

Composição de aplicações integrando a Web de Serviços e a Web de Dados

Leandro José S. Andrade
Universidade Federal da Bahia - UFBA
Salvador, Bahia - Brazil
leandrojsa@dcc.ufba.br

Cássio Prazeres
Universidade Federal da Bahia - UFBA
Salvador, Bahia - Brazil
prazer@dcc.ufba.br

ABSTRACT

No cenário atual, a Web apresenta um ambiente bastante heterogêneo de recursos e informações, a qual usuários e aplicações interagem na produção e composição de conteúdo. A fim de aprimorar ainda mais esse cenário Berners Lee propôs uma extensão da Web Atual, acrescentando-se elementos semânticos para descrição das informações: a Web Semântica. Nesse sentido, com proposta de Linked Data motivou a comunidade para criação de um projeto para publicação de dados abertos através do Projeto Linking Open Data. Esse conjunto de dados ficaram conhecidos como Web de Dados, que tornaram mais uma fonte de informações dentro do ambiente altamente heterogêneo que compõe a Web. Este trabalho tem o objetivo de desenvolver um modelo para compor aplicações associando dados oriundos da Web de Dados com Serviços Web, por meio de descrições da Web Semântica. Para tanto, objetivava-se construir um arcabouço, o qual permita criar uma abstração de Serviços Web Semânticos nas bases Linked Data.

Keywords

Composição semântica de aplicações, Linked Data, OWL-S, Serviços Web Semânticos

1. INTRODUÇÃO

O advento da Web 2.0, através da utilização de tecnologias mais interativas, possibilitou a colaboração entre usuários na produção de conteúdo apoiados por aplicações, no ambiente Web, tais como redes sociais, *wiki's* e *blogs* [10]. Além disso, a Web 2.0 agrega o desenvolvimento de outras tecnologias, as quais visam aprimorar as formas de intercâmbio de informações entre aplicações Web, tais como *Web Application Programming Interface* (Web API) [5] e Serviços Web [2]. De acordo com Martin et al. [13], uma tendência da evolução da Web está no desenvolvimento de padrões, para que serviços produzam e consumam dados através do mesmo protocolo. A padronização propicia o desenvolvimento de linguagens para descrição de serviços Web, como WSDL (*Web Services-*

Description Language) e WADL (*Web Application Description Language*), que respectivamente apresentam descrição sintática de serviços SOAP e RESTful.

Como desdobramento das pesquisas sobre Web Semântica [3], sugeriram estudos que buscavam associar descrições semânticas aos Serviços Web [14]. Nesse sentido diversas abordagens foram propostas como padrões para descrição semântica de serviços, dentre os quais destacam-se *Semantic Markup for Web Services* (OWL-S) [8] e o *Web Service Modeling Ontology* (WSMO) [19]. Posteriormente, Bizer et al. [6] propuseram um conjunto de boas práticas para publicação de dados na Web, que deu origem ao Linked Data. A massiva adoção de tais princípios na publicação de dados deu origem à chamada Web de Dados [7], que foi bastante impulsionada pelo desenvolvimento de um projeto da comunidade de pesquisadores de Web Semântica para publicação de dados abertos [6]: o Projeto *Linking Open Data*.

Assim, no contexto descrito até então, as aplicações da Web atual, cada vez mais destinam-se a disponibilizar suas funcionalidades para serem acessadas por outras aplicações [1]. Segundo O'Reilly [16], existe uma grande diversidade de fonte de dados em crescente desenvolvimento e uma demanda de aplicações que compõe dados de diferentes origens e formatos a fim de oferecer serviços mais completos. Esse tipo de *software* é também conhecido como *mashup* [1], que, em outras palavras, combinam recursos pré-existent na Web para criar novas aplicações. Todavia, para o desenvolvimento de tais aplicações é necessário um esforço de programação do desenvolvedor para compor os recursos da Web disponíveis [9]. Quando esse esforço é minimizado, consequentemente a aplicação torna-se fortemente dependente de um escopo limitado de recursos, perdendo a capacidade de dinamicamente utilizar novos serviços, que se tornam disponíveis em tempo de execução [9].

