Data Life Cycles in Future Residential Multi-Commodity Energy Management Systems

Fabian Rigoll, Christian Gitte, and Hartmut Schmeck
Institute AIFB
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
{rigoll, gitte, schmeck}@kit.edu

***Abstract*—The number of residential devices that are capable of using more than one kind of energy carrier is increasing. Managing and efficiently utilizing such multi-commodity devices is a complex task. However, such approaches allow for more energy-efficient future residential environments. Therefore, suitable energy management systems are needed that support more than one energy carrier. In this paper, we apply a data life cycle analysis in combination with an economic role model to a prototype hybrid energy use case. The interconnections between the involved stakeholders as well as the properties of the data are analyzed. The results show that the combination of a systematic analysis and an economic role model is suitable for gaining a better understanding of complex multi-commodity processes. Facilitating the comprehension of hybrid energy use cases is a first step towards efficient and effective multi-commodity energy management systems.**

***Index Terms*—Data life cycle analysis, multi-commodity energy management, smart grid, energy market communication.**

I. Introduction

Hybrid or multi-commodity energy use cases are applications where more than one type of energy is used. When it comes to single-commodity energy management, most often, electricity only is considered. However, there are other energy carriers that are relatively wide spread: natural gas as well as thermal energy comprising both cooling and heating. Combining all or a combination of those in a residential energy management system (EMS) poses a great challenge. Control as well as optimization tasks are more complex than for single-commodity applications. Nonetheless, the utilization of micro-combined heat and power plants (µCHP) for example is on the rise. Many other use cases in which the customer has to decide which energy carrier to use are conceivable. Those applications show the need for an EMS that is capable of handling multiple energy carriers in one system.

In order to be able to create such systems, the underlying processes and data flows must be well understood. In this paper, we combine an economic role model and a data flow analysis method to gain better insights on multi-commodity processes. Results can support ICT companies, service providers, and researchers by explaining interdependencies between market parties. The strong interest in energy data has been underlined recently by the \$ 3.2 billion acquisition of *Nest Labs*, a start-up for innovative energy sensors and actors [3]. Meanwhile, the term Big Data is often mentioned by ICT companies if they talk about utilities industry, but often connected to the collection of (smart) metering data. However, the need to manage large amounts of data and messages is not only connected to smart metering. It is a direct result of liberalization in energy markets and more complex business-to-business and business-to-consumer processes. First applications of demand side management respectively demand response have started. Holistic approaches for multi-commodity energy management lead to more elaborate processes.

II. Related Work

The idea of scientific data life cycles is introduced in [1]. It addresses the need for a well-defined data management process for scientific data, especially if large data volumes are created, stored and analyzed continuously. The main concerns, identity and access management, meta data handling, monitoring, modeling and optimization in data intense scenarios, are connected to the technical implementation of data life cycles. As mentioned in the introduction, the energy data life cycles presented in this paper focus on the interconnections of data flows within a typical unbundled and liberalized energy market and not on scientific life cycles. Such data flows are dominated by energy market communication defined by regulators and federal associations representing acting industries. In [2], a market communication scenario of energy supplier change processes in Germany is investigated. Roles investigates are final consumer, retailer and metering operator. A simulation is presented for data exchange processes associated with the change of an electricity supplier today (monthly change possible) and in the future (idea of hourly change). The authors state that even the possibility for an hourly change with a very high willingness to change will lead to small sets of data compared to scientific cases investigated in [1]. Ireland's electricity market and data flows of power meter data are described in [4]. In [5] the electricity market of Flanders is introduced, a balance responsible party, retailers and DSO is included. The paper also shows another implementation of market partner communication based on an EDI subset for utilities. The relevance of future markets with a stronger role of the balance responsible party is clearly stated in several research projects, e.g. in MeRegio [6] or RegModHarz [7]. Concerning the handling of large data in energy, one can find several white papers from IT industry but only minor scientific discussion. Studies from Deutsche Telekom [9] and IBM [9] testify only several Terabytes of metering data per year and region. That is the case for resolutions of 15 minutes and central storage. More challenging applications of large scale energy data to be handled can be found on the trading level [cf. 12] and when looking at communication aspects [cf. 11]. Finally, we need to distinguish the data life cycle [1] approach from the information life cycle approach introduced by [13]. In contrast to information life cycles, which focus on the value of information retrieved by data, a data life cycle is focusing on the technical basis of the life span of data.

