Semantic Web
Services Challenge
Results from the First Year
As computing becomes ubiquitous and pervasive, computing is increasingly becoming an extension of human, modifying or enhancing human experience. Today's car reacts to human perception of danger with a series of computers participating in how to handle the vehicle for human command and environmental conditions. Proliferating sensors help with observations, decision making as well as sensory modifications. The emergent semantic web will lead to machine understanding of data and help exploit heterogeneous, multi-source digital media. Emerging applications in situation monitoring and entertainment applications are resulting in development of experiential environments.

SEMANTIC WEB AND BEYOND Computing for Human Experience addresses the following goals:

- brings together forward looking research and technology that will shape our world more intimately than ever before as computing becomes an extension of human experience;
- covers all aspects of computing that is very closely tied to human perception, understanding and experience;
- brings together computing that deal with semantics, perception and experience;
- serves as the platform for exchange of both practical technologies and far reaching research.

AdditionalTitles in the Series:


Additional information about this series can be obtained from
http://www.springer.com  ISSN: 1559-7474
Semantic Web Services Challenge

Results from the First Year

Edited by

Charles Petrie
Stanford University
Stanford, CA, USA

Tiziana Margaria
University of Potsdam
Potsdam, Germany

Holger Lausen
Michal Zaremba
University of Innsbruck
Innsbruck, Austria

Springer
Foreword by James A. Hendler

Back in 2001, I published a paper called "Agents and the Semantic Web," [1] which outlined a vision of how the then-new Semantic Web technologies being explored at DARPA could be used as a mechanism for connecting descriptions of software to the dynamic content engines of the Web - essentially, outlining the idea that what we now call Web services could be tied to semantics and ontologies. The paper has become one of my more highly cited ones, but really it was more of a vision paper than a technological prescription. Luckily for those people interested, Sheila McIlraith and her colleagues published a paper in that same issue [2] (which has been even more highly cited) which outlined a technical approach for Semantic Web Services, the name by which this area has come to be known.

In the years since, Semantic Web Services have become an increasingly important part of Semantic Web research. This work has essentially co-evolved with the growing importance of service-oriented architectures (SOAs), to the industrial computing sector, providing a set of interesting, and realistic, challenges to researchers. The area has thrived on a combination of research funding for universities, especially from the EU’s framework 6 and framework 7 programs, and industrial support within corporate laboratories. In fact, since the coining of the term in 2001, a Google Scholar search on ”semantic web services” now finds over 8,000 publications that discuss some aspect of this integration of semantics and Web services, a staggering amount of research in such a short amount of time.

However, this growth in interest in the area, and the wide swath of research it engendered, also led to significant confusion over, well, what was it really all good for? Could the addition of Semantics really increase the capabilities of Web Services? Could the theoretical results from the research labs be transitioned into workable and scalable systems? Could the techniques of the researchers be made to work with the commercial languages - SOAP, WSDL, BPEL, etc. - being used in the real world? Finding papers that showed clean results on toy problems was easy, but finding practicable technologies that could be used in real world applications required more effort.

In short, we had an odd situation. With Service-Oriented Computing becoming more and more important as a means of software engineering for distributed systems,
the Semantic Web community was finding it hard and harder to explain what it was able to do or to compare the many competing approaches that were being developed. The many different approaches could not easily be reconciled on purely theoretical grounds, rather, an empirical means of evaluating their capabilities with standard testbeds was clearly needed.

Starting in the middle of this decade, a set of workshops was held to explore this issue. Researchers who felt their work was reaching a capability level that could lead to transition came together to explore how they could develop testbeds that could compare and contrast the various approaches to mediation, composition and choreography, and discover of Web Services that were being proposed. The idea of a Semantic Web Services Challenge was born, and in 2006 the first workshop to include the challenge was held. In the following year, a series of these meetings were held around the world, and a number of systems were tested. This book is the result - the first collection to pull together these results and to allow readers to evaluate the results. Whether they are professionals, interested in using this information to help shape investments in technology, or students, looking for up to date information on the application of semantics to Web Services, readers will find this book to contain a wealth of information.

Since the early days of the Semantic Web, I have often been asked to give talks about emerging trends and capabilities. I must confess that in the past two years, I have avoided discussion of Semantic Web Services due to the very confusion I mentioned earlier. However, the question I’ve been asked most often is ”what is the status of Semantic Web Services?” and until now, I didn’t really have very good answers. Thus, I am indebted to the editor of this book for all the work he has done in helping to create and run the challenge, and now in making sure that the results are documented in this book. It’ll be nice to put this topic back in my talks!

