ENABLING NETWORKED KNOWLEDGE

Stefan Decker    Manfred Hauswirth

September 2008
ENABLING NETWORKED KNOWLEDGE

Stefan Decker\textsuperscript{1} \quad Manfred Hauswirth\textsuperscript{1}

\textbf{Abstract.} Despite the enormous amounts of information the Web has made accessible, we still lack means to interconnect and link this information in a meaningful way to lift it from the level of information to the level of knowledge. Additionally, new sources of information about the physical world become available through the emerging sensor technologies. This information needs to be integrated with the existing information on the Web and in information systems which requires (light-weight) semantics as a core building block. In this position paper we discuss the potential of a global knowledge space and which research and technologies are required to enable our vision of networked knowledge.

\textsuperscript{1}Digital Enterprise Research Institute (DERI)

\textbf{Acknowledgements:} The work presented in this paper was supported (in part) by the Líon project supported by Science Foundation Ireland under Grant No. SFI/02/CE1/1131.


Copyright © 2008 by the authors
1 What is Networked Knowledge?

The wealth of information and services on today’s information infrastructures like the Internet and the Web has significantly changed everyday life and has substantially transformed the way in which business, public and private interactions are performed. The economic and social influence of the Web is enormous, enabling new business models and social change, and creating wealth. However, we have barely scratched the surface of what information technology can do for society. The Web has enabled information creation and dissemination, but has also opened the information floodgates. The enormous amount of information available has made it increasingly difficult to find, access, present and maintain information. As a consequence, we are literally drowning in information and starving for knowledge. However, systematic access to knowledge is critical for solving today’s problems - on individual and organisational as well as global levels.

Although knowledge is inherently strongly interconnected and related to people, this interconnectedness is not reflected or supported by current information infrastructures. The lack of interconnectedness hampers basic information management and problem-solving and collaboration capabilities, like finding, creating and deploying the right knowledge at the right time. Unfortunately, this is happening at a time when the problems humanity has to face are more difficult than ever (e.g., climate change, energy and resource shortages, or globalisation).

New methods are required to manage and provide access to the world’s knowledge, for individual as well as collective problem solving. The right methods and tools for interconnecting people and accessing knowledge will contribute to solving these problems by making businesses more effective, scientists more productive and bringing governments closer to their citizens. Thus, the focus on Enabling Networked Knowledge is essential.

What is Networked Knowledge and why is it important?

Besides the creation of knowledge through observation, networking of knowledge is the basic process to generate new knowledge. Networking knowledge, can produce a piece of knowledge whose information value is far beyond the mere sum of the individual pieces, i.e., it creates new knowledge. With the Web we now have a foundational infrastructure in place enabling the linking of information on a global scale. Adding meaning moves the interlinked information to the knowledge level: Web + Semantics = Networked Knowledge. Knowledge is the fuel of our increasingly digital service economy (versus manufacturing economy); linking information is the basis of economic productivity.

Fortunately, current developments are helping to achieve these goals. Originating from the Semantic Web effort, more and more interlinked information sources are becoming available online, leading to islands of networked knowledge resources and follow-up industrial interest. Due to its forward-looking investment of the European Commission, Europe is playing a central and internationally recognised role in this development. Since more and more different kinds of information sources become available, the main goal for the future is to build up on this leadership position.

As an example of a rapidly growing information space, Gartner predicts that “By 2015, wirelessly networked sensors in everything we own will form a new Web. But it will only be of value if the ‘terabyte torrent’ of data it generates can be collected, analyzed and interpreted.” [11] Making
sensor-generated information usable as a new and key source of knowledge will require its integration into the existing information space of the Web. Now is the time to tackle the next step: exploiting semantics to create an overall knowledge network bridging the islands enabling people, organisations and systems to collaborate and interoperate on a global scale, and bridging the gap between the physical world and the virtual world so that the information on the Web (the virtual world) can directly influence activities in the real world and vice versa. This integrated information space of networked knowledge will impact all parts of people’s lives.

