WONDERWEB DELIVERABLE D25: EVALUATION AND MARKET REPORT

A report on the evaluation of WonderWeb technologies and an assessment of the potential market, including guidance as to directions for further development in response to evolving industrial requirements.

TopQuadrant

Mills Davis, Dean Allemang, and Robert Coyne
Top Quadrant, Inc.
141 Howard Drive
Beaver Falls, PA 15010

mdavis@topquadrant.com
dallemang@topquadrant.com
robert@topquadrant.com

Identifier: D25
Class: Deliverable
Version: 1.0
Date: 3-12-2004
Status: Final
Distribution: Restricted
Lead Partner: VUM

IST Project 2001-33052 WonderWeb:
Ontology Infrastructure for the Semantic Web

This document forms part of a research project funded by the IST Programme of the Commission of the European Communities as project number IST-2001-33052.

For further information about WonderWeb, please contact the project co-ordinator:

Ian Horrocks
The Victoria University of Manchester
Department of Computer Science
Kilburn Building
Oxford Road
Manchester M13 9PL
Tel: +44 161 275 6154
Fax: +44 161 275 6236
Email: wonderweb-info@lists.man.ac.uk

TABLE OF CONTENTS

SECTION ONE

EXECUTIVE SUMMARY

The European Union (EU) has funded Information Society Technologies (IST) research as part of a strategy to accelerate emergence of the European knowledge economy. IST goals are to promote economic growth, foster social equity and cohesiveness, and enhance Europe's global competitiveness.

Early-on, IST identified semantic technologies as key enablers of the knowledge economy, and has committed significant funding towards their development through as series of projects undertaken as part of the FP5 and FP6 program frameworks. The WonderWeb project was conceived, funded, and carried out as part of this wave.

The WonderWeb project is a collaborative partnership between:

- University of Manchester , UK (coordinator): automated reasoning, ontology languages and applications.
- Vrije Universiteit Amsterdam , Netherlands: web applications and ontologies.
- ISTC-CNR , Trento & Rome, Italy: ontologies.
- University of Karlsruhe, Germany: web tools and infrastructure.

WonderWeb's research focus has been *ontology infrastructure for the Web*. As the WonderWeb project entered its final stage, project organizers asked US-based semantic technology consultancy, TopQuadrant, Inc., to prepare an independent report that evaluates WonderWeb technologies, assesses the potential market, and provides guidance for further development in response to evolving industry requirements.

The key findings and conclusions of this report are as follows:

1 WonderWeb project is an integral part of a vision for networked intelligence that has blossomed into a major cycle of innovation we call the "semantic wave". It encompasses a broadening array of emerging technologies, that include: the semantic web, semantic web services, semantic grid, semantic peer-to-peer (P2P), cognitive systems, agent systems, sense-making, simulation, optimization and decision support systems, adaptive systems, autonomic and autonomous systems. While R&D themes in the semantic wave are varied, they are interrelated with a common semantic core.

2 This report breaks new ground by estimating the flows of capital for semantic technologies from government-funded R&D, venture capital investment, industry-funded technology commercialization, and market adoption for the period from 2000 to 2010. Also, this report develops regional estimates for Europe, North America, and Asia as well as estimates for both horizontal and vertical markets for semantic technologies.

3 The semantic wave is building steadily and is on a track that will take it from vision to markets valued in the tens of billions of euros by the end of this decade. The initiating force from the outset of this cycle has been public sector funding of fundamental R&D. In this regard, the EU has been a leader. However, the driving force for mass market adoption will be gains in efficiency, effectiveness and other

measures of performance. This report cites research showing that overall gains in performance can be in the range of 2-10 times, and estimates that the market for semantic technologies will grow to between $40-60 billion by 2010.

4 The principal finding of this report is that WonderWeb has played a seminal and influential role in this story. Ontologies comprise key structural technology for ambient intelligence. Because WonderWeb development has addressed foundation issues such as ontology language, application program interfaces (APIs), standards, and reference ontologies, the project's results have strongly influenced the wider research community and the approaches taken to a range of functions and semantic capabilities.

5 WonderWeb has helped bring semantic web technology from research to commercial reality. WonderWeb accelerated the emergence of the Web Ontololgy Language (OWL) as a semantic web standard, which is key to enabling broad-based commercialization. WonderWeb ontology development and expertise has contributed to the EU's assertion of leadership in emerging areas, e.g., : semantic grid, pervasive computing, ambient intelligence, knowledge technologies, and smart content.

6 The WonderWeb project defines 29 deliverables. By advancing beyond formulation of ontologies, WonderWeb contributed to seven distinct technology areas, which span the stages of ontology management:

- Ontology language that represents knowledge
- Programming interfaces that allow programmatic access to ontologies
- Reference ontologies that provide ontology engineering guidance
- Tools to build and/or display ontologies
- Reasoning tools for ontologies
- Ontology servers that present ontologies on the web
- Ontology stores for persistent management of ontologies.

7 WonderWeb developed toolsets, while not intended for direct commercialization, have provided points of departure for product development. Knowledge assets, such as the DOLCE reference ontologies have an extended life, somewhat independent of the tool suites used. We found some early adopter projects that have adopted the reference ontologies directly.

8 What's next? WonderWeb and other projects have made great advances in semantic web technology. These technologies are recognized as standards and additional exciting formal and developmental work is proceeding. While a lively industry of support tools and services has emerged, issues exist that need addressing by the research community in order for industry adoption to accelerate to its full potential. We call attention to those "scruffy" issues faced by practitioners in the business world attempting to engineer practical, economical, and effective ontology-based solutions.

Table 1-1: Directions for Further Development, following this page, outlines some critical, unresolved issues that have emerged in TopQuadrant's experience as barriers to successful adoption of semantic web technologies. We recommend that systematic technology transfer research be undertaken to resolve these issues and related challenges leading to industry best practices for vendors and solution providers. This will help better meet organizational and societal goals that depend on this technology.

Challenge Area	Barrier	Research Direction
1. Mapping of legacy data to ontologies.	**System implementers need to map a spectrum of legacy datasources to ontologies.**	Research to understand how semantic mapping can be done in a way that respects the complex interaction between ontologies, legacy data, applications, work culture and business process.
2. Multi-modeling paradigm support.	**Modelers working in manufacturing (STEP), enterprise architecture (PSL), software engineering (UML and MDA) and workflow (BPM) have difficulty integrating OWL with their domain-specific modeling approaches.**	Research to determine the kinds of issues that arise when combining OWL with other modeling formalisms.
3. Logic and ontology.	**Modelers from backgrounds other than semantic web (domain modeling, object modeling, data modeling) often lack the logical sophistication to produce good OWL models.**	Research and educational development to provide best practices advice based on the details of the various standard levels of OWL (OWL-Lite, OWL-DL and OWL-Full).
4. Design Patterns for Semantic Solutions.	**Domain practitioners are typically not sufficiently trained in logic to build effective knowledge models.**	Research to determine and develop domain- and industry-specific classes of problems and solution patterns that match them.
5. Ontology Engineering lifecycle management	**Ontology modelers find a need to perform global edits on their ontologies, merging some parts that were developed separately, dividing some models into smaller parts.**	Research to determine what operations are most needed to support practical ontology development and maintenance.
6. Ontology reuse.	**Ontologies must be designed for re-use, but real-world attempts to re-use ontologies have resulted in disappointing results.**	Research into frameworks for modularizing and re-using ontologies, and constructing actual reusable ontology assets for specific domains and competencies.
7. Visualization.	**It is difficult for modelers to come to a sufficient understanding of an ontology without the ability to "see" aspects of the ontology.**	Research into visualization techniques that exploit various industry standards for schematics and information graphics could smooth adoption paths, especially for science and engineering domains.

Figure 1-1: Directions for Further Research

SECTION TWO

INTRODUCTION

In October, 2004 WonderWeb project organizers asked TopQuadrant, Inc. to prepare an independent report to assess WonderWeb developed technologies, evaluate their market potential, and provide guidance on directions for further development.

TOPQUADRANT

TopQuadrant, Inc. was chosen for this assignment because the firm's consulting practice:

- Provides solution envisioning and trusted-intermediary advisory services focused on semantic web technologies, solutions, and markets.
- Conducts frequent surveys and studies of industry requirements, semantic technology providers, and early adopters of semantic solutions.
- Provides knowledge engineering and semantic solution development services for government and industry clients
- Operates from the US, which provides independence from European Union (EU) research initiatives as well as insights regarding developments in North America and Asia.

REPORT PURPOSE

The intent of this report is to assess the relevance, quality, and importance as well as strategic and commercial impact of technologies developed through the WonderWeb project. There are four objectives:

- Fulfill contractual requirements in the project plan for deliverable D25: Evaluation and Market Report.
- Gauge the significance of WonderWeb technology research and development in the broader context of global investments being made from 2000 to 2010 to advance semantic technologies for the web, grid computing, Internet services, and knowledge and information management.
- Examine the influence WonderWeb technology has had on other R&D initiatives, product commercialization, and early adoption of semantic technologies in the market place.
- Recommend next steps for further development of WonderWeb technologies which would enhance their utility for those engaged in engineering of practical networked ontology-based applications.

AUDIENCE

The primary audience for this report includes WonderWeb partners and project participants, and Information Society Technologie (IST) program reviewers and funding authorities. Also, the report addresses issues in a way that may be of interest to a broader audience of researchers, venture capitalists, and technology providers.

REPORT THESIS

The thesis presented in this report is that WonderWeb has played a key and pivotal role in enabling, shaping, and accelerating development of a wave of enabling semantic technologies that will grow to billion dollar markets by 2010.

The following topics introduce the strategy of the report. First, we review fundamentals of the Semantic Web. Next, we highlight some of WonderWeb contributions to this field. Then, we overview the methods we've used to assess the technical and commercial impact of these contributions.

SEMANTIC WEB FUNDAMENTALS

The Semantic Web is the name given to a vision for networked intelligence that will extend the capabilities of the familiar World Wide Web to allow information to be networked in a meaningful way. The value of the Semantic Web is not limited to web pages. Semantic web technologies are also applicable to corporate databases, managed document bases, multi-media resources, and other information sources that are not today thought of as part of "the web". A web of meaningful information will revolutionize information management, providing a wide array of capabilities for information sharing and dissemination within knowledge-based organizations.

Sir Tim Berners-Lee, who is credited as the inventor of the World Wide Web, has made the realization of the Semantic Web a priority for the World Wide Web Consortium (W3C). Hence the name "Semantic Web" also applies to the technologies and standards being developed and disseminated as part of that effort.

Ontologies are at the heart of the semantic web. An ontology describes the types of information that a semantic web page can express. More importantly, it describes the relationships among these types. Prior to the Semantic Web, Ontologies were defined loosely only as a representation of a shared conceptualization and varied widely in their structure, purpose and choice of representations. One of the major technological advances in recent years that has allowed the semantic web to move from the research labs into practice was the formulation of ontologies in a consistent, sharable, and standard way.

WONDERWEB TECHNOLOGY CONTRIBUTIONS

WonderWeb has played a key role in bringing the semantic web technology from a research theme to a commercial reality. The WonderWeb project defines 29 deliverables, which we have organized into four technical contributions. In Section-3 of this report we will map the broad range of semantic functions and capabilities that have been influenced by WonderWeb's development of ontology infrastructure for the Web. In Section-4 we will show how WonderWeb contributions have had a clear and positive impact on present and future semantic technology markets.

WonderWeb effort provided advances beyond formulation of ontologies. It contributed to seven distinct technology areas, which can best be characterized in terms of stages of ontology management:

- Ontology language that represents knowledge
- Programming interfaces that allow programmatic access to ontologies
- Reference ontologies that provide ontology engineering guidance
- Tools to build and/or display ontologies
- Reasoning tools for ontologies
- Ontology servers that present ontologies on the web
- Ontology stores for persistent management of ontologies.

WONDERWEB'S IMPACT ON MARKETS AND COMMERCIALIZATION

What steps does the market analysis follow? We evaluate the commercial impact of WonderWeb contributions with the larger context of investment, technology innovation, and market adoption which, to some extent, it has helped enable, shape, and accelerate. We refer to this larger cycle (by way of analogy with the ebb and rise of technology trends) as the "Semantic Wave".

Our method begins by examining the patterns (so-called "S-curves") of technology market growth (see figure 3-1) that have repeated fairly regularly since the industrial revolution. We argue that the current advances in the Semantic Web are best understood in the context of one of these waves.

Next, we model the stages of the value chain by which these innovation cycles develop and the roles of different funding sources involved with fundamental R&D, commercialization, and mainstream market adoption, and note the interrelationships between these stages.

To size the wave, we developed tables and charts in Section-3 that tell the story of cumulative investments in semantic technologies, commercialization of products, and the growth of mainstream markets to 2010. Sources consulted in compiling data are listed in Addendum-A. Addendum-B shows composite data tables used to develop estimates together with notes about preparation. The analysis is unique for drawing upon a wide variety of trend information to cross-correlate estimates. For the period 2000-2010, we present investment estimates for R&D, venture capital, and industry commercialization for North America, Europe and Asia. We estimate market growth to 2010 in several ways, by: information and communications technologies (ICT) versus semantic ones, and horizontal, vertical, and regional semantic technology markets.

We qualified projections by examining companies and commercialization activity, technical capabilities they are selling, and experiences of early adopters pursuing semantic web projects. For example, to gauge:

- Scope of WonderWeb's impact on other R&D, the market assessment examines the breadth, interrelatedness, and mutual reinforcement or amplification of different areas of R&D with semantic technologies.

- Extent, and growth rate of commercialization, the market assessment examines current productization, trends in venture capital (VC) and investment banking, and relevant activities of the larger ICT players.

- Strength of economic driving forces for adoption of semantic technology, this assessment examines research about early adoption of semantic technologies in government and industry. Cited research gives evidence for the rate of market growth that is likely in the second half of the decade.

This analysis gives us a longitudinal picture of the growth of the semantic wave. The factors that lead into the wave, and accelerate its growth include:

- Enabling semantic capabilities
- Supporting early adopters
- Promoting a wide range of research paths
- Stabilizing standardization efforts
- Contributing technology to software vendors.

We found that WonderWeb projects pioneered infrastructure that has enabled, shaped, and augmented these factors.

SOURCES

We have consulted both proprietary and public sources in preparing this report. The bibliography in Addendum-A lists publications. Also, we have drawn from interviews with researchers, technology companies, and early adopters as well as Internet research, site reviews, literature, (e.g. , we examined corporate fact sheets, investor data, customers, products and business performance). Another key source has been our experience and judgment. Additional information about the estimating methodology and sources is provided in Addendum-B.

REPORT ORGANIZATION

This report is organized into the following sections:

Executive Summary — Key findings, conclusions, and recommendations

Introduction — Purpose, scope, strategy, and methodology of the report

Evolution of the Semantic Wave — WonderWeb project is part of a major wave of innovation, which it has helped enable, shape, and accelerate.

WonderWeb in the Wave — WonderWeb technology results have contributed to broad-based standards, influenced a broad range of R&D initiatives, provided validation and confidence for commercialization, and accelerated adoption of semantic technologies.

Addenda — Bibliography, Data tables, Directions for further research, profile of Top Quadrant.

SECTION THREE

EVOLUTION OF THE SEMANTIC WAVE

FROM VISION TO MAINSTREAM MARKETS 2000 — 2010

In gauging the technology and market significance of the WonderWeb project, it helps to consider the larger context of investment, technology innovation, and market adoption which, to some extent, this project has helped enable, shape, and accelerate. "Semantic Wave" is the term we use to designate this larger cycle. The following topics explore the evolution of the semantic wave from vision to mainstream markets during the first decade of the 21st century. Specifically, we trace the following stories and relationships to the WonderWeb project:

- *Long waves of innovation* — Major conceptual advances that spur new industries occur about twice a century. The semantic wave is a key part of "distributed intelligence."

- *Anatomy of the semantic wave* — We discuss the transition from vision to mainstream markets as a global value web linking fundamental research, venture capital, and early adoption. By focusing on structural enabling technology, WonderWeb R&D has played a seminal role in this story.

- *Economics of the semantic wave* — We follow the money, mapping inputs and outputs of the building wave including estimates of funding for R&D, investments for commercialization, and rates of market adoption from 2000 to 2010. Seen from this angle, WonderWeb research has generated significant leverage from a relatively small investment.

- *Semantic technology investments* — We chart the story of investment: by whom, where, and how much, and in what time frames. WonderWeb results have provided confidence to make further investments.

- *Semantic technology R&D* — We examine what the money invested has gone towards. R&D themes in the semantic wave are varied, yet interrelated with a common semantic core. Because WonderWeb research addresses these structural foundations (such as ontology languages, APIs, and standards), it's results have influenced approaches by others to a range of functions and semantic capabilities.

- *Commercialization of semantic technology* — It's still an early market, but the signs of life are significant. We look at who is commercializing products and what capabilities they're developing. WonderWeb R&D has influenced product developments.

- *Early adoption of semantic technology* — This is how the rubber meets the road and products become whole solutions. We report on where in business semantic technologies are being applied, for what sorts of applications, and, most importantly, what business value return has been measured. Some WonderWeb deliverables have been directly incorporated by early adopters.

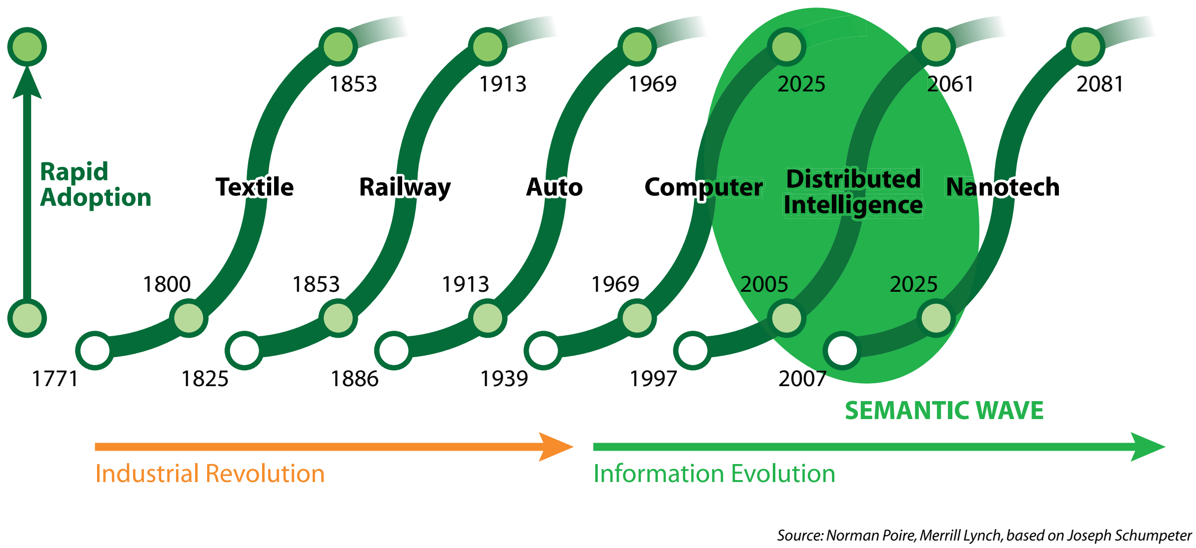


Figure 3-1: Long Waves of Innovation

Today we are at the intersection of three major innovation advances: one nearing its end, one that will continue another 10–20 years, and one that is just starting. These innovation waves spur enormous investments and radically alter the economics of affected industries. As with the computer wave, the current one, "distributed intelligence" is affecting virtually all industries.