Professor James A. Hendler
Rensselaer Polytechnic Institute, NY, USA
August 11, 2008

References

Foreword by Michael L. Brodie

This book documents lessons learned in the first year of the Semantic Web Services Challenge - the first significant step towards the creation of a set of benchmarks and processes by which to define and measure the performance and correctness of semantic web services at web-scale. The Semantic Web Services Challenge is at an early stage in the development of semantic web services benchmarks for the Web of services, similar to that of the Database Derby in mid-1980s for relational databases. As with relational databases and the Web of documents, the Web of services will have an enormous impact on our increasingly digital lives that in turn will become increasingly dependent on the underlying technologies including semantic web services. Unlike relational databases, the Web of services posses, as does each Next Generation of Computing, qualitatively greater challenges in defining and achieving performance and correctness. The lessons told in this book are not only guides down that path, they will be the tales told of the origins of the Web of services, just as today we tell tales of Ted Codd and origins of relational databases and of Jim Gray and the origins of relational benchmarks.

Let me try to put this book in an historical perspective and tell you why this topic is important and challenging.

In our increasingly digital world the Web has become an integral part of our professional and personal lives. “Just Google it!” can be heard from Anchorage to Zambia, from executives in boardrooms to 13 year olds in grade schools, to grannies in kitchens. We are constantly amazed at the Web, the largest man-made artifact ever created, with over 30 billion Web pages - 5 pages for every man, woman, and child on the planet. More information is added to the Web yearly than has been created in the preceding 5,000 years and will grow by a factor of six from 2007 to 2010 with more than 11 billion searches each month, almost double the world’s population. While the size and the growth of the Web constantly amaze us, what may be more striking is the impact of the Web on our world.

Given the significance and size of the Web, how readily can we find what we need? A search for my name produces 227,000 pages. A search for my company produces 98 million pages. While the first entries are often what I want, how would an automated process select the correct one for a specific purpose? This imprecision
prohibits current web search technology for automated business interactions. The Semantic Web vision is to augment web resources with meta-data to improve web search and facilitate automated interactions between web resources.

Service-orientation is emerging as the paradigm of the Next-Generation of computing. Happily, it will take some time - time to figure out what we are doing. With its origins in the 1970s in abstract data types, objects, and containers, the notion of a service - an interoperable, composable, reusable, and remotely invokable function - will evolve, mature, and persist. So services are not Johnnie-come-latelys. The typically overblown estimates of new technology adoption and growth are far from reality for services in 2007, a year that marked the first significant adoption of services since their introduction in 2000. In 2008, 1,000 services is considered large even for very large enterprises. Converting the major systems of a large enterprise might result in 1 million services of which less than 50% would be published externally resulting in a Web of services two orders of magnitude larger than the Web of documents.

As web and service technologies evolve, the web will move from a Web of documents to a Web of functional and data services in which a web page may contain 5 to 10 services. Services will move the web from the surface web, information available to current web browsers, to the deep web, information and services in databases and systems that underlie current web pages. The deep web is estimated to be 500 times larger than the surface web.

The move from the surface Web of documents to a deep Web of services not only increases the search space by a factor of 500, it also leads to a qualitatively new form of computing. The scale pushes the web beyond the size and complexity that can be dealt with by humans. The Web of services requires a services automation solution in which services interact without human intervention. The vision is for services to achieve a goal by discovering services that meet a requirement and negotiate the use of that service or even adaptation of the service or composition with other services that collectively meet the requirement. The Semantic Web Services vision is to enhance web services with meta-data to enable automated service discovery, selection, negotiation, mediation, adaptation, composition, invocation, and monitoring.

As we enter the Web of services we face two great challenges - scale and automation. While these challenges are familiar, the scale and complexity of this new computing environment make these problems qualitatively different from past computing environments. Scale poses challenges of performance while automation poses challenges of correctness.

In the past, performance issues were primarily addressed with hardware and systems engineering solutions. While these are still fruitful, we are now looking to software for solutions. Augmenting web services with semantics is a software solution emerging from the Semantic Technologies community, a community with little experience in engineering solutions especially at web-scale. Indeed few communities have experience with web-scale computing. Hence, the semantic web services space is novel in many ways and may require more sophisticated measures of engineering as well as of performance and correctness.

It is easy to envisage services interacting dynamically to discover other services with which to negotiate, adapt, and compose, and then to invoke to achieve a re-
quirement. It is quite another matter to specify correctness in this context, let alone achieve it in implementations.