Assim, esta proposta de dissertação tem como objetivo prover a composição automática de aplicações envolvendo Serviços Web e Web de Dados. Esse problema é bastante atual e relevante, visto que diversos autores têm proposto trabalhos [1, 9, 4, 23, 15, 22] na direção de resolver os problemas envolvidos na integração de serviços e dados. Dessa forma, este trabalho pretende investigar e propor solução para viabilizar a utilização dos Serviços Web juntamente com a Web de Dados de modo automático, permitindo a composição de novas aplicações utilizando essas diferentes fontes de dados.

Com essa proposta, é possível enriquecer ainda mais o desenvolvimento de *mashups*, gerando aplicações que combinam dados da crescente nuvem LOD com Serviços Web tradicionais através da descrição semântica dos serviços. Em linhas gerais, a proposta prevê elaborar Serviços Web a partir de requisições representadas em OWL-S, de modo que dados semânticos publicados na Web de Dados possam ser utilizados na composição dos serviços de forma automática. Assim, dada uma requisição de um serviço por uma aplicação ou por um usuário, descrita semanticamente, busca-se encontrar algum serviço disponível e compatível com a solicitação. Caso nenhum serviço seja encontrado, aplica-se técnicas para tentar produzir esse serviço a partir de serviços menores ou de dados advindos da nuvem LOD, que conjuntamente atendam a requisição.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Serviços Web Semânticos

Martin et al.[13] expõem que a ausência, ou a pouca descrição, do funcionamento semântico dos Serviços Web, que os tornam compreensíveis por agentes de *software* caracteriza-se como uma das suas principais limitações. Essa primeira geração de serviços web representa bem seus aspectos sintáticos, todavia somente as descrições sintáticas dos serviços não são suficientes para prover uma compreensão semântica, que permita esse tipo de interpretação por agente de software. Para isso, pesquisadores têm desenvolvido linguagens, arquiteturas e abordagens similares, trabalhos que se caracterizam como serviços web semânticos, que comumente é definido como o enriquecimento das descrições dos serviços web através do uso de anotações da Web Semântica. Eles tem como objetivo permitir a automação de tarefas como descoberta, composição e execução de serviços web [17].

A *Ontology Web Language for Services* (OWL-S) [11] é uma linguagem para descrição semântica de serviços, baseada em ontologias descritas em OWL (*Web Ontology Language*), que tem como principais objetivos: (a) compor um *framework* representativo para descrição semântica do serviço; (b) viabilizar a automatização no gerenciamento de serviços, tal como sua utilização por agentes de software; (c) integrar os padrões existentes dos serviços web com a web semântica; (d) por fim, ser abrangente o suficiente para dar suporte a todo ciclo de vida do serviço web [12].

2.2 Linked Data

Devido a esta fragmentação e ausência de semântica da descrição dos documentos, Bizer et al. propuseram boas práticas para publicação e conexão de dados estruturados na Web denominada Linked Data [6]. Os princípios Linked Data visam a padronização da identificação, da recuperação e da representação dos dados, para permitir a descoberta novas fontes de dados em tempo de execução [6].

As informações publicadas com os princípios Linked Data ficaram conhecidas como Web de Dados [6], onde URIs não mais identificam documentos, elas apontam para recursos, que representam entidades do mundo real, como pessoas, países e categorias de objetos, usando a linguagem Resource Description Framework (RDF): uma linguagem abstrata e genérica para metadados na Web. O modelo RDF é descrito em triplas no formato sujeito, predicado e objeto. Para buscas semânticas na Web de Dados utiliza-se *Simple Protocol*

and RDF Query Language (SPARQL), que é a linguagem de consultas padrão para bases RDF [18].