- *Growth of mainstream markets* — We look ahead to size and project the transition to mainstream markets to 2010. We present perspectives on horizontal, regional and vertical markets. WonderWeb's legacy, in part, is to have helped accelerate these markets.

LONG WAVES OF INNOVATION

Looking back over the past two centuries, major conceptual advances that power economic growth seem to occur about twice a century. Joseph Schumpeter, an Austrian-born economist, noted long waves of industrial activity in the 1940s. More recently, Merrill Lynch analyst Norman Poire sketched out a diagram (see Figure 3-1) that illustrates Schumpeter's concept. We've added a "you are here" overlay to Poire's diagram to indicate the current intersection of waves of innovation that comprise the "semantic wave".

When traced back to the Industrial Revolution in 18th-century England, Schumpeter noticed that waves of innovation ebbed and flowed every 50–60 years. Each fresh wave had brought with it a "new economy" that led to investment and excess, followed by a shakeout—but, ultimately, as The Economist concluded, left the world a richer and better place ("A Crunch of Gears," Economist, Sept. 29, 2001). The chart shows six long waves. Inventions in cotton- spinning, iron-making, and steam power propelled the first boom. It lasted from the 1780s to the 1840s. The second wave arrived with innovations in steelmaking and railways, lasting for half a century before running out of steam around 1900. Electrification and the internal-combustion engine powered the third 50- year wave. The fourth industrial wave was launched in the early 1950s on the back of petrochemicals, electronics, computing and aerospace. The fifth wave, distributed intelligence, started in the 1970s with the precursors of the Internet. It continued with the adoption of client-server corporate networking, and rapidly accelerated following the introduction of the World Wide Web. In the wake of the dot-com shakeout, this wave is shifting into a new growth gear. That's right: Far from being over, the current wave has probably another 35 years to go. Meanwhile, a sixth wave is forming that will be powered by nanotechnology, bioscience and clean energies.

New surges of economic activity tend to play out in four distinct phases. The first phase is a period of rapid innovation as practical applications of seminal inventions emerge. The next phase brings rapid growth as successful participants—whether in cotton, railways, motorcars, electrical goods or petrochemicals— enjoy fat margins, set standards, kill off weaker rivals and establish themselves as leaders of the pack. (In the IT space, we might think of Cisco, Intel and Microsoft as leaders today; but will they continue their dominance during the next wave?) In the third phase, the market matures and the dominant firms hunker down for slower growth, which is happening now with the PC. The final phase is a short and sharp decline that occurs when the next set of technologies start jostling for the attention of investors.

ANATOMY OF THE SEMANTIC WAVE

In Figure 3-2 we map the mechanism by which the vision of the semantic wave is being transformed to mainstream markets. The vision of the semantic wave has broad societal and economic implications. Its realization entails a collaborative, multi-stage process depicted by layers in the diagram. These include fundamental research, commercialization, enterprise and consumer solutions, and mainstream adoption. Value stages tap different sources of capital, have differing investment criteria, produce different categories of output, and have differing criteria for gauging outcome success.

Stages interconnect in multiple ways. Later stages depend on the results of earlier ones. Value amplifies as we move from stage to stage. The vision becomes the basis for government funding of basic R&D. Successful technology demonstration validates concepts and triggers commercialization. Viable products are prerequisite for successful early adoption. Solutions proven by early adopters provide the references that assure mainstream buyers that the technology will produce practical benefit.

Funding sources shift and levels of investment increase as R&D moves from one stage to the next:

- Investments in fundamental R&D have high risk and a relatively long gestation period. Government funding plays a major role in the early stages. Studies point to numerous examples of public sector investments that have led to billion dollar markets, for example, reports from the National Research Council of Canada [CSTB-2002]. For the semantic wave, the progression appears to be 15 years from vision to billion dollar markets. Moreover, market impacts are expected to reach trillion dollar levels by 2020.

- Venture capital and technology commercialization focuses on market timing as much as technology readiness. More is invested with the expectation that positive returns from the market will be seen in a relatively short period.

- Early adopters seek strategic advantages from moving quickly to embrace new technologies and services. Technology costs represent only a fraction of the total investment. The measure of success is the impact that the semantic solution has on business performance or consumer values.

- Mainstream market investments are less speculative, but larger scale. They focus on making affordable investments that have demonstrable ROI in reasonable time frames with manageable risks.

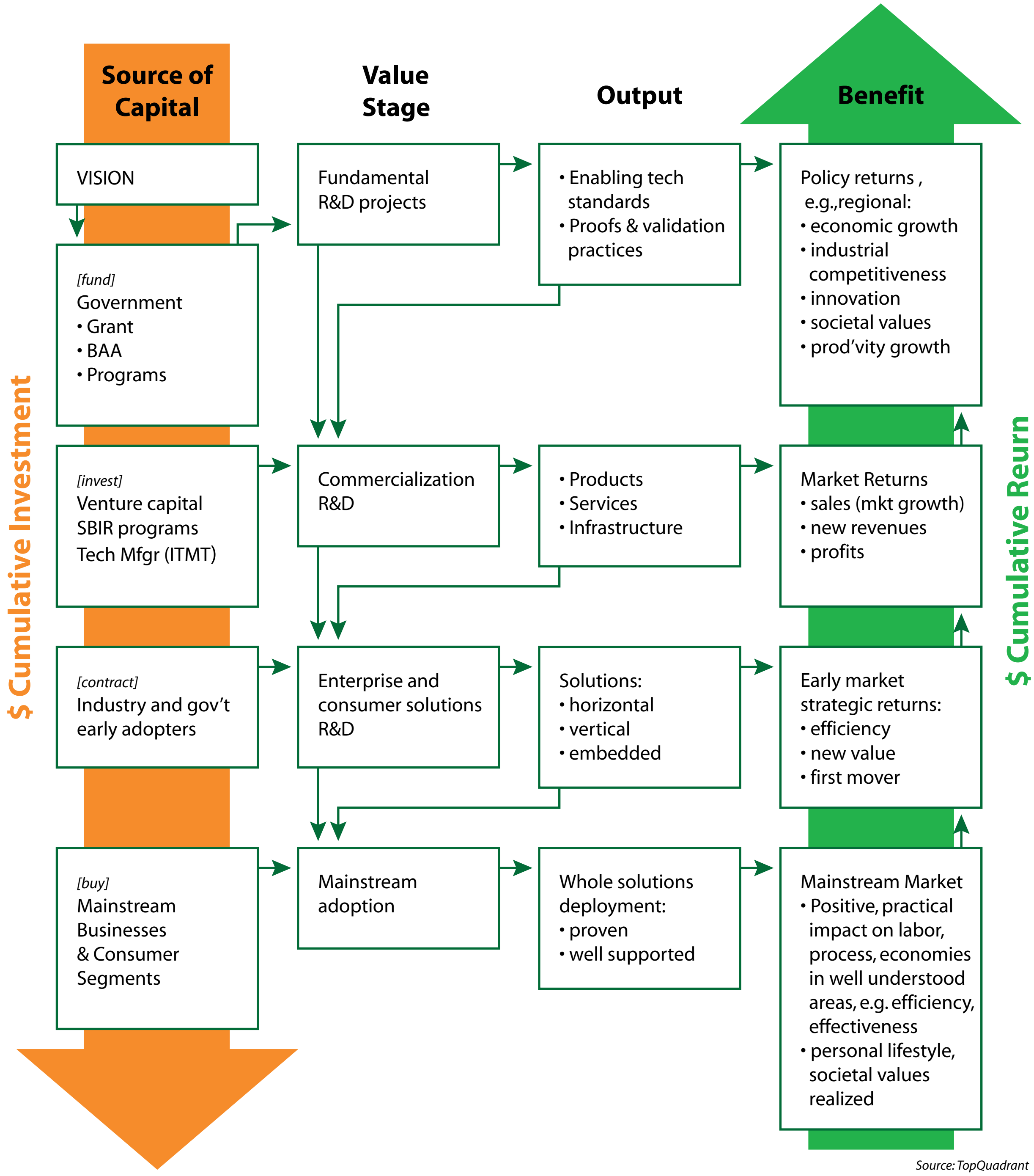
What role has WonderWeb been playing in all of this? By focusing on the development of structural enabling technology for networked ontology, WonderWeb has played a seminal part in this story.

ECONOMICS OF THE SEMANTIC WAVE

The second dimension of the semantic wave to consider is the cumulative flow of investment funding, revenues from commercialization, and economic returns to adopters. Figure 3-3 depicts this flow as a return on investment (ROI) curve that spans the period form 2000 to 2010.

The basic shape of this curve is what is important. It charts the birth of the semantic industries. Between 2000 to present, governments, venture capitalists, and technology manufacturers have made modest, but significant investments. These have led to some commercialization and early returns which we are beginning to measure in the market place. In the near-term future, 2005 to 2007, the curve trends upwards. From 2008 to 2010 the curve points to accelerated growth to a multibillion dollar mainstream market for semantic technologies, solutions and services that delivers strong return on investment for sellers and buyers alike. The data tables used to develop this graphic are discussed in Addendum-B.

Seen from this angle, WonderWeb results have generated a significant multiplier on it's efforts, from a relatively small investment (about 2M euros).



Source: TopQuadrant

Figure 3-2: Anatomy of the Semantic Wave

This diagram presents a model for thinking about technology innovation as a value chain — a dynamo that drives new ideas into basic and applied research, then to products and soloutions, then to billion-dollar markets with significant social and economic benefits. Semantic technologies are following this model, and generating value at each stage. Also, semantic technologies are now recognized as delivering value to and forming an integral part of many other lines of investment that are deemed important priorities such as grid computing, cognitive systems, smart content, and ambient intelligence. WonderWeb has been a rewarding project for the EU because it helped jump-start and amplify returns from these value linkages.

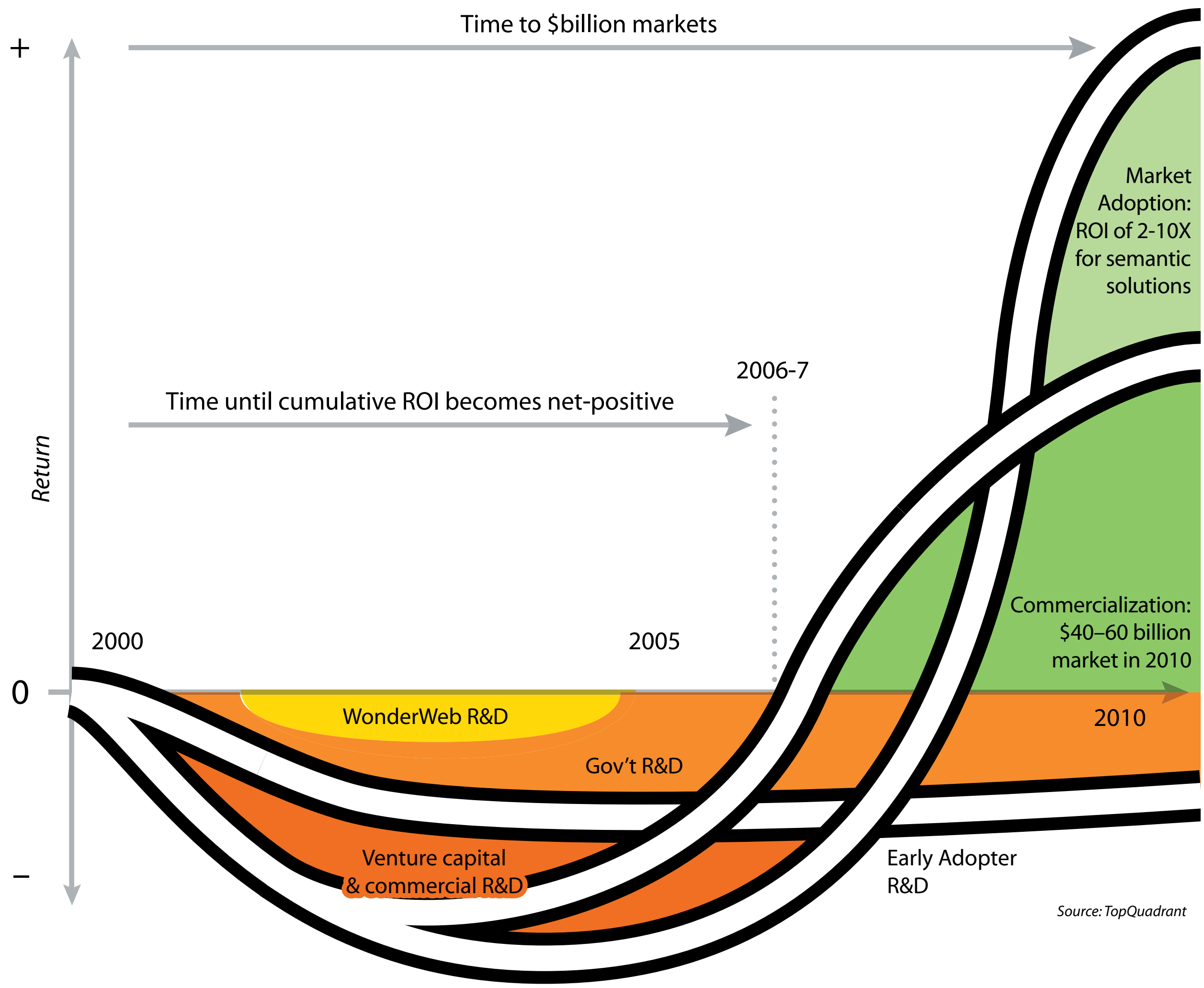


Figure 3-3: Economics of the Semantic Wave

This diagram depicts the cumulative flow of investment into and returns from semantic technology for the period 2000 to 2010. The three lines of investment include: (1) government R&D funding, (2) Venture capital investment and ICT company product/service commercialization, and (3) early adopter R&D. The two major lines of market return are: (a) returns (sales) for the technology vendors who've been commercializing the semantic technology; and (b) returns for the adopters of semantic technologies, which are measured in terms of gains in efficiency, effectiveness, and edge. Positive is up. Negative is down. Time moves left to right. So long as investment outlays exceed returns the line trends down. When returns start exceeding investments (andtotal cost of ownership) then the line moves upward. The semantic wave will deliver a strong positive return on investment (ROI).

SEMANTIC TECHNOLOGY INVESTMENT

Let's consider the investment story.

What sources of funding are significant? There are three kinds of funding: government R&D, venture capital, and industry commercialization. Figure 3-4 estimates total R&D funding for semantic technologies by type of investment. To date, government R&D has represented the largest share of funding. This is changing. Venture capital is heating up, and industry R&D is increasing. Market-based investment will dominate in the second half of the decade. WonderWeb results have helped stimulate confidence on the part of stakeholders to initiate and make subsequent rounds of investment.

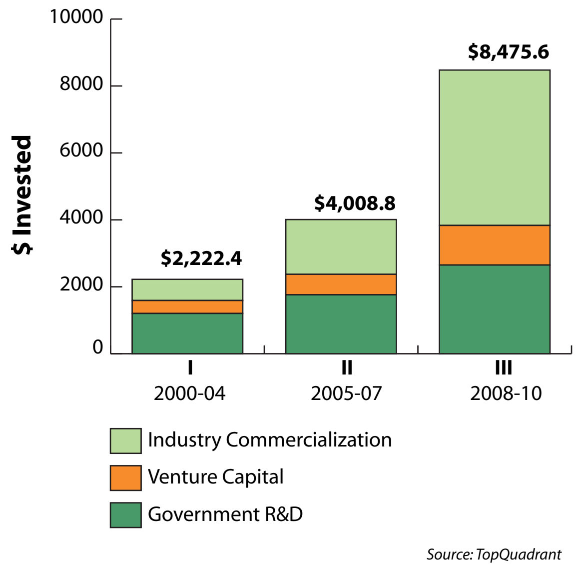


Figure 3-4: Semantic Technology Investment to 2010 by Type ($ US Billions)

This figure summarizes the magnitude of investments in semantic technologies from three sources: government R&D funding, venture capital, and industry commercialization. The series provides three snapshots: the immediate past, the near-term future, and the mid-term to 2010. Government funding of semantic technology from 2000-2004, which includes the WonderWeb project, has given the EU an edge. Venture capital and industrial commercialization will play an increasing role as the semantic wave builds.

From a global perspective, which regions are the most significant funding sources? In this analysis, we consider three regions: North America, Europe, and North America. Figure 3-5 estimates total R&D funding for semantic technologies in these regions for the period 2000 to 2010.

Development of semantic technology is being pursued globally and significant resources are being expended. Semantic technologies are universally perceived as essential to the global knowledge economy. A relationship between investment in R&D and positive economic growth is generally recognized.

An early start and relative focus on semantic and knowledge technologies has given the Europe an overall lead. Other countries, however, have been ratcheting up R&D investments (as a percentage of GDP) to make their economies more competitive in the coming era. R&D, it seems, is becoming a new global "arms race," with both US and Asia gaining.

R&D strategies and emphasis differ in different regions:

North America: US, Canada

More than half of US government R&D is focused on security, intelligence and defense. Cognitive systems, semantic web, agents, [e.g. DARPA], net-centric operations, semantic command and control, enterprise integration, semantic interoperability, autonomous operations, intelligence Q&A, cyber infrastructure, and semantic grid, are a few of the research areas receiving funding but typically unreported for security reasons.

Not generally recognized in Europe is the broad scope of semantic and knowledge technology R&D across government agencies in the US. It is significantly broader and more intense than previously reported. For example, more than a two-dozen agencies and government laboratories contract R&D through SBIR programs, and BAAs. All programs we examined with ICT components have semantic and knowledge technology research topics.

Further, federal end-user organizations are involved. Since 2003, the Federal CIO Council has convened a "semantic interoperability community of practice" (SICoP) that encompasses all Federal Agencies representing approximately $70B in annual ICT spend.

Today, semantic technologies are part of the mainstream research agenda in the US at both fundamental, and applied levels. The shift in R&D is from preoccupation with hardware aspects of new computing and communications infrastructure to address "cyber" aspects.

Europe: EU-25

The stated goal of the EU is to raise R&D investment across the EU-25 from about 1.9% in 2000 to about 3% of gross domestic product (GDP). R&D intensity in the US is about 2.8% of GDP.

The EU placed early emphasis on semantic and knowledge technologies. In the 2000-2004 period IST's FP5 and FP6 programs committed over $3B towards R&D. Of this budget, we estimate that more than 500Meuros went to fund semantic technology R&D through about 350 IST projects. In March 2004, the EU proposed doubling the IST budget for FP7 in the period 2007–2013.

