic is an associate professor in the Department of Software Engineering at the University of Belgrade, Serbia. Her research interests include semantic technologies, knowledge engineering, and technology-enhanced learn-*



Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.



## 2017 B. Ramakrishna Rau Award Call for Nominations

*Honoring contributions to the computer microarchitecture field*

**New Deadline: 1 May 2017**



Established in memory of Dr. B. (Bob) Ramakrishna Rau, the award recognizes his distinguished career in promoting and expanding the use of innovative computer microarchitecture techniques, including his innovation in compiler technology, his leadership in academic and industrial computer architecture, and his extremely high personal and ethical standards.

**WHO IS ELIGIBLE?** The candidate will have made an outstanding innovative contribution or contributions to microarchitecture, use of novel microarchitectural techniques or compiler/architecture interfacing. It is hoped, but not required, that the winner will have also contributed to

the computer microarchitecture community through teaching, mentoring, or community service.

**AWARD:** Certificate and a \$2,000 honorarium.

**PRESENTATION:** Annually presented at the ACM/IEEE International Symposium on Microarchitecture

**NOMINATION SUBMISSION:** This award requires 3 endorsements. Nominations are being accepted electronically: [www.computer.org/web/awards/rau](http://www.computer.org/web/awards/rau)

**CONTACT US:** Send any award-related questions to [awards@computer.org](mailto:awards@computer.org)

[www.computer.org/awards](http://www.computer.org/awards)