III. Data Life Cycle and Data Flow Analysis

Research in the project *Large Scale Data Management and Analysis* (LSDMA) focusses on the scientific Data Life Cycle (DLC) of different projects [1]. In the initial concept, the analysis covers the whole scientific DLC. In this paper, we use the idea of DLCs [1] and apply it to processes in hybrid energy scenarios, covering also economic aspects. In the following section we shall present the concept of a data life cycle analysis (DLCA) as well as the involved stakeholders. Besides, we show the results of such an analysis in a prototype multi-commodity use case.

A. Data Life Cycle and Data Life Cycle Analysis

The DLC model consists of six phases (see Fig. 1). In the first phase, acquisition, data are gathered, created, or measured to be used later on in the DLC. The newly created primary data are transmitted from their origin to a storage system in the second phase. The previously transmitted data are then stored, which is the third phase. Usually, most time is spent in the fourth phase, the data analysis. Primary or derived data are distributed to other systems in the fifth phase. Finally, in the sixth and last phase of one iteration, data are deleted.

A DLCA uses the aforementioned model of the DLC as an underlying scheme to gain a better understanding of the data and their flow in the system as well as the involved parties. For the DLCA, a process is divided into its separate phases and each phase is investigated individually. In the data acquisition phase the creation of the data is in the focus. Possible sources are simulations, experiments, sensors, other processes, and more. Furthermore, data formats, amount of data, and quality of data are investigated. Questions to be answered in the second phase, the transmission, comprise data origin and destiny, transmission protocols, bandwidth, formats, compression etc. The transmitted data need to be stored. Thus, it must be determined whether using databases or a file based system is appropriate. Data formats as well as location of storage also play a crucial role. Usually, the fourth phase, analysis, is the most important part of a DLC. Here, different methods and algorithms are used to derive information from the previously gathered data. A profound understanding of the involved processes is necessary to evaluate effectiveness and efficiency of the data analysis. Both primary and derived data might be of interest to other systems and stakeholders. Thus, in the fifth phase, data are distributed. To understand this step, information about involved stakeholders and quality as well as amount of data is necessary. The DLC is concluded with the deletion of primary and derived data. Triggers for deletion, the kind of data to be deleted, and the point of time when to delete are examined in this phase.

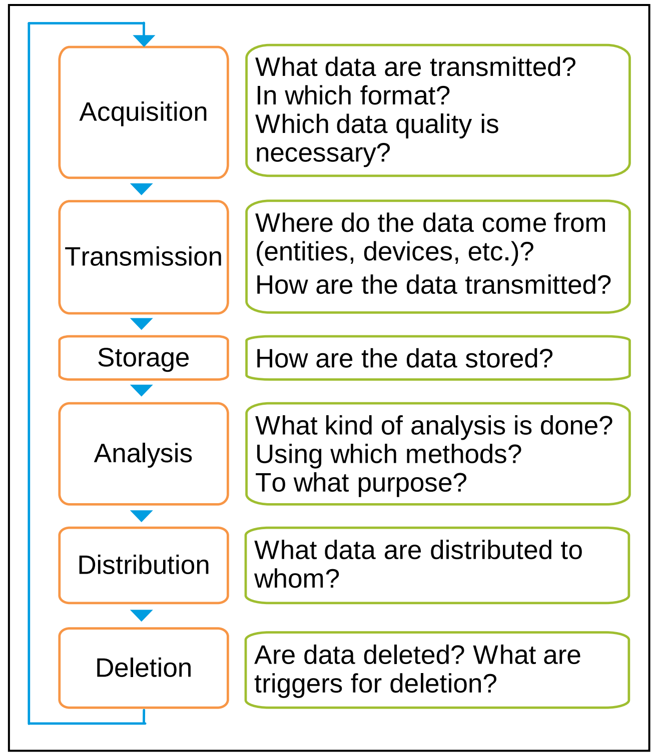

Figure 1. The six phases of a data life cycle and examples for the corresponding aspects to be analyzed.

In this paper, the basic idea of a DLCA serves as an inspiration for a different perspective. We combine an economic role model with the analysis. This allows for a more detailed view on the interconnections between different entities and data flows. Hence, during the distribution phase, data from one entity's DLC can be transferred to another entity's DLC, and thus, one DLC can be the source of data for another DLC. This leads not only to new connections between the different phases but also to connections between different DLCs. It is common that several DLCs form a closed-control-loop-like system in which data are finally fed back into the original DLC.