2.3 Trabalhos Relacionados

Lorenzo et al.[9] ressaltam em seu trabalho que a composição de *mashups* possui diferenças em relação a composição de Serviços Web, que foca principalmente na integração de negócios (processos) dos serviços, além de agregarem dados com estruturas bastante heterogêneas. Os autores também argumentam que as *mashups* abrem uma grande oportunidade para consumo de dados e serviços, contudo ainda precisa de muito investimento em pesquisas e desenvolvimento, pois além da necessidade de conhecimento em programação é fundamental ter habilidade para realizar composições entre serviços. De modo semelhante, Bianchini et al. [4] argumentam também a diferença entre composição de aplicações Web e composição de Serviços Web. Nesse estudo também citam o grande volume de *mashups* que tem surgido, juntamente com ferramentas de autoria, que visam dar suporte facilitador aos usuários, todavia o processo de combinação e integração é fortemente manual. Portanto Bianchini et al. propõem métodos e técnicas para anotações semânticas dos componentes de *mashup*, visando abstrair a representação heterogênea subjacente. Segundo Chowdhury et al. [20] as ferramentas para composição de *mashups*, apesar de disponibilizarem um ambiente gráfico e um paradigma mais simples, ainda exigem do usuário conhecimento especializado no desenvolvimento de aplicações. Analisando especialmente a ferramenta de composição de *mashups*, Yahoo! Pipes¹, os autores observam, que a realização de composição de componentes, além de ser um processo manual, não é trivial e nem intuitivo.

Seguindo mais uma linha de pesquisa na composição de Serviços Web na Web de Dados, Norton e Stadtmüller [15, 22] ressaltam abordagens que buscam associar Serviços Web aos princípios Linked Data, que buscam expô-los utilizando recursos como HTTP, RDF e SPARQL. Dessa forma, a proposta visa criar uma descrição semântica em serviços RESTful, de modo que seja possível aplicar operações de composição sobre eles a partir de tais interfaces, abstraindo suas particularidades de implementação. Com isso, as entradas e saídas são descritas em RDF utilizando ontologias populares na Web. Segundo Speiser e Harth [21] a forte iniciativa de publicação de informações seguindo os princípios Linked Data dentro da nuvem de dados ligado no projeto Linking Open Data. Apesar do grande potencial dos dados integrados da nuvem LOD, existem diversas outras fontes de informações, que não podem ter seus dados totalmente materializados abertamente na Web, pois em alguns casos estão em constantes mudanças, ou seus valores dependem de entradas específicas, ou por interesses comerciais.

3. OBJETIVO E CONTRIBUIÇÕES ESPERADAS

O objetivo deste trabalho é prover composição automática de aplicações envolvendo Serviços Web e a Web de Dados. Isto é, construir um modelo para permitir que aplicações satisfaçam suas requisições através da composição automática de dados provenientes de Serviços Web e da nuvem Linked

¹Yahoo! Pipes - <http://pipes.yahoo.com>

Data. Para tanto, pretende-se utilizar as descrições semânticas nas requisições de serviços com a linguagem OWL-S. Assim, partindo das solicitações descritas com OWL-S busca-se nas bases Linked Data (através de um arcabouço desenvolvido) dados que satisfaçam tal solicitação. Em conjunto com a extração de serviços na LOD observa-se Serviços Web semanticamente descritos, que ao serem analisados de modo integrado atendam a requisição de serviço.

Como desdobramento do objetivo principal deste trabalho, identifica-se a necessidade de descobrir serviços intermediários (subserviços), que atendam parcialmente as solicitações de Serviços Web. Isto porque como a proposta almeja compor aplicações a partir de serviços, é necessário identificar os que atendem parcialmente para posteriormente aplicar composição destes subserviços. Assim como objetivos específicos este trabalho deve prover e investigar técnicas para descoberta de (sub)serviços na Web de Serviços e na Web de Dados. Além destes destaca-se a atividade de compor os sub-serviços encontrados a fim de fornecer novas aplicações. É importante ressaltar que o presente trabalho, não visa sobrepor os estudos já realizados sobre composição e descoberta de Serviços Web. Sua contribuição concentra-se na descoberta de Serviços Web extraíndo informações semânticas na Web de Dados. Já para descoberta de Serviços Web na Web de Serviços e a composição dos subserviços será utilizado os estudos e algoritmos estabelecidos mais principais pesquisas do tema.