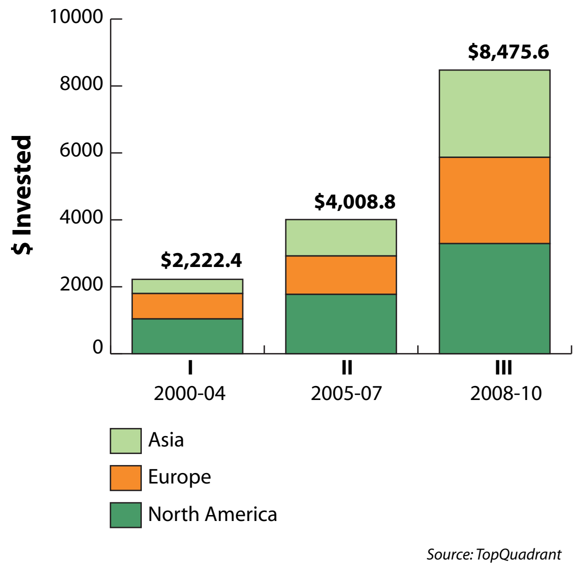


Figure 3-5: Semantic Technology Investment to 2010 by Region ($ US Billions)

This figure summarizes global investment in semantic technology by region of origin. This includes North America, Europe (EU-25), and Asia. The series distinguishes immediate past, near-term future, and mid-term future investment to 2010. Investment is growing in all regions. Funding growth in North America and Asia will challenge Europe's early leadership in semantic technologies.

Asia: South Korea, Japan, China, India

Asia is the fastest growing region. It will lead in ICT investment by the end of the decade. Tables in Addendum-B summarize data from the following countries: South Korea. Japan, China, and India.

Asian countries' ICT strategies emphasize communications infrastructure build-out, mobility, and next generation Web concepts. South Korea leads the world in broadband penetration. Japan is a close second in broadband deployment and has adopted a second-iteration of its successful eJapan strategy. China's economy ($6.5T) is second only to the US ($11T) in size. The World Bank projects that China's economy will become the world's largest by 2010-2012. While R&D intensity in China is lower than the EU, and weighted towards applied research, China has focused research on Semantic Grid as a major national effort. In the near future, India will become the world's fourth largest economy (after the US, EU, and China). It's R&D is directed internally on communications infrastructure and externally on ICT services outsourcing.

SEMANTIC TECHNOLOGY R&D

R&D themes in the semantic wave are varied, yet interrelated with a common semantic core. Because WonderWeb development addressed these foundation issues (such as ontology language, APIs, standards, and reference ontologies), its results have strongly influenced approaches by other researchers to a range of functions and semantic capabilities.

Emergence of semantic web standards

The emergence of semantic web standards accelerates application and market development. In Section 4 we will discuss the role WonderWeb has played in bringing forth web ontology standards.

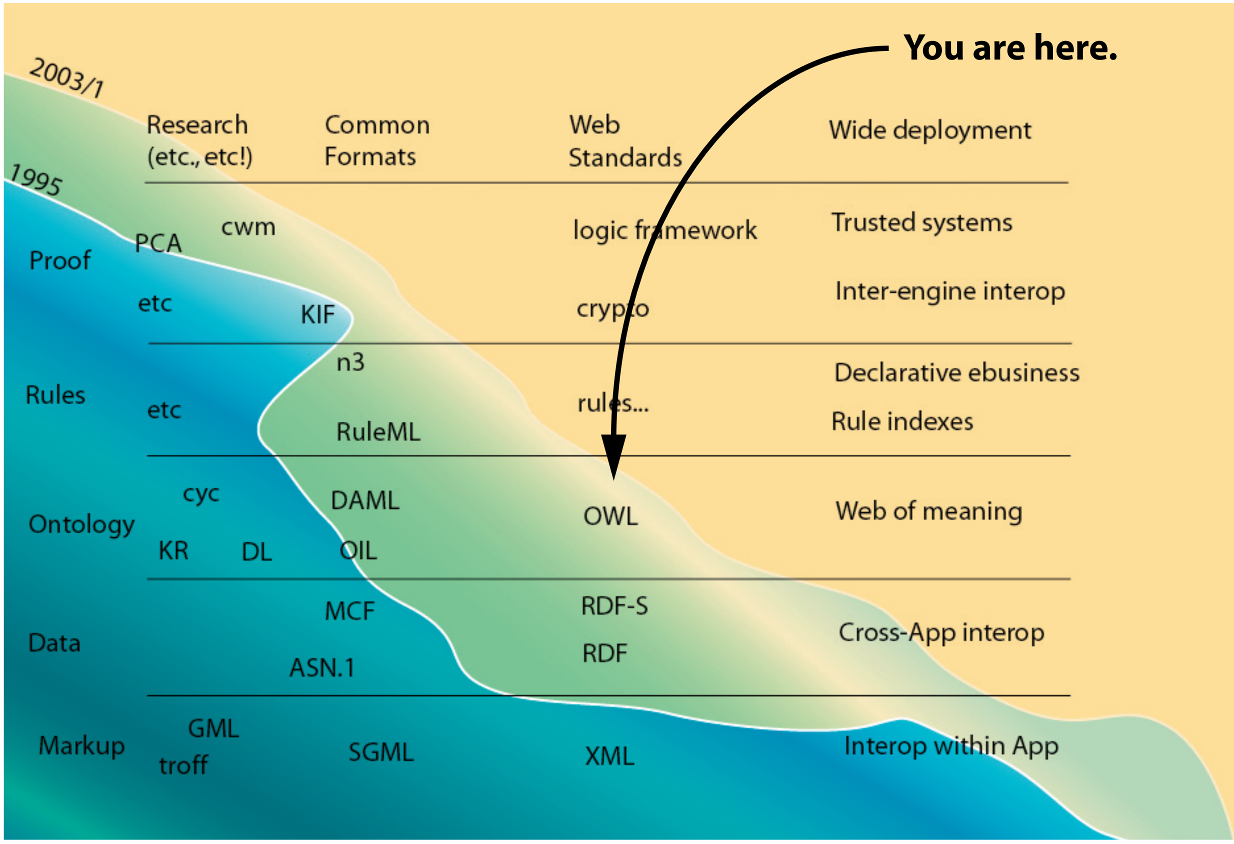
In February 2004, the World Wide Web Consortium announced final approval of two key Semantic Web technologies, the revised Resource Description Framework (RDF) and the Web Ontology Language (OWL). RDF and OWL are Semantic Web standards that provide a framework for asset management, enterprise integration and the sharing and reuse of data on the Web. These standard formats for data sharing span application, enterprise, and community boundaries—all of these different types of 'users' can share the same information, even if they don't share the same software.

This announcement marked the emergence of the Semantic Web as a broad-based, commercial-grade platform for data on the Web. The deployment of these standards in commercial products and services signals the transition of Semantic Web technology from what was largely a research and advanced development project over the last five years, to more practical technology deployed in mass market tools that enables more flexible access to data, content, and knowledge on the Web.

Semantic Web-enabled software using RDF and OWL include:

- Content creation applications: Authors can connect metadata (subject, creator, location, language, copyright status, or any other terms) with documents, making the new enhanced documents searchable.

- Tools for Web site management: Large Web sites can be managed dynamically according to content categories customized for the site managers



Source: Tim Berners-Lee, 2003

Figure 3-6: Emergence of Semantic Web Standards

In this 2003 diagram, Tim Berners-Lee charted stages of research leading to progressive emergence of the stack of semantic web standards from 1995 to present. This chart uses the metaphor of high the watermark of an advancing wave along a beachfront to describe the progression of semantic-web technologies from research to common formats, to web standards and to wide deployment.

- Software that takes advantage of both RDF and OWL: Organizations can integrate enterprise applications, publishing and subscriptions using flexible models

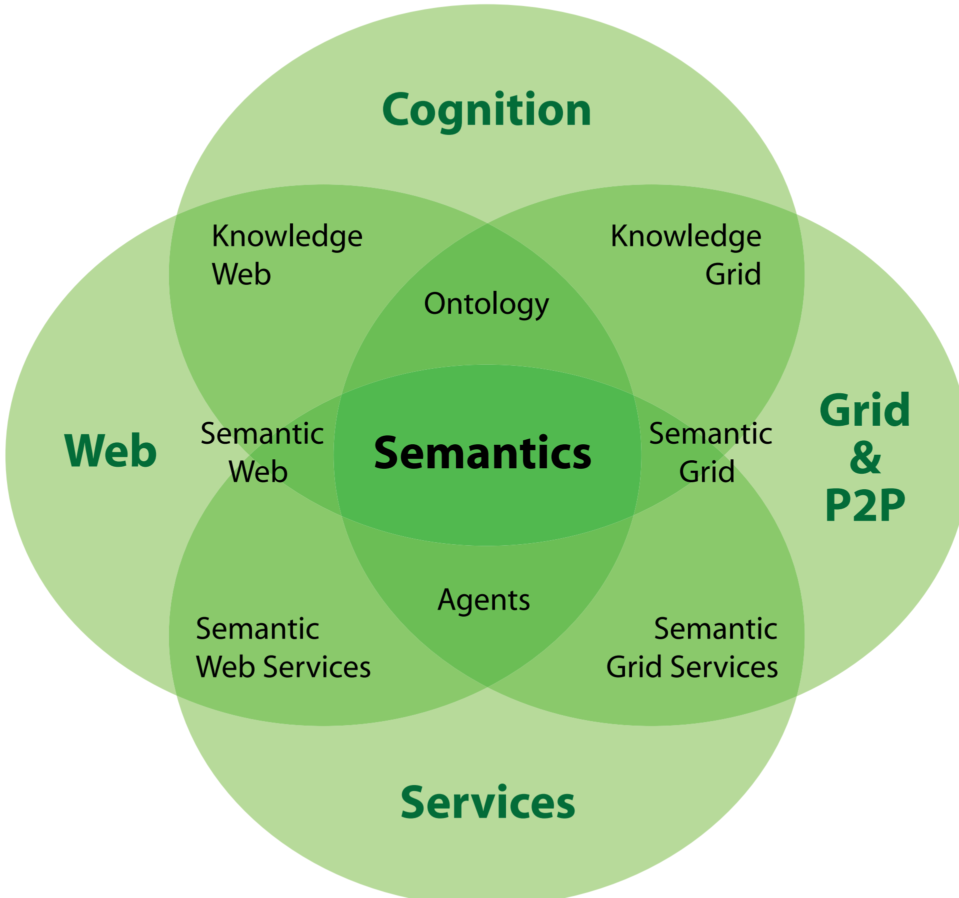
- Cross-application data reuse: RDF and OWL formats are standard, not proprietary, allowing data reuse from diverse sources.

In this 2003 diagram (Figure 3-6), Tim Berners-Lee charted stages of research leading to progressive emergence of the stack of semantic web standards from 1995 to present. This chart uses the metaphor of high the watermark of an advancing wave along a beach front to describe the progression of semantic-web technologies from research to common formats, to web standards and to wide deployment. Beyond the current position reached with the dissemination of OWL, additional research efforts continue towards open standards for enabling logic, comprehensive reasoning, and trust across ontologies and knowledge bases deployed across semantic webs, grids, and P2P infrastructures.

EU R&D themes in the semantic wave

Current EU IST research is guided by the "Ambient Intelligence (AmI)" vision where the user or smart player is placed at the centre of future developments. The focus is on technologies in which computers and networks will be integrated into the everyday environment, making a multitude of services and applications accessible through easy-to-use human interfaces.

As Figure 3-7 depicts, semantic technologies have become a central concern for broadening array of research and development initiatives. This diagram visualizes the intersections of four major develop-



Source: TopQuadrant

Figure 3-7: R&D Themes in the Semantic Wave

This diagram depicts the intersections of four major development themes in the semantic wave: cognition, the Web, grid & P2P, and services. Semantics are central to all of these. A broad range of R&D initiatives have some sort of semantic or knowledge technology component.

ment themes in the semantic wave: cognition, the Web, grid & P2P, and services. Semantics are central to all of these. These considerations serve to underscore the strategic importance and breadth of impact that WonderWeb work on ontology infrastructure has across the research community.

Researchers have examined interrelationships of European Union funded FP5 and FP6 projects with semantic and knowledge technologies:

- [Giaglis-2002] found that out of the 1725 proposals funded by the European Commission in the first six calls of the IST Programme (1998-2001), 316 (more than 18% of the total, representing close to $1 billion in funding) were marked as addressing the technological area of 'Knowledge and Information Management' (KIM).

- [Stork-2002] reviewed 23 FP6 projects that focused on development of automatic and interactive tools for making semantics explicit or for acting on semantics. Three projects, including WonderWeb, were developing general frameworks to support ontology infrastructure. Additionally, this study examined the role of semantics and knowledge technologies as enablers of intelligent web, grid, P2P, and ambient intelligence services.

- A later review of FP5 & FP6 projects [Ding-2003] reported that more than 350 projects dealt with themes that have strong semantic technology components, including: knowledge and information management, agents, ontology, simulation, cognitive systems, optimization and decision support, and supply chain management.

- Recognition of the central role of semantic and knowledge technologies across current and planned IST programs, and the need to coordinate interdependencies and potential synergies between projects, has been built in to the fabric of current programs such as Knowledge Web, next generation grid research [Obozinski-2004], and notices of upcoming calls for proposals.

Core semantic technology capabilities targeted by R&D projects between 2000-2004

Between 2000-2004, projects in the US, EU, Asia, ROW have been researching and developing a range of semantic technologies, proofs/validations, and demonstrations.

Some projects have addressed core enabling technologies while others have addressed services, horizontal applications, and vertical solution deployments to be built on semantic foundations.

Figure 3-8 highlights functions and capabilities of semantic technology that have been the focus of R&D. These functions:

- Enable widespread deployment of semantic technologies, standardize languages, APIs, frameworks, and infrastructure for knowledge representation, semantic UIs and agents.

- Discover, acquire and create semantic metadata, semantic technologies sense, discover, recognize, extract information, encode metadata, and tag content. Semantic capabilities that provide these functions include recognizers, auto-taggers, content annotators, automatic categorizers, and terminology brokers.

- Represent, organize, integrate and interoperate meanings and resources. Semantic technologies model, classify, categorize, index, cross-reference, interconnect, map, bridge, federate, store, and manage. Semantic capabilities that provide these functions include : semantic web service discovery, composition, choreography, registry, semantic web service advisor, data integrator, application integrator, process integrator, ontology builder, and semantic broker.

- Reason, interpret, infer, and answer using semantics. Semantic technologies query, search, meta-search, retrieve, match, mediate, simulate, analyze, answer, explain, plan, schedule, and optimize. Semantic capabilities that provide these functions include assistants, agent, pattern and connection explorer, match-maker, mediator, design advisor, simulator, configurator, trade-off analyzer, planner, scheduler, optimizer, and autonomous learner.

- Provision, present, communicate, and act using semantics. Semantic technologies create composite applications, browse, navigate, author text, draw graphics, create languages, and orchestrate services. Semantic capabilities that provide these functions include semantic portals, semantic browsers, application composer, concept-based searcher, context-aware retriever, expert locator, enhanced search query, social network navigator, interest-based deliverer, natural language generator, forms generator, document generator, engineering drawing generator, display and media producer.

COMMERCIALIZATION OF SEMANTIC TECHNOLOGIES

As of mid-decade, close to 60 firms are actively commercializing semantic technologies. Figure 3-9 identifies some of them. WonderWeb research has accelerated and, in some cases, directly influenced product development, as we will see in Section-4.

Venture Capital

Venture capital is heating up in the semantic technology space. Existing start-ups are moving to series-B and higher rounds of funding. More money is flowing to new start-ups as well. Our analysis of 2004 VC funding in the US [PricewaterhouseCoopers-2004] shows that 55% of deals are for ICT technology and services. Two-thirds of these are for Internet related software, services and hardware, of which between 10-20% now appear to be semantic and knowledge technology related start-ups.

Function	Activity	Semantic Capabilities
Discover, acquire, & create semantic metadata	Sense Discover Recognize Extract Encode Tag Represent	Recognizer Automated tagger Content annotator Auto-categorizer Terminology broker
Represent, organize, integrate, and interoperate meanings & resources	Model Classify Categorize Index Cross-reference Interconnect Map Bridge Federate Store Manage	Semantic service discovery Semantic web service composition Semantic service choreography Semantic content registry/advisor Semantic data integrator Semantic content integrator Semantic application integrator Semantic process integrator Ontology builder Semantic broker
Reason, interpret, infer, & answer with semantics	Query Search Meta-search Retrieve Match Mediate Simulate Analyze Answer Explain Plan Schedule Optimize	Assistant Agent Pattern & connection explorer Match maker Mediator Design advisor Simulator Configurator Trade-off analyzer Planner Scheduler Optimizer Autonomous learner
Provision, present, communicate, and act using semantics	Composite apps Browse Navigate Author text Draw graphics Create documents Orchestrate services	Semantic portal Semantic browser Application compositor Concept-based search Context-aware retrieval Expert locator Enhanced search query Social network navigator Interest-based delivery Natural language generator Forms generator Document generator Information display generator Engineering drawing generator Media generator

Source: TopQuadrant

Figure 3-8: Semantic Technology Capabilities

This table highlights functions and capabilities of semantic technology that have been the focus of R&D. The table is divided into three columns. Functions and related activities are shown to the left. Semantic capabilities are listed to the right. Capabilities package semantic technology functionality to meet business needs.

Function	Sample Solution Providers		
Discover, acquire, & create semantic meta data	Autonomy Captiva ClearForest Convera Copernic	Endeca Entrieva Factiva FAST H5 Technology	Interwoven Inxight Primus Stratify Verity
Represent, organize, integrate, & interoperate meanings & resources	Celcorp Cogito CognIT Connecterra Contivo Digital Harbor Empolis Grand Central	Enigmatec Intellidimension KFI L&C Metamatrix Miosoft Modulant Network Inference	OntologyWorks Ontoprise SchemaLogic Semagix Triple Hop Tucana Unicorn Vitria
Reason, interpret, infer, & answer using semantics	HP IBM Network Inference	NuTech OntologyWorks Ontoprise	Open Cyc Unicorn
Provision, present, communicate, and act using semantics	AT&T BEA Black Pearl BTexact	Merant Pinnacor SAP	Semaview Software AG Troux Vignette

Source: TopQuadrant

Figure 3-9: Commercialization

This table identifies approximately 60 firms that are commercializing semantic technologies This list is representative, and is by no means exhaustive. Vendors appear alphabetically grouped by major function. Since most vendors provide multiple capabilities in their product lines, we selected the category that best fit with early adopter research. Vendors comply with standards for the semantic web in varying degrees. TopQuadrant research found that companies whose proprietary technologies and capabilities were developed before semantic web standards were formally disseminated are now in process of supporting them.

Web 2.0

A key emerging direction for Internet investment in the US is called *Web 2.0*. This next wave encompasses new categories of consumer, business, and community-oriented business models. It explicitly embraces semantic technologies. Morgan Stanley's 2003 study [Meeker 2003] provides a useful structural assessment of the industry and market space as well as an investment rationale for capitalizing on this next wave.

ICT vendors

ICT vendors are researching semantic technologis. Industry giants and almost all of the major IT platform providers, integrated development environment (IDE) vendors, application software providers, content and database management software companies, and telecom firms have started research programs with semantic web and knowledge technology themes. In the next cycle (2005-2007) they are expected to begin to enter the market with commercial products.