Hypothesis

*It is our central hypothesis that collaborative access to networked knowledge assists humans, organisations and systems with their individual as well as collective problem solving, creating solutions to problems that were previously thought insolvable, and enabling innovation and increased productivity on individual, organisational and global levels.*

*In our opinion research needs to aim to:*

1. *develop the tools and techniques for creating, managing and exploiting networks of knowledge;*
2. *produce real-world networks of knowledge that provide maximum gains over the coming years for human, organisational and systems problem solving;*
3. *validate the hypothesis; and*
4. *create standards supporting industrial adaptation.*

This overall research vision is broken down into three overall complementary research strands, which form the *Networked Knowledge House* (see Figure 1).

![Networked Knowledge House](image-url)
Social Semantic Information Spaces deal with organization, linking, and management of knowledge on the Web. Semantic Reality addresses the integration of information from the physical world with knowledge in the virtual world (Social Semantic Information Spaces), the creation of knowledge out of information about the physical world, and efficient mechanisms to access this information at large scale via sensors. The technologies created by these basic research strands are then applied in and customized to a set of application domains, which we identified as most relevant to our work. This in turn requires research due to the specific requirements of the domains. Of course, the given list of application-oriented research domains is not comprehensive. A number of important domains are not listed, for example, environmental monitoring, traffic management and intelligent driving, logistics and tracking, or building management, to name a few, as they are beyond the scope of DERI at the moment.

In the following sections we explain the Networked Knowledge House in more detail.

2 Why Enabling Networked Knowledge?

The World Wide Web has dramatically altered the global communications and information exchange landscape, removing barriers of access, space and time from business and social transactions. The Web has created new forms of interaction and collaboration between systems, individuals and organisations. The dramatic development of the Web and the changes it has made to society are only a glimpse of the potential of a next-generation information infrastructure connecting knowledge and people. By interlinking the world's knowledge and providing an infrastructure that enables collaboration and focused exploitation of worldwide knowledge, Social Semantic Information Spaces and Semantic Reality, which will be explained in detail in the following sections, enable individuals, organisations and humanity as a whole to socialise, access services and solve problems much more effectively than we are able to today. The Web is already able to provide us with information, but lacks support for collaboration, knowledge sharing and social interaction. An information infrastructure supporting effective collaboration and augmented with interlinked and networked knowledge will support human capabilities and enable human-centric access to services and knowledge. We can already see the first glimpses of this in current online social networking sites (currently serving hundreds of millions of users), even though these sites are just data silos and do not interconnect knowledge efficiently.

Social Semantic Information Spaces and Semantic Reality as a networked knowledge infrastructure also make businesses more effective and scientists more productive by connecting them to the right people and to the right information at the right time and enabling them to recognise, collect, and exploit the relationships that exist between the knowledge entities in the world.

Vannevar Bush [3] and Doug Engelbart [5] were proposing similar infrastructures in 1945 and 1962. However, the technology available then was not advanced enough to realise their visions. Figuratively speaking, their ideas were proposing jet planes when the rest of the world had just invented the parts to build a bicycle. With the Semantic Web effort delivering standards to interconnect information globally and the Social Web showing how to collaborate on a global scale, now a window of opportunity has opened up to make these visions a reality and build a truly global networked knowledge infrastructure.
3 Social Semantic Information Spaces

One of the most visible trends on the Web is the emergence of “Social Web” (or Web 2.0) sites which facilitate the creation and gathering of knowledge through the simplification of user contributions via blogs, tagging and folksonomies, wikis, podcasts and the deployment of online social networks. The Social Web has enabled community-based knowledge acquisition, with efforts like Wikipedia demonstrating the “wisdom of the crowds” in creating the largest encyclopedia in the world. Although it is difficult to define the exact boundaries of what structures or abstractions belong to the Social Web, a common property of such sites is that they facilitate collaboration and sharing between millions of users. However, as more and more Social Web sites, communities and services come online, the lack of interoperation among them becomes obvious: the Social Web platforms create a set of isolated data silos – sites, communities and services that cannot interoperate with each other, synergies are expensive to exploit, and reuse and interlinking of data is difficult and cumbersome. The entities in the Social Web are not only data artefacts. Instead, it is a network of interrelated users and their concerns as well as content that the users are related to as producers, consumers or commentors. To enable machines to assist us with the detection and filtering of knowledge, many of these often implicit links have to be made explicit.

