Linked Services and the Future Internet

Carlos Pedrinaci¹ and John Domingue¹

Knowledge Media Institute, The Open University Walton Hall, Milton Keynes, MK7 6AA, UK (c.pedrinaci, j.b.domingue)@open.ac.uk

Abstract. It has often been argued that Web services would have a tremendous impact on the Internet, as a core enabling technology supporting a highly efficient service-based economy at a global scale. However, despite the outstanding progress in the area we are still to witness this impact. In this paper, we analyse the state of the art highlighting the main reasons we believe have hampered their uptake. Based on this analysis, we further discuss about current trends and development within other fields and argue that the recent evolution, notably Linked Data and Web 2.0 technologies, provides the missing ingredients that will lead to a new wave of services – Linked Services – that will drive the creation of complex systems in the Future Internet. Throughout the presentation of this vision we outline the main principles underpinning Linked Services and we illustrate how they can be implemented using a number of technologies and tools we have developed and are in the process of extending.

1 Introduction

The Internet has driven to a large extent the evolution of Information and Communication Technologies in the last years as well as it has been at the core of a number of outstanding socio-economic changes across the globe. Not surprisingly major efforts are devoted to both analysing and extending or improving it. The Future Internet initiative is one such effort focussing on both anticipating and devising the technologies for the forthcoming years. These efforts span from the lowest layers such as hardware and software for improving network communications to the highest layers enabling the creation of a service-based economy over the Internet.

In this position paper we shall focus on the service layer from a technical perspective. We shall highlight the existing limitations of current technologies in the area and we shall propose a set of simple principles and technologies that overcome these drawbacks. The approach described herein is centred on the integration of service technologies and Linked Data as means to both augment the potential and applicability of Linked Data and to remedy existing limitations of Web services enabling their use on a Web scale.

Throughout this paper we argue that essential to the evolution of the Internet towards enabling the creation of highly advanced applications and businesses is the systematic description of reusable software services and related information as Linked Data better enabling their discovery and interpretation. Similarly, we also argue that the emerging Web of Data is currently in need for means and technologies supporting the systematic, maintainable, and efficient development of (complex) applications, and services represent the right abstraction for this. We anticipate that this integration will mitigate the existing limitations of both services and the Web of Data, giving birth to a new wave of highly advanced applications for the Future Internet.

2 The Current State of Affairs

Web Services and the Service-Oriented Architecture are commonly lauded as a silver bullet for Enterprise Application Integration, implementation of interorganizational business processes, and even as a general solution for the development of all complex distributed applications. Despite the appealing characteristics of service-orientation principles and technologies, their uptake on a Web-scale has been significantly less prominent than initially anticipated [5]. First and foremost Web services, despite their name, are hardly a Web-oriented technology [17] but rather one suited for enterprises which so far have been reluctant to publish functionality on the Web. Secondly, from a technical perspective, current technologies are such that software developers need to devote a significant effort to discovering sets of suitable services, interpreting them, developing software that overcomes their inherent data and process mismatches, and finally combining them into a complex composite process.

Semantic Web Services (SWS) [10] have tried to overcome Web services limitations by enriching them with semantic annotations in order to better support their discovery, composition, and execution. Up until now, however, the impact of SWS on the Web has been minimal. In fact, although SWS technologies have already shown their benefits, e.g., in discovery [15], research in the area has failed to take into account the socio-economic aspects devoted to the creation and annotation of services. First, research has mostly focused on devising highly expressive conceptual models glossing over the complexity they introduce and the additional effort demanded of users. Second, SWS research has for the most part targeted WSDL/SOAP based Web services which are not prevalent on the Web [5]. As a consequence, SWS is instead a niche technology only accessible to highly trained experts and the benefits obtained are most often not considered worth the additional investment.

In parallel, the Web is currently witnessing a dramatic change with the advent of Web 2.0 [7]. Web 2.0 technologies are "socialising" the Web, putting individuals at the core of the Web as both data producers and consumers. Web 2.0 technologies have shown that collaboration over the Web can produce outstanding results with a low cost, and it is also encouraging enterprises and institutions to offer their data and services publicly at a previously unprecedented scale and pace [7,5]. Unfortunately, Web APIs are generally described using plain, unstructured HTML. As a consequence, and despite their popularity, discovering Web APIs or developing mashups that integrate disparate services in this manner suffers from a number of limitations similar to those affecting "classical" Web services, with an increased complexity since most often no machine-processable description is available.