B. Stakeholders

Our simplified role model addresses four kinds of stakeholders and several specializations (see Fig. 2). The basic description is based on the ENTSO-E model [cf. 8]. We include fundamental roles of a typical liberalized and unbundled energy market, also applicable in divisions of fully integrated classic utility companies. Considered roles have particular interests in data (see Tab. 1).

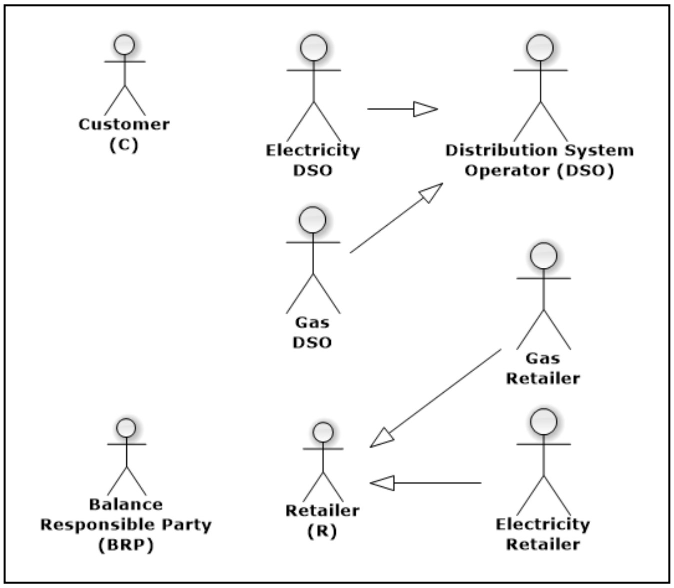

Figure 2. Role model utilized in data life cycle analysis.

	Customer	DSO	Retailer	BRP
Metering data	x	x	x	x
Billing data	x	x	x	
Operational data	x	x		
Contractual data	x		x	x

TABLE I. DATA INTEREST OF DIFFERENT STAKEHOLDERS.

1) Customer (or consumer): A (private or commercial) party that consumes electricity. Prosumers, consumers with small scale generation facilities, are included in this role, too. They are interested in metering data, billing data, and operational data of micro CHPs. Different kinds of energy, namely electricity, gas, and thermal energy, will result in different data.

2) Distribution System Operator (DSO): A party that operates one or more energy grids and is responsible for metering at customers' sites. In case of electricity-DSOs the low and medium voltage grid is addressed. In case of gas-DSOs the regional downstream infrastructure is addressed.

3) Retailer: A role responsible for gas and electricity retail. It has contracts with its consumers about delivering energy or buying energy produced at the prosumer level. The retailer is also responsible for invoicing a concerned party (billing agent function). Furthermore, retailers buy energy at the wholesale level. We think that it's reasonable to exclude existing regulations and incentives by the states to promote renewable energy and dispersed generation. As a consequence the retailer party is responsible to buy energy at the prosumer level in our model.

4) Balance Responsible Party (BRP): A party that has a contract providing financial security and identifying balance responsibility in the market. It is equivalent to "Program responsible party" in the Netherlands, equivalent to "Balance group manager" in Germany, and equivalent to "market agent" in Spain.

C. The Data Life Cycle Analysis

In the project LSDMA the data life cycle analyses focus on the technical level of different projects in many research areas. This effort results in solutions, work flows, and tools applicable to a variety of other related projects. It is our intention to adapt the method of DLCs to a multi-commodity smart grid scenario. Therefore, we expand the model of data life cycles. Different roles, for example the stakeholders mentioned above, are introduced. For each stakeholder, an individual DLC is defined and the interconnections between them are demonstrated. The combination of the DLCA and the role model results in a detailed view on the complex structure of a well-defined use case.

A very common hybrid or multi-commodity application in smart grid scenarios is the use of a micro combined heat and power plant (µCHP) with additional electrical heating. Such a device provides both heat and electrical energy. It does so by combusting fuel in an engine while driving a generator. A relatively widespread µCHP, the *Dachs* by the company *Senertec* [14], produces about 12 kW of thermal power and 5 kW of electrical power. It can thus be used in single households as well as in multi-family homes. Even though the Dachs and other µCHPs can be fueled with a variety of fuels such as gasoline, fuel oil, and natural gas, we restrict this use case to natural gas.