Social Semantic Information Spaces are a combination of the Semantic Web, the Social Web, collaborative working environments and other collaboration technologies. The goal behind Social Semantic Information Spaces is to create a universal collaboration and networked knowledge infrastructure, which interlinks all available knowledge and their creators. The resulting infrastructure would finally enable knowledge management capabilities as expressed by visionaries like Vannevar Bush and Doug Engelbart.

Figure 2 shows how Social Semantic Information Spaces fit into the current landscape: Communication and collaboration tools are augmented and made interoperable with Semantic Web technologies. The result is a network of accessible interlinked knowledge, enabling productive collaboration and knowledge management.

![Social Semantic Information Spaces](image.png)

Figure 2: Social Semantic Information Spaces
In the following we list a couple of specific examples of enabling technologies and describe where
and how they fit in the idea of a Social Semantic Information Space. All these technologies are
just starting up and further research is necessary to ensure their development into broadly adopted
technologies. However, convergence between some of the different efforts are already recognisable
today.

3.1 Semantic Social Networks

From the beginning, the Internet was a medium for connecting not only machines but people.
Email, mailing lists, the Usenet, and bulletin boards allowed people to connect and form online
social networks, typically around specific topics. Although these groups did not explicitly define
social networks, the ways people acted and reacted did so implicitly. The early Web continued this
trend. More recently, sites such as Friendster and LinkedIn have brought a different notion of online
communities by explicitly facilitating connections based on information gathered and stored in user
profiles. However, all these sites are stovepipes and lock the information in: using the social network
information for other purposes, e.g., for prioritising email as discussed in [6], requires standardised
data exchange mechanisms. Initial crystallisation points to remedy this situation are efforts like the
Friend-of-a-Friend vocabulary (FOAF\(^1\)) or the Semantically-Interlinked Online Communities initiative (SIOC\(^2\) [2]). The SIOC initiative may serve as an example how social networking information
can be interlinked with content such as online discussions taking place on blogs, message boards,
mailing lists, etc. In combination with the FOAF vocabulary for describing people and their friends,
and the Simple Knowledge Organization Systems (SKOS) model for organizing knowledge, SIOC
enables the linkage of discussion postings to other related discussions, people (via their associated
user accounts), and topics (using specific “tags” or hierarchical categories). As discussions begin to
move beyond simple text-based conversations to include audio and video content, SIOC is evolving
to describe not only conventional discussion platforms but also new Web-based communication and
content-sharing mechanisms.

Some social networking sites, such as Facebook, are also starting to provide query interfaces
to their data, which others can reuse and link to via the Semantic Web. Thus this information
becomes part of the Web of information, which may be used or reused for a variety of purposes,
providing crystallisation points for a network of knowledge.

3.2 Semantic Collaborative Technologies

Apart from the specific data representation mechanisms outlined above, other mechanisms and
technologies contribute to the emergence of Social Semantic Information Spaces on the Web. The
Social Semantic Desktop (SSD) [4] effort (materialised in the EU IP project NEPOMUK [7]) is
aiming at organising information on the desktop by using Semantic Web metadata standards. Onto-
logies capture both a shared conceptualisation of desktop data and personal mental models. RDF
serves as a common data representation format. As Web services can describe their capabilities and
interfaces in a standardized way and become Semantic Web Services, on the desktop, applications,
or rather their interfaces, can be modelled in a similar fashion. Together, these technologies provide
a means to build the semantic bridges necessary for data exchange and application integration. The

\(^1\)http://www.foaf-project.org
\(^2\)http://www.sioc-project.org
Social Semantic Desktop has the potential to transform the conventional desktop into a seamless, networked working environment, by obliterating the borders between individual applications and the physical workspace of different users.