Almost three decades ago, the Next Generation of Computing, at the time, faced similar challenges. In the early 1980’s the projected scale of relational databases was unimaginable, and like the Web of documents far exceeded its projections. As with our current Web of services we are facing unimaginable scale and complexity with novel, unproven technology and with few benchmarks. Now, as then, we require efficient, scalable solutions to problems for which we lack definitions of correctness. But this time we do not control the architecture, which is both distributed and emergent. We require objective means of testing whether the new technology solutions meets realistic performance and correctness requirements. We want to encourage innovation via competition amongst possible solutions, and the development of an objective basis of appropriate measures against which to compare them, and the standardization of accepted solutions. In addition to providing benchmarks for emerging technologies, we want to provide a focal point for such engineering challenges, discussions, and achievements where well-defined industrial problems can be used to drive and test technology solutions. Ideally semantic web services benchmarks will contribute to the development and acceptance of semantic technologies just as relational benchmarks did for relational technology.

In the early 1980’s several database benchmark activities emerged as candidates against which to measure the performance and correctness of emerging relational database technologies. One of the earliest candidates, The Database Derby, run by an emerging database magazine, was run as a series of workshops at database conferences. The importance and need for an objective database benchmark lead to the definition of the “DebitCredit” benchmark [1] by Jim Gray and 24 academic and industrial co-authors. By 1988 eight companies formed the Transaction Processing Performance Council (TPC) [2]. Since then, TPC has defined benchmarks with which to measure the performance and correctness of DBMSs and methodologies by which they are conducted, fairly and objectively. TPC benchmarks have been used to define requirements for emerging workloads such as for e-commerce, decision support, application servers, and web services.

Let me conclude with a projection and a challenge to consider as you read this book. Each Next Generation of Computing needs benchmarks - for the web services era as for the database era. Next Generation challenges will always be at a greater scale and complexity than those of the previous generation. However, the Web of services poses qualitatively greater unknowns and opportunities than did the database era. The relational model of 1970 is largely in tact today. Relational database benchmarks evolved with dramatically new workloads prompting radically new hardware, engineering, and systems technologies, all within the bounds of the relational data model.

There is no such constraining model for web services. While the lack of a single model permits opportunities for other computational and informational models, it also opens the space for engineering and technology solutions. Are there multiple computational models for web services that require multiple benchmarks? Is the Se-
mantic Web Services Challenge the entrance to a wide but single path to the future or to a myriad of paths to parallel but distinct computational futures?

Michael L. Brodie
Chief Scientist Verizon
Cambridge, MA, USA
August 19, 2008.

References

Preface

This “first year” book addresses results of the the SWS Challenge through the November 2007 workshop at Stanford University (USA). The first workshop was held at Stanford in March of 2006, setting up the organization and the drafting the methodology. The subsequent workshops were functioning evaluations: at Budva (Montenegro), Athens (USA), Innsbruck (Austria), and again at Stanford.

This series of workshops has provided a forum for discussion based on a common application. The Challenge focuses on the use of semantic annotations: participants are provided with semantics in the form of natural language text that they can formalize and use in their technologies. Being a challenge rather than a contest, workshop participants mutually evaluate and learn from each others’ approaches.

In this book, the focus is on the understanding of the technical issues in the proposed solutions, and of their tradeoffs. Therefore, solution chapters that describe in depth the technologies of the participant teams are complemented by other chapters containing pairwise comparisons of solutions. A full list of the workshops in the ongoing initiative is available at the SWS Challenge wiki\(^1\).

There have been further results, and teams, in 2007 and 2008 and we invite new teams to participate in any of the coming workshops announced on the Challenge wiki as well to join the W3C SWS Challenge Testbed Incubator during 2008.

We are grateful for the continual and substantial support from Professor Dieter Fensel of the Semantic Technologies Institute Innsbruck and Professor Michael Genesereth of the Stanford Logic Group. This book reflects the major efforts of all the technology contributors and the STII and Potsdam staff, including Omair Shafiq and Christian Winkler, who not only supported the SWS Challenge directly, but who did the final composition of this book in LaTeX.

August 2008,

Charles Petrie
Tiziana Margaria
Michal Zaremba
Holger Lausen

\(^1\) http://sws−challenge.org/
1 Introduction to the First Year of the Semantic Web Services Challenge

Charles Petrie

1.1 SWS Challenge Mission and Organization
1.2 Scope of the Challenge and this Book
1.3 Related Initiatives
1.4 SWS Challenge Evaluation Methodology
1.5 SWS Challenge Problem Scenarios
1.6 Overview of Technologies
1.7 Organization of the Book
1.8 The Challenge is Open

References

2 SWS Challenge Scenarios

Holger Lausen, Ulrich Küster, Charles Petrie, Michal Zaremba, Srdjan Komazec

2.1 Introduction
2.2 The Mediation Scenarios
2.3 The Discovery Scenarios
2.4 Summary

References

Part I Mediation Individual Solutions

3 Mediation using WSMO, WSML and WSMX

Tomas Vitvar, Maciej Zaremba, Matthew Moran, Adrian Mocan

3.1 Introduction
3.2 Execution Model