Assim, será desenvolvido um arcabouço junto ao OWL-S, que remete uma série de contribuições a fim de de construir uma solução para compor serviços web juntamente com Web Dados a partir de requisições semanticamente descritas em OWL-S. Dentro dos objetivos deste trabalho espera-se promover as seguintes contribuições:

- **Expressão de requisições OWL-S em consultas SPARQL.** Desenvolver um componente de software que dada uma pedido de serviço descrito em OWL-S, expresse, se possível, em uma consulta SPARQL. Esse ponto é fundamental para a construção desta proposta, pois é a partir da consulta SPARQL que a nuvem Linked Data será explorada na busca de dados compatíveis com a solicitação.
- **Estender a OWL-S para prover um *grounding* para Web de Dados.** O OWL-S na sua definição propõe que a sub-ontologia de *grounding* possa ser estendida se adaptando para as tecnologias utilizadas. Por exemplo OWL-S possui um *grounding* definido para serviços SOAP e outro para serviços RESTful, que respectivamente são o WSDL e WADL. Assim nossa proposta é desenvolver um *grounding* que aporte abstração de serviços web a partir extração de dados Linked Data e por consequência permitir a composição com serviços SOAP e RESTful.
- **Estender a OWL-S API² para executar o *grounding* para Web de Dados.** Essa biblioteca é a principal ferramenta na implementação de serviços OWL-S. Então, o intuito é produzir uma extensão da API elaborando um código que permita a biblioteca executar conforme o *grounding* planejado.
- **Realização de um estudo de caso utilizando bases de serviços web juntamente com bases Linked Data.** Com objetivo de validar e avaliar o desempenho da proposta, desenvolveremos um estudo de caso que englobe todas demais contribuições deste trabalho.

²OWL-S API - <http://on.cs.unibas.ch/owl-s-api>

4. METODOLOGIA

O processo de desenvolvimento deste trabalho pode ser organizado segundo as seguintes atividades: a) Revisão da literatura; b) Aplicação para gerar consulta SPARQL a partir do OWL-S; c) Desenvolvimento de um modelo de *grounding* da OWL-S para Web de Dados; d) Estender a API OWL-S para suportar o *grounding* projetado; e) Realização de estudo de caso; f) Desenvolvimento da dissertação.

A figura 1 mostra o ciclo de vida das principais atividades do projeto. É possível notar que em grande parte das tarefas seguem-se um encadeamento sequencial, pois a conclusão de uma atividade torna-se ponto de partida para outra. O ciclo formado com a seta pontilhada denota eventuais modificações no modelo a partir de observações durante a realização do estudo de caso, momento do qual as formalizações serão avaliadas. Por fim, a atividade de Revisão da literatura, não possui nenhum relacionamento explícito com as demais, pois estará presente em todo ciclo de vida do projeto. Entre a conclusão de cada etapa do projeto está previsto a submissão de artigos científicos em eventos, a fim de reforçar a relevância do trabalho dentro da comunidade acadêmica de Serviços Web Semânticos. Além disso, possibilitará importantes *feedbacks* sobre o andamento do projeto, que auxiliará no desenvolvimento do mesmo. O atual estágio do trabalho encontra-se nos itens b) e c).

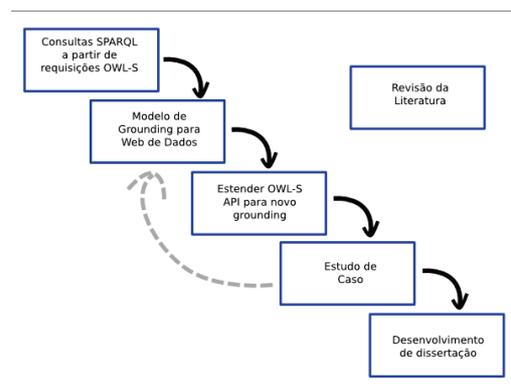


Figure 1: Ciclo de vida das atividades do projeto

5. ESTADO ATUAL DO TRABALHO

Neste trabalho, o estudo iniciou-se buscando analisar a viabilidade da proposta. Desse modo, buscou-se compreender quais os benefícios de prover composição automática de aplicações envolvendo Serviços Web e a Web de Dados através descrições semânticas em OWL-S. Primeiramente, observa-se que integrar a nuvem Linked Data com os serviços OWL-S, permite explorar a capacidade semântica desses dados, o que oportuniza a criação de aplicações envolvendo dados de diferentes fontes. Contudo, parte desta motivação já é suprida pela utilização de consultas em SPARQL, que possibilita explorar os dados semanticamente, porém não comporta a composição, sobretudo automática, com serviços de outras fontes. A construção dessa ideia propicia a criação de serviços moldáveis de acordo com a solicitações. Isso significa a criação de serviços sob demanda, a qual depende somente da disponibilidade dos dados solicitados estarem semanticamente descritos na nuvem Linked Data.