In contrast to desktop applications Wikis have become popular Web-based collaboration tools and are widely deployed to enable organizing and sharing of knowledge. Wikis gained popularity since they enable the management of online content in a quick and easy way by “group-editing” using a simple syntax. However, typically knowledge collected in Wikis cannot be reused easily automatically and is only usable for human consumption. Semantic Web techniques applied to Wikis [8] leverage semantic technologies to address this challenge. They provide means to rapidly acquire formal knowledge also by non-knowledge engineers, and to create a network of this knowledge linked with other information sources. A typical example is the Semantic MediaWiki\(^8\), which enables the evolution of Wikipedia into a reusable knowledge source enabling automatic processing and human support.

4 Semantic Reality

Until now the virtual world of information sources on the World Wide Web and activities in the real world have always been separated. However, knowledge accessible on the Web (the virtual world) may influence activities in the real world and vice versa, but these influences are usually indirect and not immediate. In contrast to this, imagine a world where:

- Cars know where the traffic jams are and traffic can be managed based on real-time input about the traffic situation.

- Medical data monitored through body sensor networks is automatically included into a patient’s electronic healthcare record. Should a critical condition be detected by these sensors, the patient can be physically located and the closest doctor on duty can be guided to the patient, whilst preparing the necessary resources in the hospital the patient is to be transferred to.

- Your calendar knows how long the queue is at your physician.

- Your travel planner knows that the train is delayed before you go to the train station.

- Or generally, scarce resources can be managed efficiently and in-time.

The advent of sensor technologies in conjunction with the Semantic Web now provides the unique opportunity to unify the real and the virtual worlds as for the first time we have the necessary infrastructures in place or large-scale deployment will happen in the short term. Their combination will enable us to build very large information spaces and infrastructures which for the first time facilitate the information-driven online integration of the physical world and computers. Similarly, as the Internet has changed the way people communicate in the virtual world, Semantic Reality extends this vision to the physical world, enabling novel ways for humans to interact with their environment and facilitating interactions among entities of the physical world (Internet of Things). The physical world will be represented in cyberspace and information on our environment

---

\(^8\)http://meta.wikimedia.org/wiki/Semantic_MediaWiki
will become ubiquitously available on the Internet. This integrated information space has a wide range of applications in monitoring, manufacturing, health, tracking and planning.

We call it Semantic Reality because due to the possible scale and the overall goal of creating networked knowledge, understanding of information, i.e., semantics, plays a central role. Whether semantics is based on statistics, logical descriptions, or hybrid approaches does not matter. In fact, we believe that a wide spectrum of approaches and their combinations will be necessary to cover the diverse requirements. Semantic Reality aims at an integrated information space very much in line with the design philosophy of the original Internet, which embraces community-driven agreement processes, emergent behaviour and self-organisation, but adding semantics as a key enabling ingredient. Without machine-processable semantics such a large-scale system cannot work to its fullest extent. Yet, semantics must be light-weight, fault-tolerant, must support dynamic change, and has to be able to deal with incomplete, wrong, and noisy information in order to be applicable and useful in a global-scale heterogeneous environment. The rationale for success could be along the lines of “a little bit of semantics gets you a long way.”

Figure 3 shows how Semantic Reality fits into the overall picture: Sensors connect the physical world to the computer, Social Semantic Information Spaces create and connect virtual worlds, and Semantic Reality integrates these two into one uniform information space which will provide novel ways of monitoring, controlling and influencing the environment, and how people and enterprises collaborate.

