Semantic Web Technologies for a Smart Energy Grid: Requirements and Challenges*

Andreas Wagner, Sebastian Speiser, and Andreas Harth

Institute AIFB, Karlsruhe Institute of Technology, Germany
{a.wagner,speiser,harth}@kit.edu

Abstract. The Smart Grid aims at making the current energy grid more efficient and eco-friendly. The Smart Grid features an IT-layer, which allows communication between a multitude of stakeholders and will have to be integrated with other "smart" systems (e.g., smart factories or smart cities) to operate effectively. Thus, many participants will be involved and will exchange large volumes of data, leading to a heterogeneous system with ad-hoc data exchange in which centralised coordination and control will be very difficult to achieve. In this paper, we show parallels between requirements for the (Semantic) Web and the Smart Grid. We argue that the communication architecture for the Smart Grid can be built upon existing (Semantic) Web technologies. We point out differences between the existing Web and the Smart Grid, thereby identifying remaining challenges.

1 Introduction

The Smart Grid – a radical redesign of the traditional energy grid – aims at profoundly changing the way how energy is created, distributed and consumed, thereby saving a considerable amount of energy [1, 2]. The envisioned Smart Grid should be [1]: (V1) flexible, i.e., fulfil current requirements, but also allow future extensions, (V2) accessible, i.e., allow access to/from all participants, (V3) reliable, i.e., assure quality of supply and (V4) economic, i.e., provide best value and allow for innovation and competition. Keeping the Smart Grid vision in mind, we wish to design a communication architecture, which achieves the above goals.

In the following, we describe a preliminary communication architecture developed in context of the Smart Grid project MeRegioMobil[1]. Our contribution is two-fold: 1) We outline requirements for a communication architecture for the Smart Grid and describe how (Semantic) Web technologies meet them. 2) We outline the remaining differences between the (Semantic) Web and the Smart Grid, thereby identifying future research problems.

The remainder of the paper is structured as follows: We present architecture requirements and an initial architecture in Section 2. In Section 3, we describe the differences between the Web and the Smart Grid and outline novel problems. We conclude with Section 4.

* This work was in part supported by the German Federal Ministry of Economics and Technology (MeRegioMobil, Grant 01ME09005).

[1] http://meregiomobil.forschung.kit.edu/

2 A Semantic Web Architecture for the Smart Grid

In this section, we present requirements for a communication architecture, which we derived from the Smart Grid vision and the literature, e.g., [1, 2]. Further, we introduce a (Semantic) Smart Grid communication architecture meeting the requirements.

- *R1 - General Requirement* A suitable architecture should incorporate a layered (data access, data representation and application layers) communication stack providing different functionalities and levels of abstraction. Employing a layered architecture leads to a more flexible and versatile Smart Grid communication, as varying technologies may be integrated and functionalities can be modified or replaced (V1).
- *R2 - General Requirement* We wish an appropriate architecture to be decentralised and thus omit a single point of failure, in order to provide the desired reliability (V3).
- *R3 - Data Access Layer* In order to allow full access to/from all participants we need a naming mechanism to uniquely identify each participant (V2).
- *R4 - Data Access Layer* The Smart Grid needs flexible, open and scalable data access procedures (V1/V2/V4). Flexibility means that a communication architecture should be able to facilitate heterogeneous participants employing hardware of lower or higher specification. Further, procedures only available under restrictive licenses to a selected number of participants might hinder innovation. Thus, standards should be open and royalty-free. As huge amounts of data are handled within the Smart Grid, data access procedures should be light-weight, i.e., scale well w.r.t. the data volume.
- *R5 - Data Representation Layer* We need structured and machine interpretable data models for representation of data semantics and context, in order to allow flexible application and business logic at higher layers (V1).
- *R6 - Data Representation Layer* Data semantics may be used for data integration, thereby fostering the access of heterogeneous participants (e.g., employing different data schemas) (V1/V2).
- *R7 - Application Layer* We have to support participants in making (automated) decisions, i.e., provide the means to express application and business logic (V4).
- *R8 - Application Layer* For allowing decision making based on logic, we have to fulfil (complex) information needs, thus we need to provide (semantic) querying features (V2).
- *R9 - Application Layer* Last, via logic we have to ensure data security and privacy, i.e., safeguard the grid against attacks and enable data protection mechanisms (V3).

There have been various proposals for a communication architecture for the Smart Grid, e.g., [1, 2]. In these works, the authors aim at a top-down architecture design approach employing a wide spectrum of both open and proprietary protocols. However, we aim at a concrete communication architecture, based on open and royalty-free standards, which are already applied in similar networks such as the Web. Further, while stating in [2] that a semantic layer (providing data semantics and context) is needed, no suitable standards are identified. As a solution, we advocate the use of semantic technologies to provide machine-interpretable data, thereby enabling advanced Smart Grid applications and processes.