Com esse entendimento, segue-se para a análise da viabilidade técnica da proposta. O primeiro passo foi planejar um modelo geral para a descoberta de serviços com descrições em OWL-S nas bases Linked Data. Assim, adotou-se um modelo baseado na criação de uma consulta SPARQL a partir da análise das entradas, saídas, pré-condições e efeitos da descrição OWL-S, o qual a consulta criada deve expressar um sentido semelhante à requisição do serviço. Com a consulta em mãos verifica-se nas bases Linked Data se existe algum conjunto de dados capaz de satisfazer a solicitação.

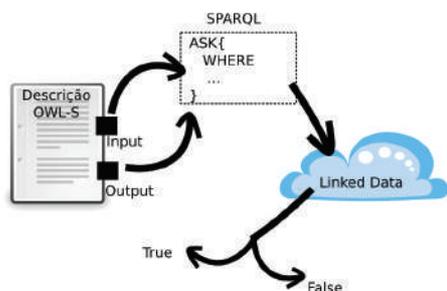


Figure 2: Proposta para utilização da Web de Dados para descobrir serviços com descrição OWL-S

A figura 2 apresenta uma visão geral da estrutura da solução proposta. Note que, a partir da descrição do serviço em OWL-S é desenvolvida uma consulta SPARQL do tipo ASK (retorna *true* ou *false* se existir um ou mais dados de acordo com a requisição) a partir das ontologias das entradas e saídas. Após isso, o próximo passo é a aplicação da consulta em um SPARQL *endpoint* para verificar se existem dados que correspondem a solicitação. Em caso de positivo, pode-se compor um serviço e utilizar essa base de dados para as respostas. Apesar da estrutura inicial, existem alguns desafios associados a essa transcrição. Primeiramente, a expressividade semântica da linguagem OWL-S é superior à capacidade da SPARQL. Isto acarreta a necessidade de identificar quais tipos de requisições estão habilitadas para realização da expressão de consulta. Assim, para uma avaliação experimental, foi selecionado um perfil de requisição OWL-S simples, que permita a geração de consultas. Dessa forma, foram selecionadas as solicitações de serviços onde o elemento de entrada é pertencente (é uma propriedade) do elemento de saída, por exemplo a entrada sendo uma música e a saída o compositor; ou entrada o título de um filme e a saída a instância do filme em si.

6. CONSIDERAÇÕES FINAIS

As transformações vividas pela Web no contexto atual, apresentam um cenário com dados advindos de diversas fontes e aplicações que visam prover suas funcionalidades mesclando informações de origem distintas [10]. Um dos fortes caminhos de pesquisa, visa associar a Web Semântica, a fim de inserir na descrição dos serviços abstrações que permitam automatizar o processo de descoberta e composição de aplicações. Portanto este trabalho almeja prover a composição automática de aplicações envolvendo Serviços Web e a Web de Dados, através de descrições semânticas com a linguagem OWL-S. Esta proposta espera trazer importantes contribuições para área de Serviços Web Semânticos e nos estudos de composição de aplicações associando Serviços Web a Web de

Dados.