![Diagram of Semantic Reality](image)

*Figure 3: Semantic Reality*

The ultimate goal of Semantic Reality is “to deliver the right knowledge to the right people at the right time.” This requires the adequate description of information, people and their requirements, and a temporal view on data sources, be they “real” or “virtual”, i.e., a unified model of evolution of (integrated) information sources, thus moving from a static to a dynamic model of the Web and the physical world. This is currently taken into account only to a limited extent on the Web and only for “closed” applications, e.g., RSS feeds or blogs.
Semantic Reality shares several goals and properties with ubiquitous and pervasive computing and ambient intelligence. Though drawing on a large body of work in sensor networks, embedded systems, ubiquitous and pervasive computing, ambient intelligence, networking, distributed systems, distributed information systems, artificial intelligence, software engineering, social networking and collaboration, and Semantic Web, Semantic Reality is different from these research domains as it pushes the boundaries further by aiming at large-scale integration of (possibly isolated) information islands and the integration of systems, which requires the central use of semantics for information-driven integration and a uniform/universal, but light-weight semantic model of information sources and information.

The sheer size of the possible systems poses quite novel and unique challenges. Semantic Reality systems can only be built, deployed, and maintained if a large degree of self-organization and automatization capabilities are being built into the infrastructures and their constituents, enabling automated deployment (plug-and-play), automated (re-) configuration, automated component and information integration, and tailored information delivery based on user context and needs in a service-oriented way. The previous characteristics require semantic descriptions as a central ingredient: User requirements and contexts, the constituents of the system, the dynamic data (streams) they produce, their functionalities, and requirements – all need to be described using light-weight semantic mechanisms to enable a machine-understandable information space of real-world entities and their dynamic communication processes on a scale which is beyond the current size of the Internet.

In the following, we briefly discuss some of the core challenges and hint at possible strategies to address them.

Large-scale and open semantic infrastructures and flexible abstractions are required to enable the large-scale design, deployment and integration of sensor/actuator networks and their data. The integration has to happen on both the technical (data and network access) as well as on the semantic level (“What does the (stream) data provided actually mean?”). The infrastructure has to be open and easily extensible to address the heterogeneity issues which go far beyond those seen to date on the Internet. The infrastructure will draw on key enabling technologies such as (semantic) overlay networks using P2P technology to achieve scalability and light-weight semantic formats based on RDF and microformats. Middleware systems such as the Global Sensor Network (GSN) platform [1] are examples aiming at the development of a general-purpose middleware supporting these requirements. GSN is work-in-progress and provides a flexible middleware layer which abstracts from the underlying, heterogeneous sensor network technologies, supports fast and simple deployment and addition of new platforms, facilitates efficient distributed query processing and combination of sensor data, provides support for sensor mobility, and enables the dynamic adaption of the system configuration during runtime with minimal (zeroprogramming) effort. The GSN implementation is available from http://gsn.sourceforge.net/.

Query processing, reasoning, and planning based on real-world sensor information will be core functionalities to exploit the full potential of Semantic Reality. The key research problems to overcome are the very large scale, the number of distributed information sources, the time-dependency of the produced data (streams), and the fact that the data is unreliable and noisy. For query processing this means to support distributed query processing and load-balancing at large scales with only incomplete views on the state of the overall system. In this context, dis-
distributed event-based infrastructures are of specific interest ("reactive" queries). Users should be able to register expressive, semantic "patterns" of interest and be notified by the system as soon as information satisfying their interests becomes available.

Also, new approaches for distributed reasoning and reasoning on time-dependant information, taking into account modalities and being based on an open-world assumption will be necessary. The size and the physical distribution of data will require new approaches combing logical and statistical approaches which will have to trade logical correctness with statistical guarantees and expressivity with scalability. Essentially, the goal is to enable "The World is the Database" scenarios with support for structured querying, integrated views (real-world information with virtual information), aggregation and analyses, and open, distributed reasoning over large, incomplete, and approximate data sets.