Finally, Linked Data technologies [2] have given birth to the Web of Data, "a Web of things in the world, described by data on the Web" [2]. The Web of Data, impelled by the current trend towards an open Web, has recently experimented an outstanding growth and currently provides publicly large amounts of interconnected data concerning a wide range of domains [2]. The outstanding growth of the Web of Data is urging researchers to devise means to exploit the valuable information it exposes. Among the main applications produced so far there are a number of data browsers that help people navigate through the data. There are systems that crawl, index and provide intelligent search support over the Web of data like Sindice [11] and Watson [4]. And finally, there are a few domain-specific applications such as Revyu.com [6]. Although useful these applications hardly go beyond presenting together data gathered from different sources leaving the potential of this massive data space unexploited. It is therefore becoming of crucial importance to devise ways in which smart applications that exploit the Web of Data could be systematically developed.

3 Services and the Web of Data: An Unexploited Symbiosis

We believe that services and the Web of Data are characterised by a number of complementary features that can mitigate each other's limitations and can ultimately enable the development of highly advanced applications and services as envisioned for the Future Internet [12]. In the remainder of this paper we shall refer to this new kind of services as Linked Services.

On the one hand, from a technological perspective, the evolution of the Web of Data is highlighting the fact that light weight semantics yield significant benefits that justify the investment in annotating data and deploying the necessary machinery. This initiative is contributing to generate an outstanding body of knowledge that can help to significantly reduce the effort for creating semantic annotations for services. Furthermore, it also represents a significant use case for the application of services technologies on a Web scale in order to process this wealth of data which remains nowadays largely unexploited. On the other hand, from a socio-economic perspective, the recent evolution around Web 2.0 has shown that collaboration over the Web can lead to large quantities of very useful data with a low cost. Similarly, both Web 2.0 and more recently Linked Data technologies are encouraging enterprises and institutions to offer their data and services publicly enabling open innovation on a Web scale. This new scenario provides in our view suitable technologies and data, as well as the necessary economic and social interest for the wide application of services technologies on a Web scale.

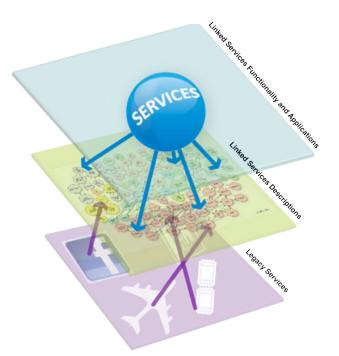


Fig. 1. Linked Services for the Web of Data.

The vision toward the next wave of services – Linked Services – previously presented in [12] and depicted in Figure 1, is based on two simple ideas: publishing service annotations in the Web of Data, and creating services for the Web of Data, i.e., services that process Linked Data and generate Linked Data. In a nutshell, Linked Services are services described as Linked Data. Therefore, these are service descriptions whereby their inputs and outputs, their functionality, and their non-functional properties are described in terms of (reused) light weight RDFS vocabularies and exposed following Linked Data principles. In fact, as such, Linked Services descriptions represent highly valuable information which is still to be provided in the Web of Data: data about reusable functionality on the Web. Secondly, by virtue of these descriptions, Linked Services are therefore services that, with appropriate infrastructure support, can consume RDF from the Web of Data, and, if necessary, can also generate additional RDF to be fed back to the Web of Data. In other words, Linked Services constitute a processing layer on top of the wealth of information currently available in the Web of Data which remains unexploited.

4 Services on the Web of Data

Before any significant uptake of services can take place on the Web, proper mechanisms for creating, publishing and discovering services must be in place. In this respect the state of the art shows that:

- Semantics are essential to reach a sufficient level of automation during the life-cycle of services and applications,
- finding an adequate trade-off between the expressivity of the service model used and the scalability from a computational and knowledge acquisition perspective is key for a wide adoption of service technologies,
- the annotation of services should be simplified as much as possible, and "crowdsourcing" appears to be a particularly effective and cheap solution to this end,
- on the Web, light weight ontologies together with the possibility to provide custom extensions prevail against more complex models,
- any solution to deploying services that aspires to be widely adopted should build upon the various approaches and standards used on the Web, including Web APIs, RDF, and SPARQL,
- Linked Data principles [2] represent nowadays the best practice for publishing data on the Web both for human and machine consumption,
- links between publicly available datasets are essential for the scalability and the value of the data exposed.

The principles we have just highlighted have an impact in a wide range of activities during the life-cycle of services, notably during the annotation, publication, discovery, and use of Web services and Web APIs on the Web. We shall cover these in the remainder of this paper.