Using the µCHP efficiently requires a decision support system or a residential energy management system (EMS) which is capable of handling multi-commodity optimization problems. Multi-commodity use cases are of a much greater complexity than applications that make use of electricity or heat only. We aim to make multi-commodity use cases manageable by performing a DLCA combined with the role model. In the following, we describe the results of the analysis of a µCHP use case. Four different stakeholder roles are involved: a customer (C) with a µCHP, one distribution system operator (DSO) each for electricity and gas, one retailer (R) each for both electricity and gas, and a balance responsible party (BRP) for electricity only. They do not act independently but their processes are interconnected. Refer to Fig. 3 for an overview. An iteration of the standard data life cycle is shown on the left. The adjacent columns present a brief description of the respective DLC's phases itemized by the different stakeholders that are involved.

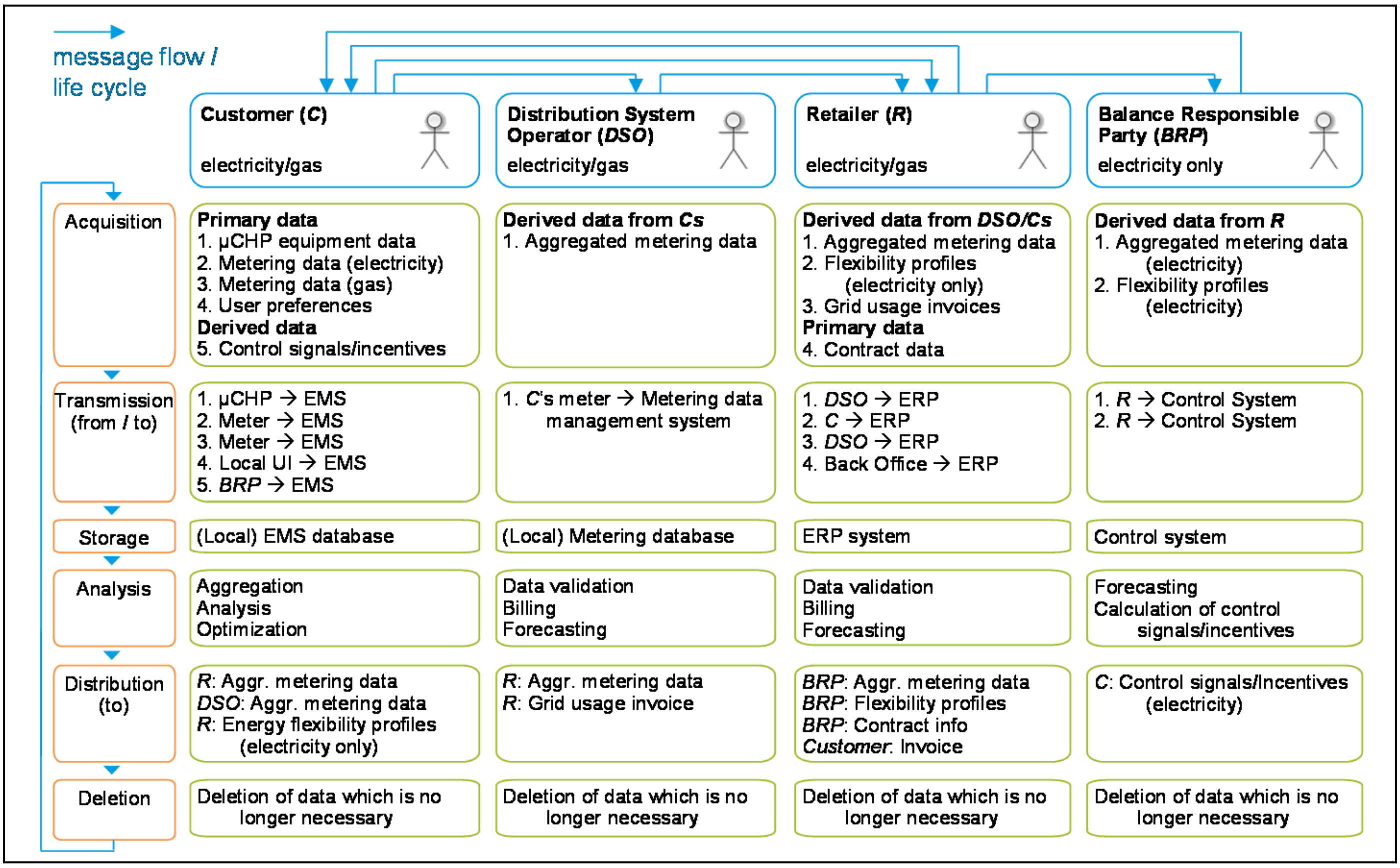

Figure 3. Interconnected data life cycles in a multi-commodity setup.