**Cross-layer integration and optimization** will play a central role due to the extremely heterogeneous environment – a wide range of sensing devices with very heterogeneous hardware and processing characteristics; information systems and architectures along with virtual information streams which considerably increase complexity – and the various and often contradicting requirements on the different system levels. For example, sensor networks are optimized for life-time and offer only primitive programming and query models. If this is combined with the "wrong" distribution approach, e.g., a distributed hash table for discovery and the "wrong" distributed query processing approach which does not limit expressivity of queries, this will lead to an inefficient system design and limit the life-time of sensors by draining their power sources because of incompatible processing strategies at the different levels.

**Semantic description and annotation** of sensors, sensor data and other data streams will enable the flexible integration of information and (distributed) discovery of information. For scalability, integrity, and privacy reasons this has to be supported in a distributed fashion, for example, through semantic peer-to-peer systems. A prerequisite for discovery is the meaningful semantic description of sensors and sensor data by the manufacturer and by the user; for example, by the manufacturer through IEEE 1451 standard compliant Transducer Electronic Data Sheet (TEDS) [9], which essentially give a (non-semantic) description of the sensor that can very easily be ontologized, or via an ontologized subset of SensorML [10] which provides standard models and an XML encoding for describing sensors and measurement processes, and by the user by extending these basic descriptions with annotations adding more information, meaning, and links to other information.

Especially the annotation of sensor data itself will be highly relevant to understand the meaning of the produced data and share this knowledge. Visualization environments to support the annotation process will be of high importance. Such environments may support simple graphical annotation up to annotation with claims and findings in the case of scientific data. This derived knowledge then can be used again in the discovery process and will help to prevent "data graveyards" where interesting (measurement) information is available but cannot be used because the knowledge about its existence and meaning has been lost (the typical "PhD student finishes" syndrome). Due to the possibly large sizes of the produced data this poses additional scalability problems. As discovery, semantic annotation has to be supported in a distributed fashion, for example, by distributed semantic Wikis.
Emergent semantics, self-organization, and plug-and-play are required to build working systems at the envisioned large scales where top-down system control, configuration, and enforcement of standards will be a very hard problem or even impossible. As we can see from the current community processes on the Web, a lot of successful de-facto standards develop bottom-up. Conversely, these processes support the incremental development of standards and knowledge. The system must be able to self-organize and adopt its behavior in a plug-and-play fashion within organizational boundaries based on semantic understanding and agreement. Semantic understanding and agreements in turn will depend on dynamic processes which support (semi-)automatic assessment of the levels of agreement and their correctness. Such emergent semantic agreements can then be used as the basis for standardization (ontologies). Conversely, semantic formats can be advanced through such processes.

Semantically enriched social network and collaboration infrastructures enable the targeted delivery of knowledge and information based on context description and actual user needs. The ubiquity of information requires means to filter and direct data streams on a need-to-know basis. The definition of user profiles, needs and contexts are key features enabling targeted information delivery and avoiding overload. Social networking information enables both - information sharing and information filtering based on interests and information needs.

Development support and tools along with experimental platforms and simulation tools will be necessary for efficient application development and testing. This means the availability of visual programming tools which support the developer in designing the acquisition, combination and integration of data. These designs then can be compiled down to the required target platforms (both sensor and back-end platforms, e.g., for business processes). To test applications, experimental testbeds along the lines of PlanetLab (http://www.planet-lab.org/) are essential as many of the characteristics of Semantic Reality systems require experimental evaluation under real-world conditions especially in terms of scale and distribution. To further evaluate applications, the integration of experiments and simulations should be supported in a seamless way, i.e., a test of an application in an experimental testbed should support the inclusion of simulation without changes to the application code. This means that parts of an application (or the complete application) should be able to run on an experimental testbed or on a simulator or any combination of those. On the application level modern paradigms such as service-oriented architectures, service mash-ups and Web 2.0-like functionalities should be available and be supported.