4.1 Semantic Description of WSDL Services and Web APIs

The semantics of a service, or software component, are typically divided into four kinds: i) data model; ii) functional semantics; iii) non-functional semantics; iv) behavioural semantics. The level of refinement with which software components can be manipulated (e.g., discovered, composed, invoked), depends directly on the level of detail with which their semantics are captured and exposed.

In our current work we build upon the Minimal Service Model (MSM) [8, 13], WSMO-Lite [18] and hRESTS/MicroWSMO [8], for capturing the semantics of services as Linked Data. In a nutshell, the MSM is a simple RDF(S) integration vocabulary which captures the maximum common denominator between existing conceptual models for services. Thus, the MSM is an integration model at the intersection of existing formalisms, able to capture the core semantics of both Web services and Web APIs in a common model, homogeneously supporting publication and discovery.

The MSM is enriched with WSMO-Lite and hRESTS/MicroWSMO. The former provides a very simple RDFS ontology together with a methodology for expressing functional and nonfunctional semantics, and an information model for WSDL services. The latter is a microformat-like¹ notation that forms the basis for our work on semantically describing Web APIs. It enables the creation of machine-processable Web API descriptions based on available HTML documentation by allowing one to structure APIs descriptions and by including pointers to the semantics of the service, operations, and data manipulated.

4.2 Supporting Services Annotation

Arguably, one of the main limitations of previous approaches to integrating services in the Semantic Web, has been the difficulty from an annotation perspective. The use of light weight ontologies as opposed to highly expressive conceptual models reduces considerably the effort involved and the amount of annotations to be provided. Additionally, and more importantly, the Web of Data is significantly changing the environment from an annotation and usage point of view.

On the one hand, the Web of Data can be used as background knowledge [4] in order to provide suitable ontologies that can be used, extended, and combined to create domain ontologies for annotating services in an easier manner. On the other hand, generating service annotations by reusing existing ontologies directly contributes to increasing services usability and presumably their uptake. For instance annotating services inputs and outputs with respect to existing vocabularies ensures the direct applicability of services over data already available as well as it allows Linked Data application developers to carry out data driven discovery of services by simply checking the input and output types of services. From a more abstract perspective, this process ensures that services modeled in this way are linked to the Web of Data as encouraged by Linked Data principles.

We are currently devoting significant efforts to creating tools that support users in the annotation of services based on the principles introduced above. One such application is SWEET [9] which is, to the best of our knowledge, the first tool that enables the creation of semantic annotations for Web APIs and RESTful services. SWEET provides user support for creating hRESTS/MicroWSMO annotations over any HTML page describing Web APIs, therefore supporting a non-intrusive incremental annotation of existing resources. The tool, assisted by Watson [4], supports users in browsing the Semantic Web while annotating services so that they can identify suitable vocabularies such as FOAF [3], and use them for the annotation. A tool called SOWER, based on the same principles but focussing on the annotation of WSDL services, has also been developed.

4.3 Publication and Discovery of Services on the Web of Data

At present, service descriptions can only be retrieved through the Web of documents, which is essentially designed for human beings, or through specific interfaces to registries such as UDDI that have failed to gain significant uptake. A fundamental step for bringing services closer to the Web is their publication

¹ See http://www.microformats.org

based on current best practices. We view service annotations as a particular kind of highly valuable data: data that informs us about existing reusable functionality exposed somewhere on the Web that processes and/or generates data. As such, services should therefore be published on the Web according to current best practices for publishing data – Linked Data principles – so that applications can easily discover and process their descriptions on the basis of the very same technologies they use for retrieving data.

In order to explore and validate these principles we have developed iServe [13], a public platform that unifies service publication and discovery on the Web through the use of Linked Data. iServe builds upon lessons learnt from research and development on the Web and on service discovery algorithms to provide a generic semantic service registry able to support advanced discovery over different kinds of services described using heterogeneous formalisms. iServe uses as its core conceptual model the MSM and it currently includes a number of import mechanisms able to deal with a large variety of Web service and Web APIs description formats.

Once imported, iServe publishes the semantic annotations of services as Linked Data and provides a number of semantic discovery methods including SPARQL, or more refined algorithms exploiting the semantic descriptions of services, RDFS inferencing, and similarity measures for more accurate results. The current implementation already shows how Web services and Web APIs can be described by means of an homogeneous conceptual model – the Minimal Service Model – and how they can be published as Linked Data, therefore better promoting their discovery based on the use of the well established and adopted Linked Data principles.