The customer's DLC could be considered the central DLC as it affects all other parties' DLCs. The µCHP is equipped with several sensors that provide information on its system state. Those data as well as metering data of both electricity and gas meters are primary data. Furthermore, user preferences are collected and control signals or incentives are received. All data are transferred from their respective origins to the customer's EMS. Usually, a local database is used as storage. Primary data are aggregated and derived values are calculated. The entirety of data is then used to analyze the current system state and optimize the use of the µCHP depending on different parameters and the input data. Some of the data are distributed to other stakeholders: The gas and electricity retailers as well as the gas and electricity DSOs receive aggregated metering data. Additionally, the electricity retailer receives flexibility profiles which contain information on load shifting and production flexibilities of the customer. The customer's EMS usually deletes primary data which are no longer necessary and keeps only derived data in its database.

Within the restricted view of this use case, the DSOs receive data from customers only. While electricity and gas DSOs might be different entities, their processes are of the same kind and thus, they are aggregated in our analysis. The aggregated metering data are transferred from the customers' systems to the DSOs' meter data management system. They are stored in a local database. As the data might be corrupt or incorrect, a data validation step is necessary. Consequently, the verified data are used for billing purposes. Obviously, a connection between customer information and metering data is necessary in order to associate both. The DSO can furthermore use data for forecasting. In the scope of this particular use case, data are sent to one other role only: the retailer. Both electricity and gas retailers receive aggregated metering data which originally come from the customers. Besides, they receive grid usage invoices. Data which is no longer necessary is deleted. The amount of data which can be deleted is rather small as aggregated metering data are not only required as proof for the billing processes but also as a basis for reliable forecasting. Therefore, they will be stored for several years.

The retailers' situation is comparable to the DSOs'. Within the focus of this use case, there are gas and electricity DSOs. Even though they might not be the same entity, they belong to the same role. Therefore, they can be aggregated. The retailers receive aggregated metering data, both from electricity and gas consumption. Flexibility profiles are received for electricity only. Additionally, grid usage invoices and contract data are acquired. The aggregated metering data are not received from the customers directly. Instead, they come from the DSOs. The customers send their electricity flexibility profiles. For gas, there are no flexibility profiles as the gas distribution grid itself offers a large buffer capacity and thus, short-term variation in demand does not have a great impact.

The DSO sends grid usage invoices to the retailer. Lastly, contract data from the back office are received. All data are transmitted to the retailers' enterprise resource planning systems (ERP). The ERP is also used as storage for the aforementioned data. Like the DSOs, the retailers are responsible for data validation, billing, and forecasting. Received data can be corrupt and thus must be checked. The billing process is for the customers, in contrast to the DSO's billing process which is for the retailer. Forecasting is done by taking all stored data into account and thus providing reliable prognoses of energy consumption. Data is distributed to two roles: the balance responsible party and the customer. An invoice is sent to the customer by the retailer, not by the DSO. Aggregated metering data, flexibility profiles, and contract information are sent to the balance responsible party. Deletion of data is comparable to the DSOs. Any data that is no longer necessary can be deleted. However, a great part of the stored data is needed as proof for billing and forecasting.

In our model, the balance responsible party (BRP) does not exist for gas as the gas grid has a rather large buffer capacity and balancing it is less difficult than balancing a power grid with strict frequency constraints. The BRP receives aggregated metering data from electricity consumption as well as electrical flexibility profiles. Even though both kinds of data are originally created by the customers, they are not received from them directly. Instead, they are transmitted to the retailers and from there to the BRP's control system where they are stored. In order to balance the grid, forecasting and calculation of price signals or incentives are done. As explained before, the frequency constraints in an electrical power grid are rather strict and thus, proper control is necessary. Even though many data are no longer directly necessary for the calculation of control signals, they can be used as a basis for better forecasting and thus improve the overall balancing process.

The described data life cycles and their interconnections are the result of the analysis of a limited use case. Furthermore, they are a limited view within a model. However, the µCHP use case is a prototypical application of multi-commodity energy use. Even though it might appear rather simple at first glance, it is quite complex and involves many different parties. Analyzing this use case facilitates understanding the complexity of multi-commodity applications.