Considering the requirements and in particular the layered architecture, one might notice strong parallels to the (Semantic) Web Stack – an adaptation of which would result in a layered and decentralised architecture (R1/R2). More precisely, we recommend an architecture as follows:

– *Data Access Layers* We advocate URIs for identification of participants (R3). We employ a TCP/IP stack with HTTP as transfer protocol for establishing a connection and accessing data (R4). However, standard Internet protocols are usually not adequate for low-power devices, due to their overhead from the various protocol headers. Thus, special protocols developed for low-power devices (e.g., sensors) may be adapted: e.g., a light-weight layered architecture such as IEEE 802.15.4 (physical and MAC layer), 6LoWPAN (internet layer, IPv6 version for IEEE 802.15.4 networks) or a single layer coupled with a middle-ware (for communication with TCP/IP networks), e.g., [3] (R4).
– *Data Representation Layers* To support a semantic understanding we advocate RDF(S) (if necessary extended with OWL features) to provide light-weight means for machine-interpretable data encoding (R5). Via Linked Data principles, data from different sources can be linked and thus integrated (R6).
– *Application Layers* Application and business logic can be represented via RIF (R7). We may use SPARQL as means to query RDF data and thereby allow the articulation of information needs (R8). Last, employing proof and trust mechanisms (together with rules), we can model constraints for the necessary data privacy and security (R9).

3 Open Challenges in the Smart Grid

In this section, we identify future research questions in a (Semantic) Smart Grid.

Challenge 1: Support Heterogeneous Participants (Data Access Layer, R4) Devices in the Smart Grid have a higher level of (technical) heterogeneity than within the Web. That is, we have to enable a flexible and light-weight (ad-hoc) integration of low-power devices (e.g., sensors and actuators) using low-level protocols with traditional information systems working on higher levels of abstraction. In fact, problems well-known within sensor networks (e.g., data uncertainty, vastness or integration) are aggravated in the Smart Grid, as we have various distributed, heterogeneous, low-power device networks (e.g., households) and various high-level applications (e.g., billing or energy consumption prediction) which depend on reliable data in real-time.

Challenge 2: Flexible Data Schema (Data Representation Layer, R5/R6) In the Semantic Web schema learning, schema design or schema alignment are well-known problems. However, in the Smart Grid there is a large number of different stakeholders (e.g., energy producers, grid operators or appliance manufacturers) having divisive backgrounds and goals. Thus, creating and enforcing a common data schema may be challenging. Additionally, in the Smart Grid we have little a-priori information about the participants and the data exchanged. For example, customers may add new or remove old appliances within their households or new service providers may participate in the markets. Added participants may contribute new kinds of data, while existing ones still except a certain data input. Thus, we need a very flexible data schema incorporating some fixed parts (modelling static aspects of the grid), while being easily expandable and (to some extend) adjustable.

Challenge 3: Large-scale Complex Event Processing (Application Layer, R7) The vast amount of data that comes from data sources within the grid has to be processed efficiently to enable smart behaviour. Billing and usage analysis can be done using conventional batch processing methods. However, the dynamic adaptation of the grid to the current situation (e.g., current energy consumption) requires real-time complex event processing on a very large scale. In particular, due to the data vastness and uncertainty (e.g., data from sensors), efficient and reliable event processing becomes very challenging. Note, in contrast to traditional Web scenarios, actions triggered in the Smart Grid have (possibly drastic) real-world effects (e.g., energy outages). Thus, we have a very low fault tolerance when making decisions.

Challenge 4: Privacy and Security (Application Layer, R9) Last, there is a strong need for privacy and security within the Smart Grid. Privacy concerns the data about individuals, e.g., information about premises, vehicles and appliances or energy consumption. Traditional access control mechanisms are helpful to block unwanted data access. However, there are many situations where initial data access is granted, but the subsequent data usage has to be restricted (e.g., restricted to few purposes or participants). Also, there may be regulations enforcing the publishing of specific data. Means for expressing usage restrictions and a technical enforcement such restrictions (e.g., at certain participants such as a metering provider) must be supported. Thus, e.g., work on WWW policies should be adapted to allow a privacy-aware grid. Further, the Smart Grid includes participants with very high security requirements (e.g., a clearinghouse or an energy provider). That is, a malicious access at such participants can have disastrous real-world effects. Thus, a communication architecture must provide strong means for securing high risk participants, while still allowing access to/from the remaining (open) grid.

4 Conclusion

In this paper, we have argued that open, royalty-free (Semantic) Web standards can provide the foundation for a Smart Grid communication architecture. Further, we listed the remaining challenges that stem from differences between the Smart Grid and the Web, i.e., support of (very) heterogeneous participants, a flexible schema, large-scale complex event processing and a strong need for privacy and security. In the future, we plan to extend our work by implementing the outlined architecture in our laboratory and conduct first field tests, thereby (on a step-by-step basis) addressing the outlined problems.

References

1. European Technology Platform - SmartGrids Vision and Strategy for Europes Electricity Networks of the Future. European Comission, 2006.
2. NIST Framework and Roadmap for Smart Grid Interoperability Standards. National Institute of Standards and Technology, 2010.
3. K. Aberer, M. Hauswirth, and A. Salehi. A middleware for fast and flexible sensor network deployment. In *VLDB '06: Proceedings of the 32nd international conference on Very large data bases*, pages 1199–1202. VLDB Endowment, 2006.