7. REFERENCES

- [1] D. Benslimane, S. Dustdar, and A. Sheth. Services mashups: The new generation of web applications. *Internet Computing, IEEE*, 12(5):13–15, 2008.
- [2] T. Berners-Lee. Services and semantics: Web architecture. W3C recommendation, World Wide Web Consortium, 2001.
- [3] T. Berners-Lee, J. H., and O. L. The semantic web. *Scientific American*, 284(5):34–43, May 2001.
- [4] D. Bianchini, V. De Antonellis, and M. Melchiori. Semantic-driven mashup design. In *Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services*, iiWAS '10. ACM, 2010.
- [5] C. Bizer. The emerging web of linked data. *IEEE Intelligent Systems*, 24(5):87–92, Sept. 2009.
- [6] C. Bizer, T. Heath, and T. Berners-Lee. Linked Data - The Story So Far. *IJSWIS*, 5(3):1–22, Mar/Mar 2009.
- [7] C. Bizer, T. Heath, K. Idehen, and T. Berners-Lee. Linked data on the web. In *Proceedings international conference on WWW, WWW '08*, pages 1265–1266, New York, NY, USA, 2008. ACM.
- [8] M. Burstein, J. Hobbs, O. Lassila, D. McDermott, S. McIlraith, S. Narayanan, M. Paolucci, B. Parsia, T. Payne, E. Sirin, N. Srinivasan, and K. Sycara. Owl-s: Semantic markup for web services. Website, November 2004.
- [9] G. Di Lorenzo, H. Hacid, H.-y. Paik, and B. Benatallah. Data integration in mashups. *SIGMOD Rec.*, 38(1):59–66, jun 2009.
- [10] M. Lytras, E. Damiani, and P. O. de Pablos, editors. *Web 2.0 : the business model*. Springer, New York, 2009.
- [11] D. Martin, M. Burstein, J. Hobbs, O. Lassila, D. McDermott, S. McIlraith, S. Narayanan, M. Paolucci, B. Parsia, T. Payne, E. Sirin, N. Srinivasan, and K. Sycara. Owl-s: Semantic markup for web services. Internet [http://www.w3.org/Submission/2004/SUBM-OWL-S-20041122/], 2004.
- [12] D. Martin, M. Burstein, D. McDermott, S. McIlraith, M. Paolucci, K. Sycara, D. McGuinness, E. Sirin, and N. Srinivasan. Bringing semantics to web services with owl-s. *World Wide Web*, 10(3):243–277, 2007.
- [13] D. Martin, M. Burstein, D. McDermott, S. McIlraith, M. Paolucci, K. Sycara, D. L. McGuinness, E. Sirin, and N. Srinivasan. Bringing semantics to web services with owl-s. *World Wide Web*, 10(3):243–277, Sept. 2007.
- [14] S. McIlraith, T. C. Son, and H. Zeng. Semantic web services. *Intelligent Systems, IEEE*, 16(2):46–53, 2001.
- [15] B. Norton and S. Stadtmuller. Scalable discovery of linked services. In *4th Intl. Workshop on REsource Discovery (RED 2011)*, Vol. 737. CEUR-WS.org., 2011.
- [16] T. O'Reilly. What is web 2.0. design patterns and business models for the next generation of software, September 2005.
- [17] T. Payne and O. Lassila. Guest editors' introduction: Semantic web services. *IEEE Intelligent Systems*, 19(4):14–15, jul 2004.
- [18] E. Prud'hommeaux and A. Seaborne. Sparql query language for rdf. Technical report, W3C, Jan. 2008.
- [19] D. Roman, U. Keller, H. Lausen, J. de Bruijn, R. Lara, M. Stollberg, A. Polleres, C. Feier, C. Bussler, and D. Fensel. Web service modeling ontology. *Appl. Ontol.*, 1(1):77–106, jan 2005.
- [20] S. Roy Chowdhury, C. Rodríguez, F. Daniel, and F. Casati. Baya: assisted mashup development as a service. In *Proceedings of the 21st international conference companion on World Wide Web, WWW '12 Companion*, pages 409–412, New York, NY, USA, 2012. ACM.
- [21] S. Speiser and A. Harth. Integrating linked data and services with linked data services. In *Proceedings of the 8th extended semantic web conference on The semantic web: research and applications - Volume Part I, ESWC'11*, pages 170–184, Berlin, Heidelberg, 2011. Springer-Verlag.
- [22] S. Stadtmuller. Composition of linked data-based restful services. In *Proceedings of the 11th international conference on The Semantic Web, ISWC'12*, pages 461–464, Berlin, Heidelberg, 2012. Springer-Verlag.
- [23] C. Wu, T. Dillon, and E. Chang. Intelligent matching for public internet web services towards semi-automatic internet services mashup. In *Web Services, 2009. ICWS 2009. IEEE International Conference on*, pages 759–766, 2009.