Integrity, confidentiality, reputation, and privacy are the key security requirements for business users and consumers. The provided information has to be resistant against technical errors and attacks, has to be stored and transported in a secure way, has to come from authentic and trustworthy sources and must ensure the privacy of its providers and users. Physical distribution can be beneficial here as it helps to avoid the creation of “Big Brother” scenarios which consumers and legislators would not tolerate.

Vertical integration of business processes ↔ middleware ↔ sensor/actuator networks relying on the above technologies and functionalities will then unleash the full potential of Semantic Reality. Sensor information, coming both from virtual and physical sources, are a key requirement for agile business processes requiring minimal human intervention.
5 Application-Oriented Research Domains

The Web has already influenced many different areas of society. The introduction of Social Semantic Information Spaces and Semantic Reality may have a similar influence, but like the Web, the transition of these new technologies into application areas is usually slow. To ensure rapid uptake and to provide maximum benefit to society, dissemination of research should focus on a number of carefully selected application research domains. These research domains investigate the adoption and uses of Social Semantic Information Spaces and Semantic Reality, combining a critical mass of technology-oriented research with the research on needs in specific application environments to initiate ground-breaking innovation. Example research domains are:

**eHealth and Life Sciences:** The objective of the eHealth and Life Sciences domain is to reduce the cost associated with the drug research and delivery process, making clinical research more efficient through data integration, and enabling patients’ self-management of disease through telehealth, e.g., remote patient monitoring. Due to the heterogeneity of the eHealth domain, semantics is a crucial ingredient in achieving this objective.

**eScience:** The objective of the eScience domain is to improve collaboration among scientists working on computationally intensive problems, carried out in highly distributed network environments. Semantic support for distributed collaboration and annotation of scientific data and publications are of particular interest in our opinion.

**Telecommunications:** The objective of the telecommunications domain is to exploit semantic technologies for enabling telecoms to develop value-added communication services that will interface humans and machines, exploit machine-readable data, and improve intra-enterprise human communication and knowledge management. Context-information generated by sensors in conjunction with virtual information and unified communication profiles is of particular interest to enable new technology-supported communication paradigms.

**eBusiness and Financial Services:** The objective of the eBusiness and Financial Services domain is to apply new technology in the key areas of extracting business meaning from unstructured information, uncovering meaning within a business context, smarter Business Information Systems that can add meaning as they operate and communicate business information.

6 An example application scenario

To illustrate the possibilities of “Networked Knowledge” we present a simple application scenario: Siobhan Hegarty who lives in Galway is pregnant with her second child. During her first pregnancy Siobhan has suffered from elevated blood sugar levels which can endanger the unborn child and the mother. The problem with elevated blood sugar levels in pregnant women is that the important characteristic which requires fast reaction is the change in the blood sugar level. Thus measuring it a few times a day is not sufficient but constant monitoring is required. Fortunately, mobile sensors are available which enable Siobhan to leave the hospital while her general practitioner (GP) stills get the relevant information. Siobhan is being equipped with a mobile blood sugar sensor which can transmit readings via Bluetooth. The device is paired with Siobhan’s mobile telephone which transmits the sensor readings via GSM. Additionally she gets a GPS device which records her position and sends it via her mobile.
Siobhan’s GP, Dr. James Mooney, enters the necessary monitoring requirements into his Care2X healthcare information system (http://www.care2x.org/) along with rules when to raise an alarm and to whom. For example, the system will call Siobhan and warn her via a synthesized message, while James is informed via a text message on his beeper which he wears all the time. The sensor readings from Siobhan’s blood sugar and GPS sensors are directly fed back into James’s Care2X system.

