

Summit: Monetizing the Semantic Web

Denny Vrandečić, Elena Simperl, Rudi Studer – KIT

The technology stack of the Semantic Web is standardizing the protocols of how to publish, format, and link data in a common space. Standardization increases the potential number of customers and uses for a dataspace. Here we discuss a number of approaches on how to monetize the Semantic Web without sacrificing the underlying principals of Linked Open Data. It is not and can not be a complete overview, but a first draft in order to offer some answers to the economic questions raised.

We define the Semantic Web analogously to the Web: it is a set of standards and principles that allow for a shared space of data (or, on the Web, documents). The Semantic Web is built on the following standards: URIs, HTTP, XML (all three already known from the hypertext Web), and RDF (as the shared data model of the Semantic Web). Further beneficial (but not necessary) standards are SPARQL, RDFa, RDFS, OWL, and RIF.

We focus on monetization venues that are original to the Semantic Web, and genuinely need Semantic Web technologies. This means that indirect approaches of monetizing the Semantic Web are not regarded: offering training courses, certificates, business conferences, books, consultancy, technology provision, etc. This does not mean that one may not expect decent profits using these venues, but that they are not sufficiently different from introducing any other kind of technology in order to be mentioned here.

- **Publishing metadata as a collateral and for Search Engine Optimization.** A company selling a product or service wants to make potential customers aware of their offerings. Since search engines are increasingly important to find products, search engine optimization has become a key technology for marketing. Search engines are becoming increasingly powerful in helping the user to exactly find what they are looking for. The Google recipe search for example allows to not only search for a recipe, e.g. for a chili, but also to then refine the recipe by the time needed for preparing the dish, the calories, or certain ingredients. Providing metadata about their products and services will allow search engines – either general search engines like Google, but also more highly specialized search engines – to offer more intelligent interfaces to sift through data, well beyond the capabilities of keyword search, and improve how Search Engines “understand” the content of a Web site.
- **Accountability and transparency.** The demands for more information about the products we buy and consume are getting increasingly louder. As the rise of fair-trade coffee or chocolate has shown, consumers are willing to pay more under certain circumstances. In many cases providing this information is simply not feasible due to the amount of information. But if instead every single product was connected to the Semantic Web using ubiquitous access to the Web via the mobile phone, the consumer could easily gain access to the data, drill down to the information they are

interested in, and perform informed decisions. Given that the information is provided in Semantic Web standards, the mobile phone may even check the compliance against the consumer's requirements automatically.

- **Freemium data.** Providing entities with minimal useful data can establish a service as an authority for an area. Adding additional data, or a certain quality of service for data, can then be added based on a paid subscription to the data source. Maybe the best known such data provider is IMDb, who offer much richer data and APIs through their pro-accounts.
- **Get the dump for free, but further services will be charged.** Providing access through SPARQL is expensive, whereas merely providing RDF dumps is cheaper by far. SPARQL is just one possible service on top of data, and like other data markets a provider could offer services on top of data that even go beyond the rich querying abilities of SPARQL.
- **Data curation and aggregation services.** Collecting and caching data from different sources, integrating it into one common framework, and providing a single, trusted access point to the data provides a very valuable service that will enable a number of applications to properly function and display the rich, integrated data from the Semantic Web. Whereas such apps have their own stream of revenue, part of it would be transferred to the underlying data curation and aggregation services. This also works with enriching data from other sources, e.g. taking UK government data and interlinking it with further data. Legal publishers have provided such services for decades, enriching and organizing legal texts for higher accessibility and reusability.

We have outlined some financial models for publishers and republishers of data. Once the data is available, it enables the creation of services and apps on top of the data, with their own streams of revenue, like ads or selling the apps. And last, but not least, using Semantic Web standards can also reduce the mid- and long-term costs of any data-intensive application, due to the usage of standards, which leads to the avoidance of vendor lock-in, and a bigger choice in tools and expertise.

We have collected a number of possible monetization paths for the Semantic Web. We do not claim that this is a complete list: we actually expect that the Semantic Web will create new business models that we cannot yet even anticipate, just as the Web did before, and continues to do. But we expect this list to provide at least a first answer to the recurring question of how to monetize the Semantic Web, and why businesses should be keen on entering this growing market, and how.