5 Services for the Web of Data

The notion of services as well-defined, independent, invokable and distributed pieces of functionality is indeed a very powerful architectural notion for developing distributed systems. Encapsulating functionality in this way provides the capacity for maintaining a loose coupling between integrated components which, when it comes to an environment like the Web, appears as a highly beneficial (if not necessary) feature. Services provide therefore a suitable architectural abstraction for the integration of processing capabilities over the Web of Data in a loosely coupled manner.

5.1 Integrating Legacy Systems

Currently a good part of the Web of Data is generated from existing databases by using tools such as D2R [2]. There is, however, a large body of information exposed through Web APIs by companies which are either not interested in offering all the information publicly on the Web given its commercial value and/or its sensitivity, or because they do not have the technical skills or interest in exposing the information as Linked Data. Similarly, in the Internet of Things, there is a growing number of streams of data provided by sensors through highly heterogeneous formats and interfaces, which exhibits considerable integration and processing limitations [16].

We previously highlighted that Linked Services are such that their inputs and outputs are RDF. As a consequence, they represent a natural means for exposing as Linked Data valuable information previously enclosed within silos, through the annotation of existing Web APIs and WSDL services. Web APIs could in this way be invoked by interpreting their semantic annotations, and RDF information could be obtained on demand. In this way, data from legacy systems, state of the art Web 2.0 sites, or sensors, which do not embrace Linked Data principles could be made available as Linked Data easily.

This approach is currently being explored in the context of a number of use cases from the European project SOA4All [5]. Our current experience, although preliminary at this stage, shows already the applicability and potential of bringing legacy systems to the Web of Data in this manner.

5.2 Processing Linked Data

Integration and fusion of disparate data coming from the Web of Data hardly takes place nowadays and therefore applications do not perform any ulterior processing of this data other than for presenting it to the user [2]. Generating new data based on what has been found or the provisioning of added-value services that exploit this data thus remains a pending issue. For instance, something as simple and useful as a unit transformation service is still to be provided for the Web of Data. There is a clear need for enabling the processing of Linked Data in ways such that application developers could conveniently apply them over data gathered at runtime to carry out computations as simple as unit transformations, more complex as deriving similarities between products or services based on the reviews published by users, or even more advanced like booking a trip.

It seems natural to approach the development of systems that process Linked Data by composing Linked Services. These services should be able to consume RDF data, carry out the concrete activity they are responsible for, and return the result, if any, in RDF as well. The invoking system could then store the result obtained or continue with the activity it is carrying out using these newly obtained RDF triples combined with additional sources of data. In a sense this is quite similar to the notion of service mashups [1] and RDF mashups [14] with the important difference that it supports generic services which are, in this case, RDF-aware and their functionality may range from RDF-specific manipulation functionality up to highly complex processing beyond data fusion.

It is worth noting in this respect the benefit brought by having service descriptions in the Web of Data. When developing applications that process Linked Data, discovering useful services could be driven by the data that needs to be manipulated. The reuse of ontologies and vocabularies would in turn contribute towards increasing the compatibility of services. In this way, Linked Data application developers would have access to an ever growing body of reusable components ready to be combined and exploited. Integrating services with the Web of Data as depicted before would give birth to a services ecosystem on top of Linked Data, whereby people would be able to collaboratively and incrementally construct complex systems by reusing the results of others, gradually supporting the development of increasingly complex applications for the Future Internet.

6 Conclusions and Outlook

Despite the appealing characteristics of service-orientation principles and technologies, their uptake on a Web-scale has been significantly less prominent than initially anticipated. The Web is witnessing a dramatic evolution with the advent of Web 2.0 and Linked Data technologies. Web 2.0 has triggered a socialisation of the Web which has placed individuals at the centre of the Web and is widely based on somewhat altruistic contributions of free data and manual labour from users. The Linked Data initiative is in turn devoted to creating what is referred to as the Web of Data, which already provides publicly large amounts of interconnected data concerning a wide range of domains described in terms of light weight ontologies for supporting automated processing.

We have explored the relationship between services and the Web of Data. In particular we have highlighted that Linked Data represent appropriate principles for publishing services on the Web. We have illustrated how Web services and RESTful services can be brought into the Web of Data by means of simple RDF vocabularies and supporting tools. We have highlighted the fact that the current evolution of the Web of Data is gathering the necessary motivation for the development of advanced applications that process Linked Data. We have outlined that Linked Services are particularly well-suited for supporting developers in creating applications that process Linked Data.

The overall vision outlined herein represents the roadmap for the research we are currently carrying out trying to expand the capabilities of the Linked Data applications as well as trying to promote and support the use of services on the Web through light weight semantic annotations. This research, like the principles it builds upon, will strive to provide data, resources, tools and engines publicly on the Web in order to eventually lead to the wider uptake of services on a Web scale.