EARLY ADOPTION OF SEMANTIC TECHNOLOGIES

To date, semantic technologies have been characterized as an R&D market, or an early market. The majority of expenditures to date have been for R&D more than for operational deployments.

Whenever an industry (or in this case, an entire economy) is about to retool to make a transition from one category of technology to another, there

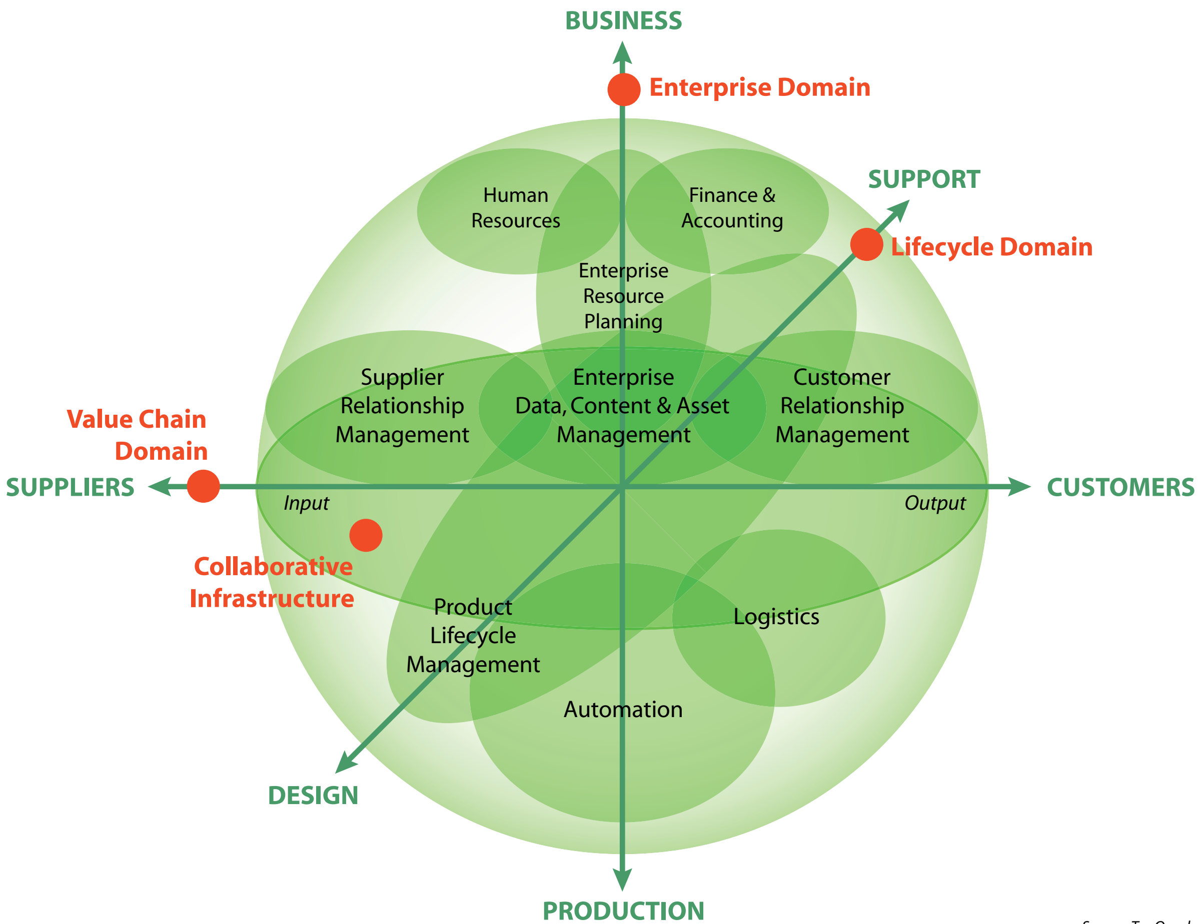


Figure 3-10: Business Domains Where Semantic Technologies Can

Semantic technologies can be applied and deliver value in multiple areas of an enterprise. The diagram is divided into several business domains:

1. *Enterprise domain* is the business<—>production axis. The business processes include enterprise resource planning, human resources, finance and accounting, and business intelligence. Production processes include execution and automation systems.
2. *Lifecycle domain* is the design<—>support axis. The key process is product lifecycle management, which includes research and development, design, engineering, manufacturing, distribution, and support.
3. *Value chain* domain is the customer—supplier axis. The key processes are customer relationship management involving marketing, sales, and customer service; and supplier relationship management, which involves planning, sourcing, making, and delivery.
4. *Collaborative infrastructure* is the integrative platform (shown here as a disc around the equator) that interlinks and supports these domains. The processes involve information input, enterprise data, content and asset management, and information output.

needs to be a compelling economic driving force. Semantic technology is still evolving. The transition to open standards is a work in progress. And, the practices for building semantic solutions are immature. Still, early adopter case examples show that semantic technologies are passing the acid test for business value. It delivers 2-10 times improvement in performance for applications addressing core business problems. Benefits of this order make a compelling value proposition.

During 2004 TopQuadrant has researched early adoption of semantic technologies by government, industry, and consumers [Davis-2004]. While it is beyond the scope of this report to review early adopter case studies in any detail, the following topics summarize some of the key findings.

Figure 3-10 depicts areas in business and government where semantic technologies can be applied and deliver value. The diagram is divided into several business domains:

- *Enterprise domain* is the business<—>production axis. The business processes include enterprise resource planning, human resources, finance and accounting, and business intelligence. Production processes include execution and automation systems.

- *Lifecycle domain* is the design<—>support axis. The key process is product lifecycle management, which includes research and development, design, engineering, manufacturing, distribution, and support.

- *Value chain domain* is the customer<—>supplier axis. The key processes are customer relationship management involving marketing, sales, and customer service; and supplier relationship management, which involves planning, sourcing, making, and delivery.

- *Collaborative infrastructure* is the integrative platform and fabric that interlinks and supports these domains. The processes involve information input, enterprise data, content and asset management, and information output.

Figure 3-11 calls out ten specific application areas where semantic technologies are being applied today by early adopters in business and government and gives examples. Listed clockwise, these areas are: (1) infrastructure and integration, (2) managing risk, (3) customer-facing services, (4) output management, (5) smart products and services, (6) design and manufacture, (7) research, (8) input management, (9) supplier facing processes, and (10) intelligence.

The research sample included 35 early adopter cases. These involved enterprise, rather than consumer applications. The cases come from many different verticals, for example: government, financial services, manufacturing, logistics, transport and communications, health services, media, and business services. They were selected because they solved a problem "semantically", (as defined in Figure 3-8). We deliberately looked for case examples that showed a significant gain in efficiency, or effectiveness, or both. Case examples did not necessarily use products based solely on semantic web standards to effect the semantic solution. OWL came out less than 6 months before the study began. Almost all involved some commercial-off-the-shelf (COTS) component.

The key finding from the TopQuadrant research is that semantic technology applications are delivering business value to early adopters. Figure 3-12 summarizes gains in performance that the study documented.

The classic motivations for investing in new technologies are basically three:

Efficiency gain — Doing the same job faster, cheaper, or with fewer resources than it was done before. The key measurement is cost savings.

Effectiveness gain — Doing a better job than the one you did before, making other resources more productive and improving attainment of mission. The key measurement is return on assets.

Business edge — Changing some aspect of what the business does, resulting in growth, new value capture, mitigation of business risk, or other strategic advantage. The key measurement is return on investment.

The acid test for any new technology or solution investment is this: do the benefits represented by these "3-Es" outweigh the costs and risks associated with making the change. If the answer is "yes", we have something to consider. If not, the matter ends. Here are some of the ways that semantic solutions are delivering business value today:

Efficiency gains

Semantic technologies can have a dramatic impact on labor hours, cycle time, inventory levels, operating cost, development time, and development cost. Case examples showed 20-90% reductions in these measures.

Semantic technologies drive efficiency gains in multiple ways, for example:

- *Semantic information access* — applied to content and data sources, semantic automation allows people to cope with information overload. Semantic solutions allow them to access, tag, extract, classify and make sense of vastly more and varied sources than ever previously possible. This matters, for example in research, intelligence, and publishing. But it impacts any area requiring search.

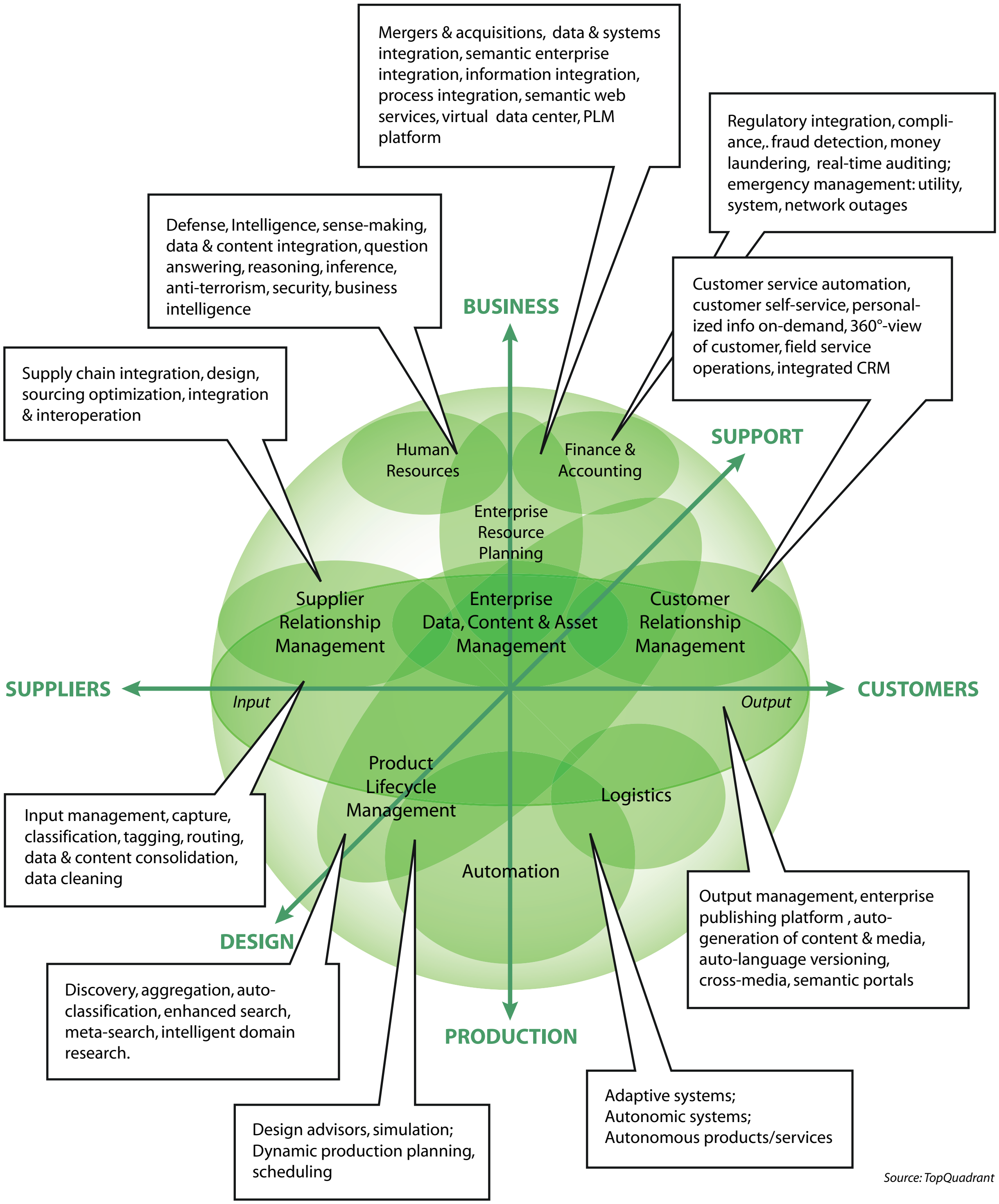


Figure 3-11: Early Adopter Application Areas

This figure calls out ten specific areas where semantic technologies are being applied today and gives examples of early adopter applications in business and government. Listed clockwise from the top, these areas are: (1) infrastructure and integration, (2) managing risk, (3) customer-facing services, (4) output management, (5) smart products and services, (6) design and manufacture, (7) research, (8) input management, (9) supplier facing processes, and (10) intelligence and security. These applications are based on findings from 35 semantic technology early adopter case studies reported by TopQuadrant in *"The Business Value of Semantic Technologies"*, September 2004.

EFFICIENCY	EFFECTIVENESS	EDGE
Cost savings Doing the same job faster, cheaper, or with fewer resources than it was done before	*Return on assets* Doing a better job than the one you did before, making other resource more productive and increasing their return on assaets and attainment of mission	*Return on investment* Changing some aspect of what the business does, resulting in growth, new value capture, mitigation of business risk, or other strategic advantage
IMPACT OF SEMANTIC TECHNOLOGIES*		
20-80% less labor hours 20-90% less cycle time 30-60% less inventory levels 20-75% less operating cost 25-80% less set-up & development time 20-85% less development cost	50-500% quality gain 2-50X productivity gain 2-10X greater number or complexity of concurrent projects, product releases & units of work handled 2-25X increased return on assets.	2-30X revenue growth 20-80% reduction in total cost of ownership 3-12 month positive return on investment 3-300X positive ROI over 3-years

*Source: TopQuadrant

Figure 3-12: Business Value

This table identifies three primary categories of business value and summarizes the potential impact that semantic technologies can have in each area, based on early adopter case studies. TopQuadrant research has documented 2-10 times improvements in key measures of business performance for semantic technology applications. This sort of value proposition is compelling for any technology.

- *Enterprise and supply chain integration* — Semantic integration changes the economics of bridging disparate and redundant data sources, and linking legacy systems together. Huge savings are possible. Today's motivation for using web services is to take cost and time and effort out of the process of integrating applications and maintaining interoperability. Exposing functionality as web services can reduce total cost of ownership (TCO) by significant percentages. However, it still requires people to construct the interfaces, even when they can find out about the service in a directory. Moving to semantic web services enhances the value proposition. Service ontologies provide a way to take people out of the loop of configuring and managing the integration. The idea here is akin to "autonomics": software with self-knowledge that can auto-configure, self-maintain, and self-manage — resulting in even better TCO.

- *Knowledge-centered manufacturing* — Semantic virtualization means driving a product life cycle (from research, design, engineering, manufacturing, through in-service maintenance and sunset) using knowledge-based models. Together with simulation and auto generation of drawings and technical document, semantic virtualization dramatically increases productivity, and improves the quality, speed, and economics of design and manufacturing.

- *Generative communications* — Another area where semantic technologies impact efficiency is generating the right content, in the right media format, delivered to the right audience, through the right media channel, with the right economics. A knowledgebase that represents the meaning of content separately from the language forms, as well as the attributes of audience, language, media, and delivery channel allows semantic technologies to automatically generate audience-specific text, drawings, documents, animations, and media displays directly from the knowledgebase. Huge savings in time, labor, and cost are possible.

Effectiveness gains

Semantic technologies can drive dramatic improvements in quality, service levels, and productivity. Combined with process improvements, these can allow existing staff to handle a greater number (or complexity) of concurrent projects, product releases, and units of work. Case examples showed increases in effectiveness and return on assets from 2-50 times.

Business edge

The strategic value of semantic technologies is the ability to create new advantages. Case examples evidenced this in several ways, for example: 2-30X revenue growth; 20-80% reduction in total cost of ownership; 3-12 month positive return on investment; 3-300X positive ROI over 3-years; mitigated risk and reduced vulnerability to fraud, liability, or litigation; and improved odds of achieving or exceeding mission objectives.

Over the past five years, early adopter experiences with semantic technologies have moved from bleeding edge experimentation to high-yield operational deployments. While early adopter experience and returns will never all be uniformly positive, we are encouraged by the numbers of positive outcomes, the diversity of applications, and the range of industry verticals represented. The semantic wave is clearly building, and now appears poised to "cross the chasm" from early adoption to mainstream markets.

GROWTH OF MAINSTREAM MARKETS

What is the market outlook for semantic technologies to the end of the decade? We present four views of the emergence of semantic technologies as a mainstream market from the period 2003 to 2010, namely:

- As a component of ICT spending
- By Horizontal segment
- By Vertical segment
- By region

Projections of semantic technology mainstream markets are based on analysis of studies by government agencies and ICT analysts. These market projections are preliminary estimates. Supporting data tables and sources are discussed in Addendum-B. Some pertinent notes follow:

- Statistics on global R&D (governments, venture capital, and business) are based on publications from OECD, AAAS, NSF, WITSA, PricewaterhouseCooper, and MorganStanley.

- Statistics on ICT markets are based on studies by WITSA, IDC, Ovum, Gartner Group, Morgan Stanley, McKinsey & Company. These examined: knowledge management access, infrastructure, and ontology-based solutions; enterprise & commercial publishing; enterprise-class applications such as ERP, CRM, BI, PLM, help desks, and portals; web services, grid computing, and pervasive computing (e.g. RFID); and the largest category, integration software and services.

- Estimates of the amount of R&D for semantic technologies are based on proprietary research by region, into the technology coverage of funded government research, the technology coverage of venture capital deals, and the commercialization activities of ICT vendors with interests in semantic technologies.

- The mainstream market estimate was based on a number of ICT-industy studies that provided the base line and rationale for end-points of the projection. They also gave insight into different horizontal application areas that would likely "go semantic" to some extent — for example the "enterprise integration" market.

Growth of semantic and information technology to 2010

Figure 3-13 compares the growth of ICT (information and communications technology) markets with markets for semantic technology. It highlights relative size (in $ billions) and growth rate. The baseline year show to the left. The forecast year is shown to the right. And the compound annual growth rates (CAGR) appear in the arrows connecting the two pie charts.