Let us assume that after some time, Siobhan’s blood sugar levels change dramatically and the alarm rules are set off. Now it is important to get Siobhan to a doctor as fast as possible, or vice versa – a doctor to Siobhan. Besides notifying Siobhan and James, the Care2X system accesses the information system of the hospital and requests a proposal, whether it is better to bring Siobhan into the hospital via an ambulance or bring a doctor to Siobhan. The hospital information system which knows the GPS position of all doctors with matching skills to help Siobhan and of all ambulances produces an optimal plan based on real-time sensor input from the traffic control system of the city. Given the current positions of available ambulances and doctors with the necessary skills, the optimal strategy is to pick up the endocrinologist Dr. Sarah O’Connor from her home with a nearby ambulance and bring her to Siobhan.

Unfortunately, while this plan was calculated two important changes to the scenario have happened: (1) No more readings from Siobhan’s GPS are received, probably because she has entered a building or because the device ran out of battery and Siobhan does not respond to calls on her mobile and (2) the last blood sugar readings show some strange and unknown pattern which neither James nor Sarah can interpret. As a reaction, the system now tries to locate Siobhan via other means: The system tries to determine her position via triangulation of her mobile and additionally informs all Bluetooth access points in the vicinity of her last position to send a message if they recognize any of her Bluetooth devices.

The strange patterns in the blood sugar readings worry James and Sarah and they decide to use their country-wide social network of clinical specialists to look for doctors who probably have already seen similar patterns. Additionally, they search medical databases on the Web for annotations describing such patterns. As a result of their search they find information which looks similar to the pattern they have seen but the result is inconclusive. In parallel, a colleague of them from Dublin who also participates in the social network they sent the symptoms to, informs them that the pattern may indicate a malfunction of the blood sugar sensor and describes his experiences.

In the meantime, Siobhan could be located by a Bluetooth access point. To be on the safe side the ambulance with Sarah on board is sent to her location and finds her in good condition. However, an examination reveals that indeed her blood sugar levels had changed dangerously and Siobhan is treated on the spot. After this successful intervention James and Sarah annotate the sensor readings to permanently store their findings. Their findings are stored in James’s Care2X system, the hospital’s information system and also made accessible to other doctors in the national infrastructure along with the actual sensor readings in a secure and anonymized way.

7 Core research topics for the next years

Core research objectives for the next years include the foundation for the creation of knowledge networks and collaboration infrastructures, which will support human capabilities and enable the human-centric access to services and knowledge on a global scale, opening up new opportunities
for individuals and organisations. Example topics include:

**Foundations for semantic collaboration:** the development of technologies supporting distributed collaboration with a focus on the Semantic Desktop and the Web. Examples include APIs and ontologies that reuse existing social networking information from sites to assess the identity and relevance of information.

**Scalable reasoning and querying facilities for knowledge:** Current knowledge bases are not able to exploit and analyse knowledge, which would be necessary in order to learn from it. To exploit the available knowledge, scalable querying and data mining mechanisms need to be developed. Additionally, dynamic data sources (streams), modalities (time, space) and noise in the data, need to be taken into account and be supported.

**Frameworks for semantic sensor networks:** Currently sensor networks and the data they produce lack semantic description, making it difficult to integrate data coming from large-scale, dynamic sensor networks with existing information. It is necessary to develop practical semantic description methods for sensors and mobile device middleware, enabling the integration of sensor data with knowledge from knowledge networks. This will be part of a more general practical and deployable semantic service-oriented architecture.

8 Creating impact

Knowledge networks are not created in a vacuum, but inside a highly dynamic information infrastructure – the Web, which provides us with a living laboratory enabling us to validate our approaches and hypothesis, and to improve our ideas.

The first way to validate the hypothesis is to study the usage of emerging networks of knowledge on the Web. Many application areas are dealing with the challenges of large, open, heterogeneous, dynamic and distributed environments. Semantics is an important cornerstone for achieving scalability of knowledge interchange and interoperability. Projects should validate this hypothesis by investigating the required research and approaches in application domains, ranging from eHealth to eGovernment to eLearning.

References