7 Acknowledgements

This work is partly funded by the EU SOA4All project. We would like to thank all the members of the SOA4All project for their contribution.

References

 Djamal Benslimane, Schahram Dustdar, and Amit Sheth. Services Mashups: The New Generation of Web Applications. *IEEE Internet Computing*, 12(5):13–15, 2008.

- Christian Bizer, Tom Heath, and Tim Berners-Lee. Linked Data The Story So Far. International Journal on Semantic Web and Information Systems (IJSWIS), 2009.
- 3. Dan Brickley and Libby Miller. FOAF Vocabulary Specification 0.91. http://xmlns.com/foaf/spec/, November 2007.
- Mathieu d'Aquin, Enrico Motta, Marta Sabou, Sofia Angeletou, Laurian Gridinoc, Vanessa Lopez, and Davide Guidi. Toward a New Generation of Semantic Web Applications. *IEEE Intelligent Systems*, 23(3):20–28, 2008.
- John Davies, John Domingue, Carlos Pedrinaci, Dieter Fensel, Rafael Gonzalez-Cabero, Morgan Potter, Marc Richardson, and Sandra Stincic. Towards the Open Service Web. *BT Technology Journal*, 26(2), 2009.
- Tom Heath and Enrico Motta. Revyu: Linking reviews and ratings into the Web of Data. Web Semantics: Science, Services and Agents on the World Wide Web, 6(4):266-273, 2008.
- James Hendler and Jennifer Golbeck. Metcalfe's law, Web 2.0, and the Semantic Web. Web Semantics: Science, Services and Agents on the World Wide Web, Jan 2008.
- Jacek Kopecký, Karthik Gomadam, and Tomas Vitvar. hRESTS: an HTML Microformat for Describing RESTful Web Services. In *The 2008 IEEE/WIC/ACM International Conference on Web Intelligence (WI2008)*, Sydney, Australia, November 2008. IEEE CS Press.
- Maria Maleshkova, Carlos Pedrinaci, and John Domingue. Supporting the Creation of Semantic RESTful Service Descriptions. In Workshop: Service Matchmaking and Resource Retrieval in the Semantic Web (SMR2) at 8th International Semantic Web Conference, 2009.
- Sheila McIlraith, Tran Son, and Honglei Zeng. Semantic web services. Intelligent Systems, IEEE, 16(2):46 – 53, Jan 2001.
- Eyal Oren, Renaud Delbru, Michele Catasta, Richard Cyganiak, Holger Stenzhorn, and Giovanni Tummarello. Sindice.com: a Document-oriented Lookup Index for Open Linked Data. *IJMSO*, 3(1):37–52, 2008.
- Carlos Pedrinaci and John Domingue. Toward the Next Wave of Services: Linked Services for the Web of Data. *Journal of Universal Computer Science*, 16(13):1694– 1719, 2010.
- 13. Carlos Pedrinaci, Dong Liu, Maria Maleshkova, Dave Lambert, Jacek Kopecký, and John Domingue. iServe: a Linked Services Publishing Platform. In Proceedings of Ontology Repositories and Editors for the Semantic Web at 7th ESWC, 2010.
- 14. Danh Le Phuoc, Axel Polleres, Manfred Hauswirth, Giovanni Tummarello, and Christian Morbidoni. Rapid Prototyping of Semantic Mash-ups Through Semantic Web Pipes. In Juan Quemada, Gonzalo León, Yoëlle S. Maarek, and Wolfgang Nejdl, editors, WWW, pages 581–590. ACM, 2009.
- Thomi Pilioura and Aphrodite Tsalgatidou. Unified Publication and Discovery of Semantic Web Services. ACM Trans. Web, 3(3):1–44, 2009.
- Amit Sheth, Cory Henson, and Satya Sahoo. Semantic Sensor Web. Internet Computing, IEEE, 12(4):78–83, 2008.
- Steve Vinoski. Putting the "Web" into Web Services: Interaction Models, Part 2. IEEE Internet Computing, 6(4):90–92, 2002.
- Tomas Vitvar, Jacek Kopecky, Jana Viskova, and Dieter Fensel. WSMO-Lite Annotations for Web Services. In Manfred Hauswirth, Manolis Koubarakis, and Sean Bechhofer, editors, *Proceedings of the 5th European Semantic Web Conference*, LNCS, Berlin, Heidelberg, June 2008. Springer Verlag.