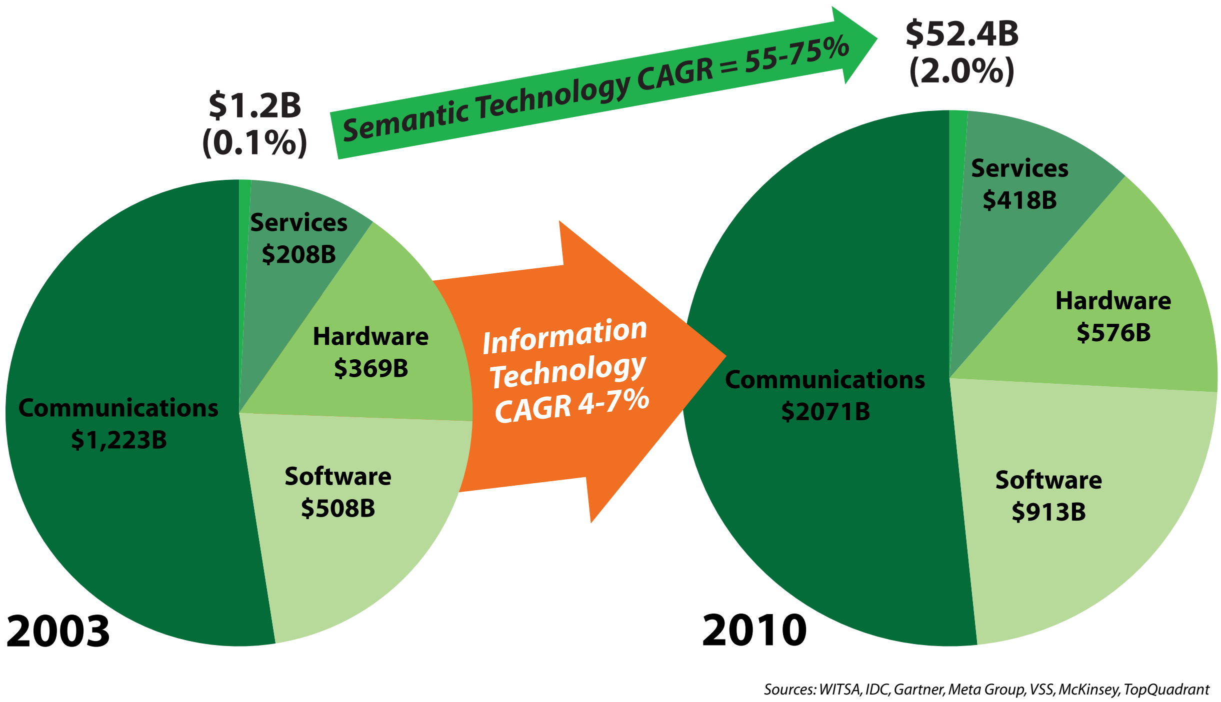


Figure 3-13: Semantic and Information Technology Growth to 2010
These two charts project the growth of mainstream markets for semantic and information (ICT) technologies to 2010. It highlights differences in relative sizes of these markets (in $ billions) and forecasts growth rates. The baseline year, 2003, is shown to the left. The forecast year, 2010, is shown to the right. Compound annual growth rates (CAGR) appear in the arrows connecting the two pie charts. The numbers for information technology represent total sales of software, services, and hardware. The numbers for semantic technology represent a composite of services and software. Supporting data tables are provided in Addendum A.

The numbers for information technology represent total sales of software, services, hardware. The numbers for semantic technology represent a composite of services and software.

Information technology is projected to grow for the rest of this decade at an overall growth rate of 4-7%, with services growing fastest, and hardware the slowest.

Semantic technology during this period is projected to grow from around $1.2 billion in 2003 to about $52.4 billion with a compound annual growth rate of 55-75%%. Given the significant ROIs being realized by early adopters of semantic solutions reported here, we believe that this level of growth will prove to be both reasonable and sustainable. Applying a sensitivity threshold of plus or minus 20%, yields a market size between $40-60B in 2010.

Semantic technology horizontal markets
Figure 3-14 presents horizontal sizing of mainstream markets for semantic technologies by functionality category in 2010.

The diagram depicts five horizontal segments as follows:

- *Discovery and access* — Sense-making, data and content mining, moving from documents to knowledge-centered processes, intelligent information access, and social networking will be growth areas. The estimate for this segment is $2.7B.

- *Reasoning* — Semantic technologies enable intelligence, real-time auditing and compliance, simulation-based "virtual" product design, engineering and manufacturing, virtual data centers, adaptive logistics, and supply chain optimization. New application categories will have huge economic benefits. The estimate for this segment is $6.1B

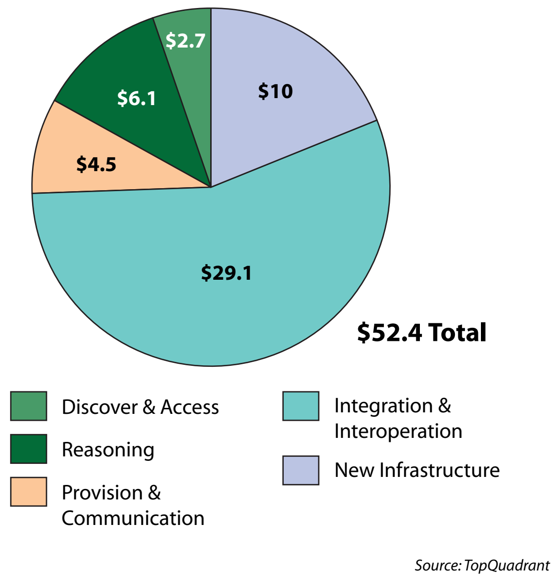


Figure 3-14: Semantic Technology Horizontal Markets to 2010 ($US Billions)

This estimates the size of horizontal mainstream markets for semantic technologies in 2010. This horizontal segmentation generally follows the earlier description of semantic technology functional categories. Semantic integration and interoperability represent the largest segment, followed by new infrastructure. Discovery and acess, reasoning, and semantic provisioning and communication are expected each to become multi-billion dollar segments.

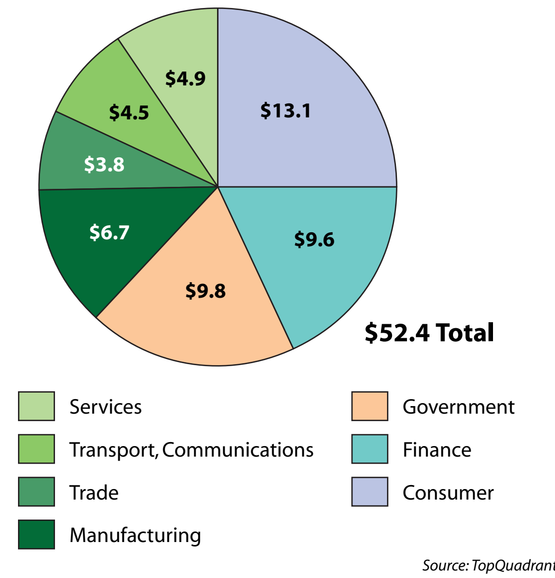


Figure 3-15 Semantic Technology Vertical Markets to 2010 ($US Billions)

This chart estimates the sizing of the top seven vertical mainstream markets for semantic technologies in 2010. Finance, government, and manufacturing represent the largest business segments. Consumer technology and services represent a significant segment for semantic technologies.

- *Provisioning & Communication* — Representing the knowledge about things separately from content and media files will spawn new categories of enterprise publishing, especially relating to product lifecycle management, professional publishing, and business information services. The estimate for this segment is $4.5B.

- *Integration & Interoperation* — By far, semantic integration will be the largest category of services and software during this decade. The estimate for this segment is $29.1B

- *Semantic infrastructure* — The emergence of semantic web services, context and situational computing, semantic grid, pervasive computing, and large-ontology reasoning engines will include new operating systems and hardware categories. The estimate for this segment is $10B.

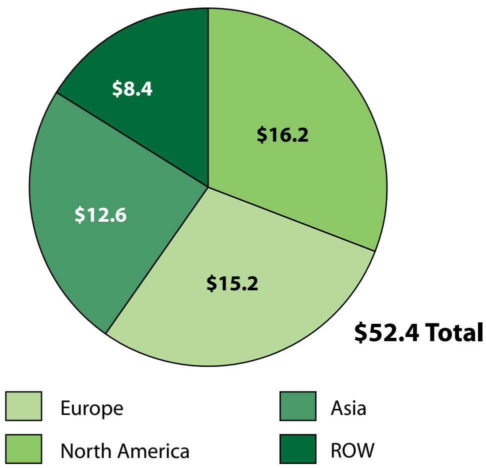
Semantic technology vertical markets

Figure 3-15 estimates vertical sizing of mainstream markets for semantic technologies by the top-7 industry segments in 2010. Industry segments are ordered by size of ICT investment in semantic technologies.

Semantic technology regional markets

Figure 3-16 estimates sizes of mainstream markets for semantic technology for Europe, North America, Asia-Pacific, and Rest of World regions.

This completes Section-3. We have seen that WonderWeb project is an integral part of a larger cycle of innovation we call the semantic wave. This wave encompasses investment in R&D, commercialization of technologies, and market adoption of new products and services. The semantic wave is building steadily. It is on a track that will take it from vision to between $40-60 billion markets by the end of this decade. The initiating force from the outset of this cycle has been public sector funding of fundamental R&D. However, the driving force for mass market adoption will be 2-10X gains in efficiency, effective-



Source: TopQuadrant

Figure 3-16: Semantic Technology Regional Markets to 2010 ($US Billions)

This charts estimates size of global mainstream markets for semantic technology in 2010 by region. Europe and North America represent the largest sectors. Asia-Pacific, and Rest of World will be the fastest growing regions.

ness and other measures of performance. We've seen that WonderWeb has played a seminal and influential role in this story by focusing on networked ontology infrastructure, which is at the core of many new lines of technology development. In the next section we examine specific WonderWeb technologies, their R&D influence, and market implications in more detail.

SECTION FOUR

WONDERWEB IN THE SEMANTIC WAVE

TECHNOLOGY AND MARKET ASSESSMENT

This section addresses the following key questions:

- What key contributions has WonderWeb made?
- How has the impact of WonderWeb results been disseminated?
- How do WonderWeb results compare with other world-class research efforts?
- How has WonderWeb influenced the world?
- What directions should future development consider taking?

BACKGROUND - ONTOLOGIES

Ontologies are at the heart of the semantic web. An ontology describes the types of information that a semantic web page can express. More importantly, it describes the relationships among these types. Ontologies were originally defined as a representation of a shared conceptualization . WonderWeb contributions, in particular the Ontology Inference Layer (OIL, a direct precursor to the W3C standard ontology language OWL) provided a concrete, formal representation of an ontology. By linking the notion of ontology with inference, OIL provided a computationally tractable method by which the value of an ontology could be delivered.

Inference mechanisms allow an ontology to specify how different types of data fit together, enabling mediation and integration capabilities. An ontology describes the network of meaningful inferences that can be made with information of certain types.

Most semantic web technologies can be characterized in terms of ontology management:

- Ontology language that represents knowledge
- Programming interfaces that allow programmatic access to ontologies
- Reference ontologies that provide ontology engineering guidance
- Tools to build and/or display ontologies
- Reasoning tools for ontologies
- Ontology servers that present ontologies on the web
- Ontology stores for persistent management of ontologies.

WonderWeb has made contributions to many of these areas.

WONDERWEB CONTRIBUTIONS

The WonderWeb project defines 29 deliverables, which we have organized into five technical contributions. Descriptions of these contributions, and comparisons to other related approaches are shown in Figure 4-1.

As Figure 4-1 shows, WonderWeb contributions span a wide range of semantic capabilities and functions. This table also shows that in some cases, WonderWeb contributions include activities (such as the OWL API) that have gone beyond activities described in the planned deliverables originally declared in the project.

The interpretation of the last column, "Other approaches", is necessarily somewhat loose, since in many cases there is no direct equivalent to the WonderWeb deliverable elsewhere in research or industry. In particular, the OWL API solves a somewhat different problem than the RDF APIs (JENA et al) listed here. The KAON server is noteworthy not for the novelty of the solution it provides (it is an integration framework), but for the fact that it is an open integration platform, designed to support modules coming from different sources. Comparable industrial platforms are, not surprisingly, focused on supporting proprietary sets of components.

IMPACT AND COMPARISON

The impact of the WonderWeb deliverables in comparison to other world-class research efforts is shown in Figure 4-2.

As shown in Figure 4-2, WonderWeb deliverables have had considerable impact on ontology tools and development in all areas of semantic web technology. The most direct impact has been the influence of the OWL standard, of which the Ontology Inference Layer (OIL) was a direct contributor. This standard has already been influential in many of the commercial tools mentioned in Figure 3-9.

WonderWeb Contribution	Deliverable	Product Category	Functions and Capabilities	Other Approaches
Oil	D1, D3, D4	Ontology Language	Represent ontologies, description logic. Reason and infer.	Topic Maps, RDFS, Ontology Works Language
OWL API		Programming interface	High-level programming interface designed for the specific needs of OWL-based applications; inferencing, versioning, parsing, serializing, and modeling.	JENA, SNOBASE, SESAME
OilED	D10	Tool to build/display ontologies	Discover, acquire and create metadata, Graphically edit description logic class structure, properties, logic statements, etc. Linkage to DL reasoner for auto classification and hierarchy construction.	Protégé, OntoEdit, Cerebra Construct, InferEd
DOLCE	D15, D16, D17, D18, D19, D20, D21, D22	Reference ontology	Modular set of reference (or so-called "upper") ontologies as a starting point and methodological guidance for ontology engineering.. Used to Discover, acquire and create metadata, and to reason and interpret.	Cyc, SUMO, SUO
KAON Server	D5, D6, D7, D8	Ontology server/store Reasoning tool	Provision, present, communicate and act. Provides an open infrastructure for combining different semantic web tools; parsers, reasoners, web deployment, query engines, etc.	Proprietary integrated systems: RDF Gateway, Cerebra, OntologyWorks, OntoEdit

Figure 4-1: Description of WonderWeb Deliverables

Contribution	Differentiators	Impacted systems/ Standards/ Projects	Comments
Oil	W3C recommendation. Oil includes sufficient logical power to represent constraints in a consistent and tractable way	W3C OWL Recommendation	OWL is the basis for a new generation of ontology-based tools Oil formed the basis of a consistent, logic-based approach to the semantic web.
OWL API	OWL API is tuned to the needs of a reasoning system. Other APIs are more like database/object store/query systems.	Used by Protégé to access DIG reasoners. SDK Ontology API is based on the OWL API Commercial vendors are largely unaware of the OWL API	Inference as the basis for ontology specification results in particular challenges and opportunities for programming interfaces. The OWL API is ahead of its time in addressing these issues.
OilED	Focus on modeling with OWL reasoning Linkage to DL reasoners. Other modeling systems were focused on frame logics or object systems.	Protégé OWL Plugin, Co-ode project	Non-trivial OWL models cannot be authored by members of most user classes. A graphical support tool for authoring OWL is necessary. The interface innovations made in the OilED tool have proven to be a usable step for making OWL accessible to a wider user class.
DOLCE	DOLCE has a modular structure that differentiates it from other "top-down" approaches DOLCE has been applied to web services (OWL-S) to disambiguate and formalize service descriptions.	Chosen by Mitre and OntologyWorks for data integration	The viability and usefulness of so-called "upper ontologies" is a controversial issue in Ontology Engineering. Dolce provides a mechanism for modularizing its reference ontologies that lowers the entry barrier to making a commitment to an upper ontology.
KAON Server	KAON server is an open framework for integrating tools from different sources. Other frameworks are proprietary, lock-in closed systems.	KAON framework is the web server for OntoWeb community. Has been used for United Nations FAO pilot projects	The KAON project includes a number of pieces (Text2Onto, the KAON portal, etc.) The KAON server is a framework that allows something of a plug-n-play environment for ontology services. Most commercial tools have a vested interest in having the system developer use their own tools, and provide their own integrated environment. The KAON environment has had most impact on research projects and prototypes.

Figure 4-2: Impact of Particular WonderWeb Deliverables

Another major influence of WonderWeb results has been through Stanford University's Protégé project and the AKT Co-ode project. Protégé has become the de-facto standard software for creating, visualizing, and editing ontologies. Even though Protégé is an open-source project, it is quickly achieving a commercial level of reliability. The Co-ode project is a collaborative project including some WonderWeb partners and Stanford University, with the aim of producing a "best of breed" ontology visualization and editing environment for OWL. One result of this collaboration is the new Protégé capability to handle OWL ontologies. The influence of OilED on this capability is clear to anyone who has used both tools.

DOLCE is the WonderWeb contribution to the field of reference ontologies. This is a controversial area, in that the usefulness of reference ontologies is not uniformly accepted by modeling practitioners. DOLCE is a relative newcomer to the field (Cyc, the grand dame of reference ontologies, dates back to 1985). Nevertheless, we are encouraged by the fact that the user list for DOLCE is quite long, and includes companies like OntologyWorks [Andersen-2004] who have demanding requirements for ontologies.

SUMMARY OF WONDERWEB PROJECT CONTRIBUTIONS

- Results from WonderWeb have had significant impact on other Research and Development projects in the EU, US and globally. For example, through Co-ode, OilED has had significant impact on the Protégé platform.

- WonderWeb technologies have formed the basis of and provided the impetus for significant commercialization. The OWL recommendation is already supported by many vendors, and others have announced future support for OWL.

- WonderWeb technologies have been embraced by early adopters. The United Nations and Boeing are prime examples.

- WonderWeb developments have exceeded original plan in important respects. Contributions of DOLCE to semantic web services, and the insights of the OWL API went beyond the original expectations of the WonderWeb project.

These contributions define the technology trends behind the semantic web vision today. As such, WonderWeb products have created much of the infrastructure for the modern semantic web.

WONDERWEB IMPACT – BROAD IMPLICATIONS

The adoption of OWL as a W3C Recommendation in Feb. 2004 was a milestone event, whose impact was felt well beyond the scope of W3C specifications. The idea of creating standards for modeling was not new; in fact, there is no shortage of modeling languages for various industries and applications.

Each of these modeling languages or approaches has identified a need to represent complex logical constraints between the elements that they model. Many of these approaches have even given names to planned language extensions to address this issue. However, efforts to gain agreement on constraint representation standards have faced difficulties, and few approaches have achieved success.

The OWL recommendation was ground-breaking in three ways:

1. It provided a means of representing constraint information,

2. It provided a rigorous, logical semantics for that representation, and

3. It was approved by a major, world-wide standards body.

All three of these features were instrumental in making OWL attractive to other modeling communities. Its ability to represent constraints made it a candidate solution for a long-standing and recognized problem. Its rigorous foundation made it possible to determine the precise semantics of combinations of OWL with other modeling formalisms, and its acceptance by a major standards body relieved other modeling consortia of the time-consuming and painful task of coming to a consensus in their own communities.

Table 4-3 shows various modeling communities, approaches or standards efforts and the degree to which they have embraced or approached OWL to satisfy their constraint modeling needs. Given that many of these efforts pre-date OWL ratification by many years, and these investigations are underway less than half a year afterwards, this is a very rapid uptake.

Modeling standard	Industry/focus	OWL activity
Topic Maps	Content management	One major vendor and Topic Map consortium founding member (Mondeca) has recommended the use of OWL in conjunction with Topic Maps to satisfy constraint representation needs.
UML/OCL	Software	Sources inside one premier vendor have expressed dissatisfaction with the precision of the semantics of UML (despite efforts to formalize the constraint language OCL). In its ODM effort, OMG is exploring use of OWL to remedy this issue.
STEP	Manufacturing	Key players in the commercialization and dissemination of STEP (EuroSTEP and NASA) are evaluating the use of OWL as a standard and rigorous way to model constraints.
BPEL4WS	Enterprise Architecture	A standards effort is underway to use OWL to satisfy the logical needs of process mapping.
OWL-S	Web Services	The emerging W3C Web Services effort officially embraced OWL-based PSL from NIST as a means for specifying semantic agent mappings.
DCML	Network Infrastructure	Cisco and others are pursuing the use of OWL to provide constraint modeling for the specification of data centers.
ebXML	eBusiness	The Repository Interface Manager (RIM) of ebXML is evaluating the use of OWL for specification and management of business services.

Figure 4-3: Modeling approaches embracing OWL for constraint representation.

WONDERWEB CONTRIBUTION — PUTTING THE "O" IN "OWL"

Ontologies have been studied in Artificial Intelligence since long before the beginning of the effort that culminated in the OWL standard. But these studies had not resulted in a clear, technologically actionable, notion of ontology. At that time, the best agreement on the definition of "ontology" was as "a specification of a conceptualization." This idea was too vague to bring unity to research into ontologies.

