# Semantic Reality –
# Connecting the Real and the Virtual World
## Position Paper

Manfred Hauswirth, Stefan Decker
Digital Enterprise Research Institute (DERI)
National University of Ireland, Galway
manfred.hauswirth@deri.org, stefan.decker@deri.org
http://www.manfredhauswirth.org/, http://www.stefandecker.org/
http://www.deri.ie/

## 1  What does Semantic Reality mean?

Until now the virtual world of information sources on the World Wide Web and the activities in the real world have always been separated. However, information on the Web (the virtual world) may influence the 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

- your car knows where the traffic jams are;

- your calendar knows how long the queue is at your physician;

- your shopping list knows that the milk is out at your local store;

- your travel planner knows that the train is delayed before you go to the train station;

- you know that your guests will be late before they call you;

- or more generally, scarce resources can be managed efficiently.

The advent of sensor technologies and the Semantic Web provide the unique opportunity to unify the real and the virtual worlds. It enables the building of very large infrastructures which for the first time facilitate the information-driven real-time integration of the physical world and computers on a global scale. The sheer

size of the possible infrastructures pose quite novel and unique challenges as the vision can only be engineered and deployed if a large degree of self-organization and automatizing capabilities are being built into the system and its 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. This requires semantic descriptions of the user needs and contexts, and of the system's constituents, the data streams they produce, their functionalities, and their requirements 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. We call this information space *Semantic Reality* as it comprises the virtual and the real world und requires the (machine-processsable) understanding of both.

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 will become ubiquitously available on the Internet. This integrated information space has a wide range of applications in monitoring, manufacturing, health, tracking and planning.

## 2   Research problems and enabling technologies

Though drawing on a large body of work in sensor networks, embedded systems, 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 areas as it targets the integration of all these domains on a large scale and has to take into account time and space as the intrinsic limitations of the physical world. Semantic Reality will provide an integrated information space very much along the design philosophy of the original Internet, which embraces community-driven agreement processes, emergent behavior, and self-organization, but adding semantics as a key enabling ingredient.

To achieve the overall vision of Semantic Reality a number of technological issues need to be addressed:

**Large-scale and open semantic infrastructures and flexible abstractions** are required to enable the 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.

**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 context 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.

**Semantic description and annotation** of sensors, sensor data and any 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 distributed semantic Wikis. Especially the annotation of sensor data itself will be highly relevant to understand the meaning of the produced data and share this knowledge. Due to the possibly large sizes of the produced data this poses additional scalability problems.

**Emergent semantics, self-organization, and plug-and-play** are required key characteristics to build a working system. In the Semantic Reality, top-down system control, configuration, and enforcement of standards will be a very hard problem. As we can see from the current community processes on the Web, a lot of successful de-facto standards develop bottom up. Additionally, these processes support the incremental development of standards and knowledge. To obtain meaningful results 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 again will depend on dynamic processes which support (semi-)automatic assessment of the levels of agreement and understanding and their correctness. Such emergent semantic agreements which are very closely related to folksonomies can then be used as the basis for standardization. Conversely, semantic formats can be advanced through such processes.

**Query processing, reasoning, and planning** based on real-world sensor informa-

tion bases will be core functionalities to exploit the full potential of Semantic Reality. However, the size and the physical distribution of data will require new approaches which will have to trade logical correctness with statistical guarantees.

**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 the Semantic Reality.

Semantics clearly is a core pillar to materialize the vision of Semantic Reality. Without machine-processable semantics such a large-scale system cannot work to its fullest extent. Yet, semantics must be light-weight, fault-tolerant, and must support dynamic change to be applicable and useful in a global-scale heterogeneous environment. The rationale behind Semantic Reality to be successful could be the old rule of thumb "A little bit of semantics gets you a long way."