IV. Conclusion

Worldwide, the on-going liberalization and unbundling of energy markets result in more complex market communication processes. In the future, changes in energy policy making will focus on increasing renewable energy usage. Most likely, dispersed generation will grow. The integration of such generation facilities requires more flexibility at the customer's site. First approaches of demand side management respectively demand response have already been implemented. New incentive systems will build on dependable and regular messaging between market parties, especially at customer's site. Higher energy efficiency can be reached if we consider multi-commodity energy management also in residential scenarios. Multi-commodity, also known as hybrid energy, optimization approaches based on residential energy management systems can provide more flexibility than traditional I-shaped solutions. However, in such a setup the communication and processing effort will continue to increase. The data life cycle analysis as well as the economic role model presented in this paper contributes to a better understanding of roles and processes involved in multi-commodity scenarios.

Based on the fundamentals of data life cycles, introduced in scientific data handling, we investigated a first real world scenario and presented a valuable contribution to further development in energy data handling. Besides, our investigation can help to identify other research topics. Especially, the distribution of data to several other market parties is of interest for privacy and data usage control approaches. Market-ready optimization solutions for residential energy management systems are not yet available and companies dealing with market communication support systems do not focus on data handling at customer's site so far. We believe that well known data life cycles based on national market implementation will be necessary if processes and data handling need to be efficient.

Acknowledgment

We gratefully acknowledge the financial support from the Helmholtz Association of German Research Centers, their project *Large Scale Data Management and Analysis* (LSDMA), and the *EIT ICT Labs* which provided the environment for this paper.

References

[1] C. C. Jung, M. Gasthuber, A. Giesler, M. Hardt, J. Meyer, F. Rigoll, K. Schwarz, R. Stotzka and A. Streit, "Optimization of data life cycles" in *Proc. 2013 CHEP of CHEP2013. In Press.*

[2] C. Gitte, N. Renkamp and H. Schmeck, "*Investigation of Data Exchange Processes for Electricity Supplier Change in German Smart Grid Scenarios*" unpublished, presented at the 1st Smarter Europe Conf. at E-World Energy & Water 2014, Essen, Germany, 2014.

[3] P. Olson, "Nest Gives Google Its Next Big Data Play: Energy" in *Forbes Website*. 13th January 14, Accessed 8th May 14 [Online]. Available http://www.forbes.com/sites/parmyolson/2014/01/13/nest-gives-google-its-next-big-data-play-energy/

[4] T. Bean, "*The Integration of Smart Meters Into Electrical Grids to Ensure Maximum Benefit for Consumers, Generators and Network Operators*". Masters Dissertation. Dublin Institute of Technology, 2010.

[5] T. van Craenenbroeck, B. de Wispelaere Vreg, "*Service level agreements and regulatory aspects of data communication between DGO's and suppliers*" 2005. CIRED 2005. 18th International Conference and Exhibition on Electricity Distribution, 6-9 June 2005.

[6] L. Hillemacher, K. Hufendiek, V. Bertsch, H. Wiechmann, J. Gratenau, P. Jochem, and W. Fichtner, "*Ein Rollenmodell zur Einbindung der Endkunden in eine smarte Energiewelt*" in ZfE – Zeitschrift für Energiewirtschaft 37(3), pp. 195-210, 2013.

[7] K. Rohrig, "*RegModHarz*". In Arnold Picot und Karl-Heinz Neumann (ed.), "*E-Energy*", Heidelberg: Springer, 2009, pp. 73–83.

[8] ENTSO-E, "*The Harmonized Electricity Market Role Model.*" Brussels: ENTSO-E, 2011.

[9] IBM, "*Managing Big Data for Smart Grids and Smart Meters*" May 2012

[10] Deutsche Telekom, "Smarter Energy Management Big Data für Energieversorger" in T-Systems *Website*. Accessed 12th May 14 [Online]. Available http://www.t-systems.de/loesungen/smarter-energy-management-big-data-fuer-energieversorger/1162956

[11] C. Herdt, and C. Schlenzig, "*Das Stromnetz wird schlau Chancen und Herausforderungen im Smart Grid für die IT-Branche*" in VKSI Magazin No. 5, November 2011.

[12] P. P. Tallon, "*Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost*" Computer , vol.46, no.6, pp.32,38, June 2013

[13] R. S. Taylor, "Value-Added Processes in the Information Life Cycle", in " Journal of the American Society for Information Science", vol. 33, issue 5, 1982, pp. 341-346.

[14] Senertec, "The Dachs - The Microcogeneration - Technical Data" Accessed 13th May 14 [Online]. Available http://www.derdachs.de/