The Ontology Inference Layer work of WonderWeb took a stand on a rigorous definition of ontology, and related it to the idea of inferencing. The result was a more formal definition of ontology as a model in a logical framework, capable of specifying certain well-defined implications. It was this notion of an ontology defined by inference that was incorporated into the DAML+OIL language, which later became known as OWL.

All of the WonderWeb deliverables mentioned above contributed to demonstrating the viability of this approach. The abstract syntax and formal semantics of the OIL language provided the foundation for the rigor of the OWL language. OilED and the OWL API demonstrated that models in a logical form could be accessed by users and software. The KAON framework demonstrated that these models could serve as a basis for distributed (web) applications. DOLCE showed that this approach was capable of representing non-trivial and comprehensive knowledge models.

The value of the OWL standard, and its rapid uptake by other modeling communities, is the result of a laudable effort by the W3C Semantic Web working group, and is the result of synergy between the many efforts that contributed to it. No single contribution is responsible for this success. But the contributions of WonderWeb, through OIL, and as demonstrated by the other WonderWeb deliverables, were instrumental to providing OWL with the formal rigor that has made it so attractive to the modeling efforts listed in Figure 4-3.

CONCLUSIONS

WonderWeb and other projects have made great advances in semantic web technology, to the point where the technologies have been recognized as standards and a lively industry of support tools and services has grown up. We cannot say that the formal and developmental work is complete, but in order to facilitate the industry take off, it is time for

Challenge Area	Barrier	Research Issue
1. Mapping of legacy data to ontologies.	**System implementers need to map a spectrum of legacy datasources to ontologies.**	Research is required to understand how semantic mapping can be done in a way that respects the complex interaction between ontologies, legacy data, applications, work culture and business process.
2. Multi-modeling paradigm support.	**Modelers working in manufacturing (STEP), enterprise architecture (PSL), software engineering (UML and MDA) and workflow (BPM) have difficulty integrating OWL with their domain-specific modeling approaches.**	Research is required to determine the kinds of issues that arise when combining OWL with other modeling formalisms.
3. Logic and ontology.	**Modelers from backgrounds other than semantic web (domain modeling, object modeling, data modeling) often lack the logical sophistication to produce good OWL models.**	Research and educational development is needed to provide best practices advice based on the details of the various standard levels of OWL (OWL-Lite, OWL-DL and OWL-Full).
4. Design Patterns for Semantic Solutions.	**Domain practitioners are typically not sufficiently trained in logic to build effective knowledge models.**	Research is required to determine and develop domain- and industry-specific classes of problems and solution patterns that match them.
5. Ontology Engineering lifecycle management	**Ontology modelers find a need to perform global edits on their ontologies, merging some parts that were developed separately, dividing some models into smaller parts.**	Research is needed to determine what operations are most needed to support a realistic ontology development effort
6. Ontology reuse.	**Ontologies must be designed for re-use, but real-world attempts to re-use ontologies have resulted in disappointing results.**	The required research in this area encompasses both frameworks for modularizing and re-using ontologies, and constructing actual reusable ontology assets for specific domains and competencies.
7. Visualization.	**It is difficult for modelers to come to a sufficient understanding of an ontology without the ability to "see" aspects of the ontology.**	Research into visualization techniques that conform to various industry standard schematics could smooth adoption paths, especially for engineering domains.

Figure 4-4: Factors that Impact Technology Adoption

the research community to turn its attention to the scruffy issues of the real business world.

Addendum-C: Directions for Further Development outlines some critical, unresolved issues that have emerged in our experience as major barriers to adoption of semantic web technologies. In many cases, research efforts are already underway to address these and related issues. Figure 4-4 provides a concise summary of a selected set of such issues, that we chose to elaborate in Addendum-C in terms of *barriers to adoption, open research questions, indications of related research underway,* and *associated development issues to be handled by industry and vendors.* In all cases, we recommend that systematic technology transfer research be undertaken on these and related issues and challenges that are emerging to determine the appropriate approach(es) that can lead to industry best practices for vendors and solution providers. We argue that if such issues are not discovered and addressed in a well-framed stage of technology transition research, they will pose significant barriers to the adoption of semantic technologies in service of identified critical organizational and societal goals that depend on this technology.

Semantic web technologies show great promise, but they also face serious challenges to adoption. They represent a new paradigm of system development, so they will face challenges from developers. They attempt to formalize terminology that is already in use by practitioner communities; this means that they face challenges with respect to cultural accepted terms and concepts. Is it possible for the formal constructs of ontologies to capture the scruffy nuances that are typical in the everyday workplace? Any real work setting or business practice has embedded in it some form of contradiction, which is dealt with through "the human angle" or "touch" parts of the business. Is it possible for any formal system to be made flexible enough to continue to provide value, while respecting the ethnographic reality of the workplace? We conclude with a caution in service of a hope and an appeal to the next round of action:

> If questions such as those above are not answered (or if the answer to them is "no"), then the critics of the Semantic Web [Shirkey-2003] will turn out to be correct – the formal capabilities of the semantic web will remain an academic exercise with little applicability to the real world of people and their information.
>
> If, on the other hand, the answers to these questions are 'yes', then the resulting resolutions and support will lead the way to a thriving semantic web industry. It is the duty of semantic web researchers to find these answers.

ADDENDUM-A

BIBLIOGRAPHY

[1] Allemang, Dean, Polikoff, Irene, Hodgson, Ralph, Keller, Paul, Duley, Jason and Chang, Paul. (2004). COVE – Collaborative Ontology Visualization and Evolution, to appear IEEE Aerospace Conference 2005.

[2] Andersen, William. (2004) Application of Formal Ontology to Database Schema Alignment, Information Interpretation and Integration Conference I3CON 2004. http://www.atl.lmco.com/projects/ontology/i3con/ppt/Andersen.ppt

[3] Auriol, L., Wyckoff, A, (2004). *EU R&D Intensity Targets of 3%: Implications for the EU and the US*. Presentation given to the OECD Breakfast Series in partnership with NABE.

[4] Behrendt, W., Geser, G., Mulrenin, A., Reich, S., (2003). *EP2010—The Future of Electronic Publishing Towards 2010: Recommendations for Research*. Conducted by Salzburg Research Forschungsgesellschaft mbH on behalf of the European Commission, Information Society DG – E2, Knowledge Management and Content Creation. http://ep2010.salzburgresearch.at/

[5] Boch, W. (2003). *EU Grid Research and the European Research Area*. Presentation to UK/FR Grid Workshop 3-4 November 2003, London. DG Information Society

[6] Clark, P. (2004) *Knowledge-base/Ontology Projects Worldwide*. Web page with links to over 250 knowledge base and ontology projects worldwide. http://www.cs.utexas.edu/users/mfkb/related.html

[7] Cohen, R. (2004). *The Impact of Grid Computing: The Results of the ESI Grid Study*. Economic Strategy Institute. Presentation to the GT'04 VIP Summit. May 2004.

[8] Computer Science and Telecommunications Board, National Research Council. (2002). *Information Technology Research, Innovation, and E-Government* .National Academy Press, Washington, D.C.

[9] Dahlman, C., Utz, A. (2004). *India and the Knowledge Economy: Leveraging Strengths and Opportunities*. World Bank Institute, New Delhi, November 2004.

[10] Damjanovic, V., Devedzic, V., Djuric, D., Gasevic, D. (2004). *Framework for Analyzing Ontology Development Tools*. AIS SIGSEMIS Bulletin Vol. 1 No. 3, October 2004, pp 43-47.

[11] Davis, M., (2004) *The Business Value of Semantic Technology*. TopQuadrant, Inc. Washington, DC. September 2004. Special report to the Semantic Technologies for E-Government Conference.

[12] Davis, M., Walter, M. (2004) *The Next Wave: Revolutions in Process and Content*. The Seybold Report Vol. 3, No. 15 • November 17, 2003

[13] Ding, Y., Fensel, D., and Stork, H. (2003). *The Semantic Web: from Concept to Percept*. Austrian Artificial Intelligence Journal. 21(4). 2003.

[14] Denny, M. (2004). *Ontology Editor Survey 2004 Table*. Michael Denny <denn@suffolk.lib.ny.us> July, 2004. Reviews 94 ontology editing tools.

[15] Expert Group. (2004). *Next Generation Grids 2: Requirements and Options for European Grids Research 2005-2010 and Beyond.* NGG2 Final Report. European Commission. July 2004.

[16] Fensel, D. (2003). *Semantic Web - EU perspectives. Presentation to EU Information Society Technologies.* Business Informatics Group, Vrije Universiteit, Amsterdam.

[17] Giaglis, G. (2002) *Directions and trends in knowledge management research: Results from an empirical analysis of European projects.* In V. Hlupic, editor, Knowledge and Business Process Management. Idea Group Publishing, 2002.

[18] Geldof, M. (2004). *The Semantic Grid: Will Semantic Web and Grid Go Hand-in-Hand?,* European Commission DG Information Society, Unit "Grid Technologies", June 2004.

[19] Goldstein, H. (2002). *They Might Be Giants: Special R&D Report.* Top-100 Industry R&D Spenders. IEEE Spectrum. September 2002. http://www.spectrum.ieee.org/WEBONLY/publicfeature/sep02/intro.html

[20] Greaves, M. (2003). *DAML PI Meeting Kickoff.* DARPA Washington DC. October 2003.

[21] Harsha, P. (2004) *Computing Research in the Fy2005 Budget Request.* Computing Research Association, Washington DC 2004.

[22] Hendler, J. (2003). *Semantic Web: Models and Services in Support of e-Science.* Maryland Information and Network Dynamics Laboratory (MIND), Semantic Web Agents Project (SWAP). Presentation.

[23] Hendler, J. (2003). *Teraflops, Petabytes and Exalinks — Science and the Semantic Web.* A presentation to e-Science UK 2003. University of Maryland. http://owl.mindswap.org

[24] Interagency Working Group on Information Technology Research, (2003). *Networking and Information Technology Research and Development —Advanced Foundations for American Innovation: Supplement to the President's FY 2004 Budget,* Office of Science and Technology Policy, Washington DC,.

[25] Keller, Berrios, Carvalho, Hall, Rich, Sturken, Swanson & Wolfe, (2004) *Semantic Organizer: A Customizable Semantic Repository for Distributed NASA Project Teams.* ISWC 2004.

[26] Khuong, V. (2004) *ICT and Global Economic Growth: Contribution, Impact, and Policy Implications.* Doctoral Thesis. Harvard University. July 2004.

[27] Lallana, E. (2003). *Comparative Analysis of ICT Policies and e-Strategies in Asia.* Asia-Pacific Development Information Programme, United Nations Development Programme. October, 2003. Nine Asian countries are included in this study: India, Japan, Republic of Korea, Malaysia, Nepal, Philippines, Singapore, Sri Lanka and Vietnam.

[28] Meeker, M. (2004) *Internet in China — October 2004.* Presentation to the Web 2.0 Conference, San Francisco CA, October 2004.

[29] Meeker, M., Pitz, B. (2004) *Internet: An Update from the Digital World.* Morgan Stanley Global Equity Research, October 2004.

[30] Meeker, M., Pitz, B., Motoyama, Y., Martin, I. (2003) *Internet, Technology, Media & Telecom: Global Insights — Global Market Sizing of TMT Products and Services.* Morgan Stanley, Global Equity Research, September 2003.

[31] Michels, D. (2004) *U.S. Trends in Venture Capital Financing.* Kirkpatrick & Lockhart LLP, San Francisco, CA USA. Presentation to Autovisionale.

[32] Miller, H. (2004) *Digital Planet 2004: The Global Information Economy.* World Information Technology and Services Alliance (WITSA). October, 2004.

[33] Noy, N. F. & Musen, N. F. (2000). *PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment.* Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-2000), Austin, TX. Available as SMI technical report SMI-2000-0831 (2000)

[34] NSF (2003). *Revolutionizing Science and Engineering Through Cyberinfrastructure.* Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. January 2003.

[35] National Science Board (2004). *Science and Engineering Indicators 2004.* Volume 1 & 2. National Science Foundation, Washington DC, May 2004.

[36] Obozinski, V. (2004) *The Information Society Strategic Objective on Grid Technologies in EU RTD Framework Programmes.* Grid Services Engineering and Management International Conference 29-30 September 2004.

[37] OECD (2004). *OECD Information Technology Outlook 2004: Highlights.* Organization for Economic Cooperation and Development, Paris, France 2004.

[38] Ohlms, C (2002). *The Business Potential of Ontology-based Knowledge Management. McKinsey & Company.* Presentation to KnowTech2002.

[39] PricewaterhouseCoopers (2004). *Global Private Equity 2004 — Final Presentation.* November 2004.

[40] Pulvermacher , M., Stoutenburg, S., Semy, S. (2004) *Netcentric Semantic Linking: An Approach for Enterprise Semantic Interoperability.* Mitre Technical Report, Revison 1. Mitre Corporations, October 2004. Also, Toward the Use of an Upper Ontology for US Government and US Military Domains: An Evaluation, MITRE report, Sept. 2004.

[41] Rader, M. et al. (2004). *Key Factors Driving the Future Information Society in the European Research Area: Synthesis Report on the FISTERA Thematic Network Study (2002-2004).* Institute for Prospective Technological Studies. European Commission Joint Research Centre, 2004.

[42] Sheehan, J. (2004). *Globalization of R&D: An OECD Perspective.* Presentation to the GUIRR Meeting on Globalization, Competitiveness and Workforce, Washington, DC, 3-4 February 2004. OECD Science & Technology Policy Division.

[43] Shirkey, Clay. *The Semantic Web, Syllogism, and Worldview.* "Networks, Economics, and Culture" mailing list, 2003. http://www.shirky.com/writings/semantic_syllogism.html

[44] Stork, H. (2002) *Webs, Grids and Knowledge Spaces - Programmes, Projects and Prospects.* Journal of Universal Computer Science. Volume 8 (2002) Issue 9.

[45] SWAP Consortium (2003). *SWAP Market Survey.* Semantic Web and Peer-to-Peer Project Deliverable D8_3. May 2003.

[46] Teicht, A., Koizumi, K. (2004). *Introduction to R&D in the FY2005 Budget.* American Association for the Advancement of Science. Washington, DC, 2004.

[47] W3C Semantic Web Activity Statement (2004), http://www.w3.org/2001/sw/Activity

[48] Zimmerman, R. (2004). *2nd Joint Call IST/NMP.* Presentation at Information Day. Bilbao, June 2004. European Commission – DG Information Society. Call for proposals on nanotechnology, nanophotonics, biosensors, and manufacturing innovation. 02 March 2004. Directorate for Science, Technology and Industry, OECD.

ADDENDUM-B

DATA TABLES

ESTIMATING THE SIZE AND SHAPE OF THE SEMANTIC WAVE

This report breaks new ground by estimating the flows of capital for semantic technologies from government-funded R&D, venture capital investment, industry-funded technology commercialization, and market adoption for the period from 2000 to 2010.

Also, this report develops regional estimates for Europe, North America, and Asia as well as estimates for both horizontal and vertical markets for semantic technologies. The analysis is unique for drawing upon a wide variety of trend information to cross-correlate estimates.

These notes describes basic features of the methodology we followed to prepare data tables which were the source for charts and graphics presented in Section-3 of this report. About the two data tables:

The first estimates global investment in semantic technologies from 2000-2010. For this period, we present investment estimates for R&D, venture capital, and industry commercialization for North America, Europe and Asia.

The second table estimates semantic technology vertical markets from 2000 –2010. We estimate market growth to 2010 in several ways, by: information and communications technologies (ICT) versus semantic ones, and horizontal, vertical, and regional semantic technology markets.

We have qualified projections by examining companies and commercialization activity, technical capabilities they are selling, and experiences of early adopters pursing semantic web projects.

For example:

- To gauge the extent, and growth rate of commercialization, the market assessment examines current productization, trends in venture capital (VC) and investment banking, and relevant activities of the larger ICT players.

- To gauge the strength of economic driving forces for adoption of semantic technology, this assessment examines research about early adoption of semantic technologies in government and industry. Cited research gives evidence for the rate market growth that is likely in the second half of the decade.

NOTES REGARDING SOURCES FOR DATA TABLES

Capital flows for semantic technology

There are two broad categories: (1) investments to develop and commercialize semantic technologies; and (2) returns from sales, and deployment of semantic technologies for use.

Semantic technology investment

Investment in semantic technology comes from three sources: (1) government R&D for fundamental and applied research; (2) venture capital investment; and (3) ICT industry commercialization of semantic technology products and services.

Semantic technology returns

Returns from semantic technologies come from two sources: (1) revenues from sale of semantic technology products and services; and (2) returns on investment (ROI) from deployment of semantic technology solutions, over and above acquisition costs and total costs of ownership (TCO).

Aggregate investment in semantic technology
For any given period of time, the capital invested towards semantic technologies is a small subset of funds going for Information and Communication Technology (ICT), especially Internet-related infrastructure, products and services. In turn, ICT investments are a subset of all R&D and capital investment. R&D investment, in turn, is a subset of the Gross Domestic Product (GDP) of the country or region.

Semantic technology markets
For any given period of time, the monies expended for semantic technology infrastructure, products and services is a subset of the spend for ICT technologies and services, which, in turn, is a subset of the Gross Domestic Product (GDP).

Sources we consulted about semantic technologies, research trends, and regional market trends include: [Meeker-2003, Meeker-2004], [Roder-2004], [Davis-2004, Davis-2003], [Behrendt-2003], [Obozinski-2004] , [Dohlman-2004], [Lallana-2003].

Government R&D
Public sector funding for fundamental and applied research is most important during the early stages of a cycle of innovation, and less important as market forces for commercialization and market adoption take hold. Investments in fundamental research have long maturation cycles and are high risk.

We examined R&D in different regions, levels of R&D spending, R&D themes and programs, R&D spending projections, and other trends.

Sources we consulted about government R&D trends include: the World Bank [Dohlman-2004, Lallana-2004], the National Science Board [NSB-2004], and OECD [Auriol-2004, Sheehan-2004], Computer Science Technology Board [CSTB-2002], Association for the Advancement of Science [Treicht-2004], and [Harsha-2004].

Sources we consulted about government R&D for semantic technologies include: EU researchers [Giaglis-2002, Stork-2002, Fensel-2003, Ding-2003], US researchers [NSF-2003, Interagency-2003].

Venture capital
Investment capital for new companies to launch products and services plays a key role in the commercialization cycle of new technologies. The time horizon for venture capital is shorter than for government R&D, typically measured in 3-5 years to reach a liquidity event.

Sources we consulted about private equity markets and venture capital investment include: OECD, [NSB-2004], [Michels-2004], [Pricewaterhouse-2004].

Sources we consulted about venture capital investment in semantic technologies include [Pricewaterhouse-2004], including survey databases, and contacts with numerous, [Meeker-2004] .

ICT R&D for new products and services
Investments by established companies to develop new products and services tend to have a near-term focus. Market timing is a key consideration.

Sources we consulted about commercialization of products and services include TopQuadrant proprietary databases, [Davis-2004].

Early adopter markets
Monies spent to acquire semantic technology solutions by enterprises seeking first-mover strategic advantages for their business.

Sources we consulted about levels of activity, costs, and business benefits of semantic technology solutions include [Clark-2004], [Davis-2004], [Davis-2003],

Mainstream markets
Monies spent to acquire proven semantic solutions by enterprises seeking practical benefits in well-understood areas of their business.

Sources we consulted about global ICT markets and the impact of new technologies include: [Miller-2004], [OECD-2004].

Sources we consulted about horizontal and vertical ICT market projections for ERP, CRM, BI, PLM, help desks, portals, web services, grid computing, enterprise integration include: IDC, MetaGroup, WITSA, Gartner, and Ovum.

GLOBAL INVESTMENT IN SEMANTIC TECHNOLOGIES — 2000 TO 2010

NORTH AMERICA

United States, Canada	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Cumulative
GDP	$10,627,644.0	$10,918,695.0	$11,305,162.0	$11,754,446.0	$12,107,079.4	$12,470,291.8	$12,844,400.5	$13,229,732.5	$13,626,624.5	$14,035,423.2	$14,456,485.9	$137,375,984.9
All R&D (around 2.8% of GDP)	373,781.9	313,424.8	294,619.5	303,865.6	361,429.2	379,625.0	395,375.7	410,637.9	427,852.0	440,958.9	459,932.4	$4,161,503.0
All Semantic R&D	138.9	159.8	175.9	212.9	352.3	457.4	561.1	754.5	960.5	1,087.1	1,243.4	$6,103.8
Government-funded R&D	66,000.0	72,000.0	78,000.0	80,340.0	126,507.0	131,960.0	134,231.0	135,230.0	137,347.0	137,388.0	139,723.6	$1,238,726.6
ICT R&D (@33%)	19,800.0	21,600.0	23,400.0	24,102.0	37,952.1	39,588.0	40,269.3	40,569.0	41,204.1	41,216.4	41,917.1	$371,618.0
Semantic R&D	39.6	64.8	93.6	120.5	189.8	237.5	241.6	324.6	329.6	350.3	377.3	$2,369.2
Semantic % of ICT R&D	0.2%	0.3%	0.4%	0.5%	0.5%	0.6%	0.6%	0.8%	0.8%	0.9%	0.9%	
Venture funded R&D	107,781.9	42,919.8	21,619.5	18,775.6	19,934.7	21,928.1	24,120.9	26,533.0	29,186.3	29,186.3	32,105.0	$374,091.3
ICT R&D (@55%)	59,280.1	23,605.9	11,890.7	10,326.6	10,964.1	12,060.5	13,266.5	14,593.2	16,052.5	16,052.5	17,657.7	$205,750.2
Semantic R&D	59.3	35.4	23.8	31.0	65.8	84.4	106.1	131.3	160.5	160.5	88.3	$946.5
Semantic %ICT	0.1%	0.2%	0.2%	0.3%	0.6%	0.7%	0.8%	0.9%	1.0%	1.0%	0.5%	
Industry funded R&D	200,000.0	198,505.0	195,000.0	204,750.0	214,987.5	225,736.9	237,023.7	248,874.9	261,318.6	274,384.6	288,103.8	$2,548,685.0
ICT R&D (@33%)	40,000.0	59,551.5	58,500.0	61,425.0	64,496.3	67,721.1	71,107.1	74,662.5	78,395.6	82,315.4	86,431.1	$744,605.5
Semantic R&D	40.0	59.6	58.5	61.4	96.7	135.4	213.3	298.6	470.4	576.2	777.9	$2,788.1
Semantic as %ICT (0.05–2.0%)	0.1%	0.1%	0.1%	0.1%	0.2%	0.2%	0.3%	0.4%	0.6%	0.7%	0.9%	

EUROPE

EU-25	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Cumulative
GDP	$9,283,921.3	$9,674,091.1	$9,876,541.7	$10,202,150.1	$10,508,214.6	$10,823,461.1	$11,148,164.9	$11,482,609.8	$11,827,088.1	$12,181,900.8	$12,547,357.8	$119,555,501.3
All R&D (from 1.9 to 3.0 of GDP)	$176,394.5	$188,644.8	$187,654.3	$198,941.9	$210,164.3	$227,292.7	$245,259.6	$287,065.2	$319,331.4	$353,275.1	$376,420.7	$2,770,444.6
All Semantic R&D	$144.4	$141.4	$134.3	$142.2	$197.1	$276.5	$362.7	$508.7	$673.8	$865.9	$1,037.8	$4,484.8
Government-funded R&D	$59,974.1	$64,139.2	$63,802.5	$67,640.3	$71,455.9	$77,279.5	$83,388.3	$97,602.2	$108,572.7	$120,113.5	$127,983.0	$941,951.2
ICT R&D (@33%)	$17,992.2	$19,241.8	$19,140.7	$20,292.1	$21,436.8	$23,183.9	$25,016.5	$29,280.7	$32,571.8	$36,034.1	$38,394.9	$282,585.3
Semantic R&D	$90.0	$96.2	$95.7	$101.5	$128.6	$162.3	$175.1	$205.0	$260.6	$288.3	$307.2	$1,910.3
Semantic as %ICT (0.05–2.0%)	0.5%	0.5%	0.5%	0.5%	0.6%	0.7%	0.7%	0.7%	0.8%	0.8%	0.8%	
Venture funded R&D (@5-7% R&D)	$13,548.1	$11,290.1	$10,651.0	$11,140.7	$11,769.2	$12,728.4	$17,168.2	$20,094.6	$22,353.2	$24,729.3	$26,349.5	$181,822.1
VC ICT R&D (@50%)	$6,774.0	$5,645.0	$5,325.5	$5,570.4	$5,884.6	$6,364.2	$8,584.1	$10,047.3	$11,176.6	$12,364.6	$13,174.7	$90,911.1
Semantic R&D	$33.9	$22.6	$16.0	$16.7	$17.7	$31.8	$42.9	$100.5	$111.8	$160.7	$197.6	$752.1
Semantic as %ICT (0.05–2.0%)	0.5%	0.4%	0.3%	0.3%	0.3%	0.5%	0.5%	1.0%	1.0%	1.3%	1.5%	
Industry funded R&D	$102,872.3	$113,215.5	$113,200.8	$120,160.9	$126,939.2	$137,284.8	$144,703.2	$169,368.5	$188,405.5	$208,432.3	$222,088.2	$1,646,671.3
ICT R&D (@20%)	$20,574.5	$22,643.1	$22,640.2	$24,032.2	$25,387.8	$27,457.0	$28,940.6	$33,873.7	$37,681.1	$41,686.5	$44,417.6	$329,334.3
Semantic R&D	$20.6	$22.6	$22.6	$24.0	$50.8	$82.4	$144.7	$203.2	$301.4	$416.9	$533.0	$1,822.3
Semantic as %ICT (0.05–2.0%)	0.1%	0.1%	0.1%	0.1%	0.2%	0.3%	0.5%	0.6%	0.8%	1.0%	1.2%	

ASIA

China, Japan, India, S. Korea	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Cumulative
GDP	$10,621,110.4	$11,299,053.6	$12,149,520.0	$13,064,000.0	$13,978,480.0	$14,956,973.6	$16,003,961.8	$17,124,239.1	$18,322,935.8	$19,605,541.3	$20,977,929.2	$168,103,744.7
All R&D (from 1.7 to 2.0 of GDP)	$180,558.9	$192,083.9	$206,541.8	$222,088.0	$237,634.2	$269,225.5	$288,071.3	$308,236.3	$348,135.8	$372,505.3	$419,558.6	$3,044,639.6
All Semantic R&D	$57.8	$56.4	$73.3	$78.8	$156.9	$247.8	$345.7	$494.4	$659.0	$870.5	$1,077.6	$4,118.3
Government-funded R&D	$50,556.5	$53,783.5	$57,831.7	$62,184.6	$66,537.6	$75,383.1	$80,660.0	$86,306.2	$97,478.0	$104,301.5	$117,476.4	$852,499.1
ICT R&D (@33%)	$15,166.9	$16,135.0	$17,349.5	$18,655.4	$19,961.3	$22,614.9	$24,198.0	$25,891.8	$29,243.4	$31,290.4	$35,242.9	$255,749.7
Semantic R&D	$15.2	$16.1	$34.7	$37.3	$79.8	$113.1	$121.0	$181.2	$204.7	$250.3	$281.9	$1,335.4
Semantic as %ICT (0.01–0.8%)	0.1%	0.1%	0.2%	0.2%	0.4%	0.5%	0.5%	0.7%	0.7%	0.8%	0.8%	
Venture funded R&D (@3-5% R&D)	$7,232.2	$7,021.5	$6,817.0	$7,328.9	$7,841.9	$9,692.1	$11,522.9	$12,329.5	$15,666.1	$18,625.3	$13,845.4	$117,922.7
VC ICT R&D (@50%)	$3,616.1	$3,510.8	$3,408.5	$3,664.5	$3,921.0	$4,846.1	$5,761.4	$6,164.7	$7,833.1	$9,312.6	$6,922.7	$58,961.4
Semantic R&D	$18.1	$14.0	$10.2	$11.0	$11.8	$24.2	$28.8	$61.6	$78.3	$121.1	$103.8	$483.0
Semantic as %ICT (0.05–2.0%)	0.5%	0.4%	0.3%	0.3%	0.3%	0.5%	0.5%	1.0%	1.0%	1.3%	1.5%	
Industry funded R&D	$122,770.2	$131,278.9	$141,893.1	$152,574.5	$163,254.7	$184,150.3	$195,888.5	$209,600.7	$234,991.7	$249,578.5	$288,236.7	$2,074,217.8
ICT R&D (@20%)	$24,554.0	$26,255.8	$28,378.6	$30,514.9	$32,650.9	$36,830.1	$39,177.7	$41,920.1	$46,998.3	$49,915.7	$57,647.3	$414,843.6
Semantic R&D	$24.6	$26.3	$28.4	$30.5	$65.3	$110.5	$195.9	$251.5	$376.0	$499.2	$691.8	$2,299.8
Semantic as %ICT (0.05–2.0%)	0.1%	0.1%	0.1%	0.1%	0.2%	0.3%	0.5%	0.6%	0.8%	1.0%	1.2%	

TOTAL

All regions	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Cumulative
GDP	$30,532,675.7	$31,891,839.7	$33,331,223.7	$35,020,596.1	$36,593,774.0	$38,250,726.4	$39,996,527.2	$41,836,581.5	$43,776,648.5	$45,822,865.3	$47,981,773.0	$425,035,231.0
All R&D	$730,735.3	$694,153.5	$688,815.6	$724,895.5	$809,227.6	$876,143.2	$928,706.6	$1,005,939.5	$1,095,319.2	$1,166,739.3	$1,255,911.7	$9,976,587.1
All Semantic R&D	$341.1	$357.6	$383.5	$433.9	$706.2	$981.7	$1,269.5	$1,757.6	$2,293.3	$2,823.5	$3,358.8	$14,706.8
Government-funded R&D	$176,530.6	$189,922.7	$199,634.2	$210,164.9	$264,500.4	$284,622.7	$298,279.2	$319,138.3	$343,397.7	$361,803.0	$385,183.0	$3,033,176.8
ICT R&D	$52,959.2	$56,976.8	$59,890.3	$63,049.5	$79,350.1	$85,386.8	$89,483.8	$95,741.5	$103,019.3	$108,540.9	$115,554.9	$909,953.1
Semantic R&D	$144.7	$177.1	$224.0	$259.3	$398.2	$512.9	$537.7	$710.8	$794.9	$888.9	$966.4	$5,615.0
Venture funded R&D	$128,562.1	$61,231.4	$39,087.5	$37,245.3	$39,545.8	$44,348.6	$52,812.0	$58,957.1	$67,205.7	$72,540.9	$72,299.9	$673,836.2
VC ICT R&D	$69,670.2	$32,761.7	$20,624.7	$19,561.4	$20,769.6	$23,270.7	$27,612.0	$30,805.2	$35,062.1	$37,729.8	$37,755.2	$355,622.7
Semantic R&D	$111.2	$72.0	$50.0	$58.7	$95.2	$140.5	$177.9	$293.5	$350.6	$442.3	$389.8	$2,181.6
Industry funded R&D	$425,642.5	$442,999.4	$450,094.0	$477,485.4	$505,181.4	$547,171.9	$577,615.4	$627,844.1	$684,715.8	$732,395.4	$798,428.8	$6,269,574.1
ICT R&D	$85,128.5	$108,450.4	$109,518.8	$115,972.1	$122,535.0	$132,008.1	$139,225.5	$150,456.3	$163,075.0	$173,917.5	$188,496.1	$1,488,783.3
Semantic R&D	$85.1	$108.5	$109.5	$116.0	$212.8	$328.3	$553.9	$753.4	$1,147.8	$1,492.2	$2,002.7	$6,910.2

Source: TopQuadrant, Inc.

1. Statistics on global investment and R&D (including government R&D, venture capital, and business) are based on publications from OECD, AAAS, NSF, WITSA, CIA, World Bank, PriceWaterhouseCooper, and MorganStanley.

2. Statistics on ICT markets are based on studies by WITSA, IDC, Ovum, Gartner Group, Morgan Stanley, McKinsey & Company. These examined: knowledge management access, infrastructure, and ontology-based solutions; enterprise & commercial publishing; enterprise-class applications such as ERP, CRM, BI, PLM, help desks, and portals; web services, grid computing, and pervasive computing (e.g. RFID); and the largest category, integration software and services.

3. Estimates of the amount of R&D for semantic technologies are based on proprietary research by region into the technology coverage of funded government research, the technology coverage of venture capital deals, and the commercialization activities of ICT vendors with interests in semantic technologies.

4. Estimates of mainstream markets for semantic technology is based on proprietary research into overall ICT adoption rates, horizontal and vertical adoption of semantic technologies, and regional rates of adoption. ICT-industy studies that provided the base line and rationale for end-points of the projection. They also gave insight into different horizontal application areas that would likely "go semantic" to some extent — for example the "enterprise integration" market.

SEMANTIC TECHNOLOGY VERTICAL MARKETS TO 2010 ($US Billions)

ICT Market Segment	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Cumulative
Consumer												
ICT Spend	$462.184	$496.972	$534.378	$574.600	$622.866	$675.187	$731.903	$793.383	$860.027	$932.269	$1,010.580	$7,694.348
SemTech	$0.005	$0.005	$0.005	$0.006	$0.031	$0.068	$0.366	$0.793	$2.580	$8.390	$13.138	$25.387
% semtech	0.001%	0.001%	0.001%	0.001%	0.005%	0.010%	0.050%	0.100%	0.300%	0.900%	1.300%	
Finance, Business												
ICT Spend	$368.127	$379.904	$392.057	$404.600	$436.159	$470.179	$506.853	$546.388	$589.006	$634.948	$684.474	$5,412.695
SemTech	$0.004	$0.004	$0.004	$0.081	$0.131	$0.235	$0.507	$2.732	$4.712	$6.984	$9.583	$24.976
% semtech	0.001%	0.001%	0.001%	0.020%	0.030%	0.050%	0.100%	0.500%	0.800%	1.100%	1.400%	
Government												
ICT Spend	$331.712	$352.510	$374.612	$398.100	$431.540	$467.790	$507.084	$549.679	$595.852	$645.904	$700.160	$5,354.943
SemTech	$0.003	$0.004	$0.075	$0.119	$0.259	$0.468	$1.521	$2.748	$4.767	$7.105	$9.802	$26.872
% semtech	0.001%	0.001%	0.020%	0.030%	0.060%	0.100%	0.300%	0.500%	0.800%	1.100%	1.400%	
Manufacturing												
ICT Spend	$307.460	$311.195	$314.974	$318.800	$341.435	$365.677	$391.640	$419.446	$449.227	$481.122	$515.282	$4,216.257
SemTech	$0.003	$0.003	$0.003	$0.080	$0.171	$0.366	$0.783	$2.097	$3.145	$4.811	$6.699	$18.160
% semtech	0.001%	0.001%	0.001%	0.025%	0.050%	0.100%	0.200%	0.500%	0.700%	1.000%	1.300%	
Trade												
ICT Spend	$248.761	$251.020	$253.300	$255.600	$270.169	$285.569	$301.846	$319.052	$337.237	$356.460	$376.778	$3,255.792
SemTech	$0.002	$0.003	$0.003	$0.026	$0.081	$0.003	$0.003	$1.276	$2.698	$3.208	$3.768	$11.070
% semtech	0.001%	0.001%	0.001%	0.010%	0.030%	0.001%	0.001%	0.400%	0.800%	0.900%	1.000%	
Transport, Comm.												
ICT Spend	$165.302	$175.108	$185.496	$196.500	$212.810	$230.473	$249.602	$270.319	$292.755	$317.054	$343.370	$2,638.788
SemTech	$0.002	$0.002	$0.002	$0.020	$0.043	$0.115	$0.250	$0.811	$2.049	$3.171	$4.464	$10.927
% semtech	0.001%	0.001%	0.001%	0.010%	0.020%	0.050%	0.100%	0.300%	0.700%	1.000%	1.300%	
Services												
ICT Spend	$138.480	$145.462	$152.796	$160.500	$179.279	$200.254	$223.684	$249.855	$279.088	$311.741	$348.215	$2,389.352
SemTech	$0.001	$0.001	$0.002	$0.032	$0.054	$0.100	$0.224	$1.249	$2.233	$3.429	$4.875	$12.200
% semtech	0.001%	0.001%	0.001%	0.020%	0.030%	0.050%	0.100%	0.500%	0.800%	1.100%	1.400%	
Total ICT Spend	$2,022.025	$2,112.191	$2,207.707	$2,309.064	$2,495.029	$2,696.486	$2,916.274	$3,159.856	$3,425.425	$3,716.668	$4,031.277	$31,092.003
Total SemTech	$0.020	$0.021	$0.093	$0.363	$0.769	$1.354	$3.654	$11.707	$22.183	$37.099	$52.328	$129.592

Source: TopQuadrant, Inc.

Statistics on ICT markets are based on studies by WITSA, IDC, Ovum, Gartner Group, Morgan Stanley, McKinsey & Company. These examined: knowledge management access, infrastructure, and ontology-based solutions; enterprise & commercial publishing; enterprise-class applications such as ERP, CRM, BI, PLM, help desks, and portals; web services, grid computing, and pervasive computing (e.g. RFID); and the largest category, integration software and services.

Estimates of mainstream markets for semantic technology is based on proprietary resarch into overall ICT adoption rates, horizontal and vertical adoption of semantic technologies, and regional rates of adoption. ICT-industy studies that provided the base line and rationale for end-points of the projection. They also gave insight into different horizontal application areas that would likely "go semantic" to some extent -- for example the "enterprise integration" market.

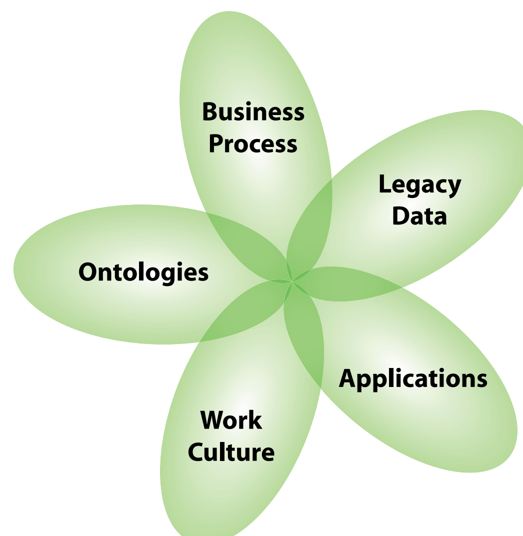
ADDENDUM-C

DIRECTIONS FOR FURTHER DEVELOPMENT

The research in the WonderWeb project to date has been largely infrastructural, with development of languages, standards, frameworks, interfaces and tools to deliver semantic capabilities to enable the adoption of semantic technologies. While we will not claim that the infrastructural work is complete, our own experiences in helping organizations adopt semantic technologies [Allemang-2004], along with experiences reported by others [Mitre-2004, Keller-2004] indicate that in addition to technical challenges, there are serious social, organizational, and business process issues involved in the successful adoption of semantic technologies.

In this section, we will outline our best interpretation of the major outstanding barriers between the current state of the art of semantic web technologies and successful use and adoption. Each point in this section is based on an observed need of some customer group we have encountered in our own practice or reported elsewhere. We have filtered each issue based on its suitability as an item on a future research agenda. It is not our place or intention to propose solution approaches, but rather to raise the issues we have observed.

It is important to keep in mind that technology adoption in any case is a multi-faceted problem, touching on business processes, work practices/culture and supporting applications. In the case of semantic technologies, the role of legacy data sources is particularly important to the story. Any solution that has a chance of being adopted will need to address organizational change issues in all of these various dimensions. As Figure C-1 suggests, all of these factors are inter-related; an effective technology application and adoption strategy cannot treat any of them in isolation.



Source: TopQuadrant

Figure C-1: Inter-related aspects impacting technology adoption.

TECHNOLOGY TRANSITION RESEARCH

In general, we recommend that the research that needs to be done at this point in the development of semantic web technologies is in the area of *technology transition research,* that grey area between academic research and commercial/government development. The current state of adoption of semantic web technology is that the change agents and early adopters in a wide range of organizations (spanning government and private sector, including industries like finance, health care, insurance, national security, energy and aerospace) have identified the need for some of the capabilities and supported changes that semantic web technologies promise to deliver. These early adopters are currently working with semantic web technology vendors, solution providers, or on their own, to craft pilot projects and demonstrators that will show the value of this new approach to the decision makers in their organizations. The result of this work is a selection of experience reports and solution stories, some more successful than others, of the applicability and application of semantic technologies.

But neither the customer organizations, nor the technology vendors, nor the solution providers are in a position to provide what the field really needs at this stage: systematic evaluation of the applicability of current semantic technologies and standards to solving the real business problems that are being faced by organizations today. In our own review of deployed semantic solutions, we have found that almost all deployed solutions finesse some semantic aspect of their solution in favor of some conventional work-around. This is not to criticize these efforts; vendors and solution providers have, quite appropriately, a first responsibility to solve a problem, not to explore the possibilities or to assume the unknown risks entailed in the application of an emerging (and likely disruptive) technology. Conventional solutions resolve many deployment issues by virtue of the maturity of the tools that support them.

What role does that leave for the research community? In contrast to solution providers and end customers, researchers have a primary responsibility to undertake the systematic exploration of a phenomenon. The phenomenon that now needs to be studied is the applicability of semantic technologies to real-world problems. This is not a contradiction. Systematic research utilizing real-world legacy data, business processes, work culture and legacy applications can be carried out by performing longitudinal studies, assembling user groups, making controlled comparisons and constructing demonstrator projects that give priority to technology evaluation, exploration of concerns and unearthing of potentially hidden risks and unresolved complexities rather than solution delivery per se.

There are unfortunately few examples of work of this sort. One example was performed in 2004 at Mitre [Mitre-2004]. The research outlines a detailed case study of a demonstration project in which real legacy data and workflow was used in an exploration of various options for using a description logic reasoner to provide semantic mediation between the data sources. The goals of the study were to examine the viability of the semantic approach, using OWL-DL in particular, to semantic data integration. The findings provided invaluable feedback to the vendor whose tool was used in the study, and the final report provides deep insights into the process of constructing a semantic integration solution for anyone who wants to attempt a similar project. A research program that produces systematic studies of this sort will provide the guidance that the semantic web field needs at this time to move from the laboratory into real practice.

There is an understandable temptation to subscribe to the "just do it" approach to technology transfer – the belief that the only way for a technology to transfer into practice is by providing solutions. While this is true to some extent, it would be a mistake to judge the success of research efforts on their direct ability to result in a business solution. The role of technology transition research is not simply to be a vanguard solution provider, but to act as an essential bridge to real solutions and adoption with the key recognition that technological bottlenecks are routinely underestimated. In this role, the next stage of research should more systematically study the options that a new technology affords, surface and resolve remaining risks, uncertainties and complexities that may be 'showstoppers', and derive insights into the strengths and weaknesses of different approaches. This technology exploration-centric approach is a luxury that solution providers and vendors can scarcely afford; but it is an essential function for successful technology transfer.

THE ISSUES

In this section, we outline some critical, unresolved issues that have emerged in our experience as major barriers to adoption of semantic web technologies. In many cases, research efforts are already underway to address these issues. In all cases, we recommend systematic technology transfer research as described above to determine the appropriate approach(es) that can lead to industry best practices for vendors and solution providers. No order of importance or priority is intended by the ordering of the issues. Further, this brief list of issues is not by any means provided as an exhaustive list but to highlight the higher level concerns outlined in the introduction to this section. In this sense, it only begins to outline numerous remaining crucial research areas and topics that we believe will emerge and can be formulated. They will reflect the reality that semantic technology adoption, in addition to impacting a number of closely interrelated aspects as depicted in Figure C-1, still faces any number of hidden or subtle technical challenges that result from the interaction of operations across the lifecyle of development and use, or from the cross-product of semantic technologies with other capabilities required for industrial strength solutions (e.g., grid computing). If such issues are not discovered and addressed in a well-framed stage of technology transition research, they will pose significant barriers to the adoption of semantic technologies in service of identified mission critical organizational and societal goals that depend on this technology.

For each issue we identify, we will outline the following points:

Barriers faced by real-world adopters of semantic web technologies, described with examples when possible.

Research questions that need to be resolved to overcome the barrier (these are recommendations for the future of WonderWeb.)

Indications of research efforts that might already be addressing some of these issues.

Development issues needed to address this barrier (these are points that should be handled by industry.)

1. Mapping of legacy data to ontologies.	
Barriers	In semantic integration applications, system implementers need to map legacy data to ontologies. The mapping needs in this situation run the gamut from schema mapping to recovery of ontology information from taxonomies, thesauruses and controlled vocabularies that are already encoded in various forms (databases, spreadsheets, web pages, Word outlines, bookmark files, etc.).
Research Questions	Research is required to understand how semantic mapping can be done in a way that respects the complex interaction between ontologies, legacy data, applications, work culture and business process. For instance, in most business processes, it is absolutely impossible to make a change to legacy data structures, even if they are logically inconsistent (the inconsistencies being dealt with by some accompanying business process). The existing infrastructure provides the raw materials for semantic integration, but at present the process is fragile, and there are actually too many options for a solution. The Mitre case study mentioned above relates a project in which a set of legacy data sources were mapped together through a Description Logic (DL) reasoner. The results revealed a tight coupling between the ontology design and the database design, and a fragility of the system. A small problem in the deployment of the DL reasoner (later repaired by the vendor) resulted in an unacceptable dependency between the ontology and the data schema. This fragility is unacceptable for industrial-scale applications of semantic technology. It is glib to blame the vendor in this case study for having delivered a "faulty" product; the field at large needs to have a comprehensive understanding of how the details of its tools (in this case the reasoner, but also mapping tools, modeling tools, display tools, languages, interfaces, frameworks are subject to this issue) impact the deployability of showcase solutions. In a mature technology, the solution patterns are well-known and (relatively) easily repeated by a trained practitioner. Semantic integration is still at the stage of tools, not solutions.
Research Efforts	The Mitre study is the only research effort we know of addressing these issues. At present, these issues are being worked out by customers of vendors of semantic integration tools, without systematic study or record of the results. This kind of work cannot be carried out by vendors or even consultants, whose charter it is to deliver solutions. They must be carried out within a research agenda, where solution options can be systematically explored.
Development Issues	Vendors must provide tools that can integrate smoothly with legacy data sources.

2. Multi-modeling paradigm support.	
Barriers	Modelers working in manufacturing (STEP), enterprise architecture (PSL), software engineering (UML and MDA) and workflow (BPM) are already committed to modeling systems that have shown some degree of success. The practitioners in these fields need a way to represent constraints in their models. Many of them recognize that OWL provides the representational capabilities they require. Yet there is still difficulty integrating OWL with their domain-specific modeling approaches.
Research Questions	Research is required to determine the kinds of issues that arise when combining OWL with other modeling formalisms. Just as is the case in semantic integration, there is already legacy commitment to these other modeling systems. In what ways can these be incompatible with the semantics of OWL, and what can be done to mediate those incompatibilities? Also, how can we quantify the value that OWL brings to these efforts, to motivate funding and adoption support?
Research Efforts	OWL/S is a W3C effort supporting process formalism. Other efforts are currently being carried out by semantic web enthusiasts in each of these communities.
Development Issues	Vendors will need to integrate OWL-native capabilities into toolkits that support these standards

3. Logic and ontology.	
Barriers	Modelers from backgrounds other than semantic web (domain modeling, object modeling, data modeling) often lack the logical sophistication to produce good OWL models. Commercial and academic training and guidance efforts to date typically aim at general modeling principles, independent of the particulars of any modeling formalism. It is easy to see why this would be the case; in advance of a standard for modeling, there is lower risk to developing generic training. Unfortunately, this leaves modelers at a loss for guidance that exploits the particular features of the OWL standards.
Research Questions	Research and educational development is needed to provide best practices advice based on the details of the various standard levels of OWL (OWL-Lite, OWL-DL and OWL-Full). Detailed best practices at this level are expensive and risky. However, with the ratification of standards in the past year, the risk level of this kind of research is now low enough to make this a good research prospect. For more basic research, on the horizon are other frameworks such as Frame Logics, Topic Maps, constraints/preferences (CP-Net) ideas, functional programming, and domain specific languages) which can also impact ontology modeling methodologies
Research Efforts	The W3C Best Practices Committee is working on efforts in this direction for the W3C standards.
Development Issues	Many vendors have already committed to some form of OWL as their modeling paradigm of choice.

4. Design Patterns for Semantic Solutions.	
Barriers	A classic problem in knowledge modeling (the so-called "knowledge acquisition bottleneck") has to do with the fact that domain practitioners understand the domain, and how it interacts with work culture, business process and legacy systems, but are typically not sufficiently trained in logic to build effective knowledge models.
Research Questions	Research is required to determine and develop domain- and industry-specific classes of problems and solution patterns that match them. These patterns should include enough guidance to build a logically sound model without requiring intense logical modeling skills on the part of the modeling practitioner. In order to be useful, these patterns must be responsive to the array of issues (business process, work culture, etc.) outlined above. The ability to link these patterns to extant information sources (legacy data and applications) is also essential to their adoption in real-world settings. Can they cope with the shagginess of real-world workplaces, with their idiosyncratic legacy data structures and applications, and the work processes that have grown up around them? The validity of such patterns, however they are delivered, must be determined by systematic tests – do they really help modelers without logic training build high quality models, more quickly? User groups, longitudinal studies, controlled experiments and case studies are required to make a compelling case. In the absence of such controlled studies, arguments about "upper ontologies" will continue to seem like esoteric arguments about frames dancing on heads of pins to most practitioners
Research Efforts	Some so-called "upper ontology" (or perhaps more properly, "reference ontology") efforts attempt to address this issue by encoding domain information and methodology into an ontology or a library of ontology models. There are also efforts to built automatic translators from legacy data sources (database schemas, information architectures, controlled vocabularies, etc.) to ontologies.
Development Issues	Vendors need to produce knowledge modeling and deployment tools that insulate the domain experts as much as possible from the detailed syntax and other system management issues of ontology development.

5. Ontology Engineering lifecycle management	
Barriers	During ontology development, modelers find a need to perform global edits on their ontologies, merging some parts that were developed separately, dividing some models into smaller parts, managing multiple namespaces, managing components and refactoring. All of these operations are familiar from other kinds of software modeling (e.g., object oriented modeling, data base schemas, etc.)
Research Questions	Research is needed to determine what operations are most needed to support a realistic ontology development effort, including user groups and real-world case studies. Methods are needed to determine when a factored or decomposed ontology "is the same as" (i.e., expresses the same facts as) the original.
Research Efforts	The Prompt plug-in for Protégé addresses some of these issues.
Development Issues	Vendors of ontology tools need to provide a suite of capabilities in a package that is familiar to developers (e.g., Eclipse), and a way to integrate them into a complete semantic solution.

6. Ontology reuse.	
Barriers	A large part of the value proposition for ontologies lies in the possibility to re-use ontologies from one project to the next. This implies that ontologies must be designed for re-use. However, real-world attempts to re-use ontologies have resulted in disappointing results. In some cases, even ontologies that have been vetted by being parts of real-world applications prove to be incomplete or inconsistent (typically because some part of the ontology was taken off the critical path in the solution by replacement with a conventional technology), or as was found in [Mitre-2004], the re-usable ontologies required considerable adjustment to fit into the application at hand.
Research Questions	The problem of designing artifacts for use in situations yet to be specified is a classic problem of software re-use, and it has not gone away just because the artifacts are ontologies instead of conventional software components. The required research in this area is at two levels; first, at the level of a framework for modularizing and re-using ontologies, and second, at the level of constructing actual reusable ontology assets for specific domains and competencies.
Research Efforts	Research on upper ontologies and ontology libraries addresses some of these issues, but typically is not sensitive to the profound impact that different ontology capability requirements have on ontology development.
Development Issues	Vendors will need to provide support for what is effectively a new paradigm of system development, based on the ontology reuse results of this research.

7. Visualization.	
Barriers	Modelers often complain of how difficult it is come to a sufficient understanding of an ontology in their domain or application. These complaints usually come in the form of a request to be able to "see" some aspect of the ontology. A wide range of approaches to this problem (including parabolic, graph layout, interactive, 3-D, multi-media, etc.) still leaves large segments of user populations cold.
Research Questions	Research to extend the already impressive pantheon of ontology visualization techniques is not likely to address this problem. The root of the issue is that various user segments have differing and sometime idiosyncratic needs when it comes to visualization. Especially for engineering domains, research into visualization techniques that conform to various industry standard schematics could smooth adoption paths by presenting the complexity of an ontology in a form that is already familiar to professionals in a particular field.
Research Efforts	Research of this sort is currently being carried out by technology vendors in their own fields.
Development Issues	Vendors will need to provide robust, professional versions of these capabilities integrated into semantic technology platforms.

TOPQUADRANT, INC.

ABOUT THE FIRM

TopQuadrant, Inc. is a leading consultancy focused on the emergence of semantic web technologies and solutions. We provide complete semantic solution services, industry research and advanced technology R&D.

As semantic solution providers, we envision, design, construct and deploy solutions that exploit semantic technology standards, tools and products, semantic web-grid-services and knowledge-based technologies.

As independent researchers, we examine markets, assess technologies, and publish results. In advanced technology R&D, we conceptualize, design and validate next-generation intelligent systems using semantic and knowledge-based technologies for a range of application areas.

TopQuadrant serves as a trusted intermediary to help enterprises envision, architect, plan and realize knowledge-based solutions that deliver significant ROI. Our professionals bring expertise in artificial intelligence, object technology, adaptive systems, ontology engineering, knowledge and content management, publishing and media, semantic web and grid computing, and methodologies for knowledge, software and systems engineering.

TopQuadrant has developed a set of unique tools, methodologies, and services that jump-start successful building of semantic solutions, and help clients transition to next generation, semantically integrated systems while sustaining and optimizing their investments in current technologies. These include:

- Semantic Solution Development Services, including optimal technology and vendor selection, ontology development, and full implementation of semantic solutions.
- Solution Envisioning, scenario-driven workshops to explore system options and design future solutions through analogies and examples using Capability Cases.
- TopMind™, executive briefings on semantic technology; hands-on trainings in semantic web standards, languages, tools and ontology development.
- TopDrawer™, a comprehensive knowledge-base of semantic technology capability cases — solution patterns for ontology-based applications.
- TopConnexion™, a multi-company knowledge service that conducts research; publishes case studies, technical assessments, and whitepapers; and produces workshops and conferences.

ABOUT THE AUTHORS:

Mills Davis

Mills Davis, managing director with TopQuadrant, brings more than 20 years experience as a consultant, industry analyst, and professional services executive. Mills specializes in strategic applications of next-wave semantic, information, content, and media technologies. His consulting clients include leading technology manufacturers, Global 2000 corporations, and government agencies in the Americas, Europe, and Asia. A researcher and industry analyst, Mills has authored more than 100 reports, whitepapers, articles, and industry studies. He is a frequent speaker at industry events, and his work has appeared in more than 30 trade magazines and journals.

Dean Allemang

Dean Allemang specializes in innovative applications of knowledge technology, and brings to TopQuadrant over 15 years experience in research, deployment and development of knowledge-based systems. Prior to joining TopQuadrant, Dr. Allemang was the Vice-President of Customer Applications at Synquiry Technologies, where he helped Synquiry's customers understand how the use of semantic technologies could provide measurable benefit in their business processes. Dr. Allemang has filed two patents on the application of graph matching algorithms to the problems of semantic information interchange. In the Technology Transfer group at Swiss Telecom he co-invented patented technology for high-level analysis of network switching failures. He is a co-author of the Organization Domain Modeling method, which addresses cultural and social obstacles to semantic modeling, as well as technological ones. Dr. Allemang combines a strong formal background (MSc in Mathematics, University of Cambridge, PhD in Computer Science, Ohio State University) with years of experience applying knowledge-based technologies to real business problems.

Robert Coyne

Robert Coyne is an Executive Partner of TopQuadrant, Inc., a consulting company specializing in knowledge-based solutions, the semantic WEB, semantic technology trainings, ontology engineering and solution envisioning. Prior to joining TopQuadrant, Robert was Executive Vice President & Chief Technology Officer of Solution Technology International, Inc. where he co-designed a complete, web-based business-to-business-to-consumer (B2B2C) 'straight through processing' platform for the insurance industry. As a Senior Consultant at IBM Global Services from 1995-98, he was co-architect of and a major content contributor to IBM's object-oriented methodology for client engagements, and served as a knowledge management /collaboration expert for the global Object-technology Practices. Dr. Coyne received a Ph.D. (1991) in Computer-Aided Design at Carnegie Mellon University where he held appointments through 1995 as a Research Faculty at the NSF funded Engineering Design Research Center, lecturer in the School of Computer Science and Adjunct Faculty in the Department of Architecture. He has authored over 30 technical papers, journal articles and book chapters and is co-author of a forthcoming book "Solution Envisioning with Capability Cases", to be published by Addison-Wesley Professional Series in Spring, 2005.