Collaborative Environmental Data Management Framework for Microsoft Excel

Jörn Freiheit[1], Ramona Görner[2], Jost Becker[3], Frank Fuchs-Kittowski[4]

Abstract

In many companies Microsoft Excel is often used as a CEMIS for managing environmental data. In collaborative scenarios this single-user application often causes problems. In this article, the multi user framework Quexolver is presented, which supports the simultaneous processing of environmental data in Excel by multiple users. The multi-user capability is provided by introducing a role and rights management concept for Excel users and a client server architecture with a central database. The use of web services provided by the framework (server) enables the communication between the individual Excel applications (client) and the database. First experiences with Quexolver based on a case study from the environmental protection are also described.

1. Introduction

The need to detect and evaluate environmental information using IT-systems in companies is constantly increasing. One the one hand, this is due to in stricter laws, standards and guidelines [3]. On the other hand, companies want to access potential to increase efficiency [11] and enhance environmental awareness to prove their reputation among their customers [4]. Unlike tasks in companies which are absolutely necessary, such as human resources and accounting or technical-specific deeds, special software called “Corporate Environmental Management Information Systems” (CEMIS), made for the care and treatment of environmental data, is only used seldom applied [9, 11]. Instead, Microsoft Excel is often used for this task [7]. This is due to the simplicity in collecting and analyzing data, as well as the various methods available for presenting it.

However, the many advantages of Excel are offset by some disadvantages. The collection of data from several sources by multiple users leads to an effort that grows disproportionately with the number of users. For this purpose, either additional communication between users or a central data consolidation is needed. Excel does not support versioning, concurrent acquisition of data and has no role concept. When a spreadsheet is used by multiple users, the complex dependencies of data are often not recognizable and will be thereby in many cases destroyed.

This is where Quexolver enters. The framework was developed to use the advantages of Excel, particularly in the case of data analysis, and to support the collection of data by multiple users. Therefore, Quexolver extends the functionality of Excel by its ability to refer data from respectively transmitted data to a central database. Excel acts as a frontend, through which data can be entered, processed and presented in the usual ways. With Quexolver, necessary functionalities in a multi-user client-server infrastructure are provided: multi-user capabilities, network capability, versioning, roles and rights concept and consistent data storage in a relational database.

[1] HTW Berlin, Wilhelminenhofstr. 75a, 12459 Berlin, Germany, Informatics & Busines, Joern.Freiheit@htw-berlin.de

[2] HTW Berlin, Wilhelminenhofstr. 75a, 12459 Berlin, Germany, Environmental Informatics, Ramona.Goerner@htw-berlin.de

[3] binaere bauten gmbh, Oranienburger Str. 4, 10178 Berlin, Germany, j.becker@binaere-bauten.de

[4] HTW Berlin, Wilhelminenhofstr. 75 A, 12459 Berlin, Germany, Environmental Informatics, Frank.Fuchs-Kittowski@htw-berlin.de

This paper describes the architecture of Quexolver (chapter 3), explains the methodology to develop Excel applications through the use of Quexolver (chapter 4) and refers first experiences from a pilot scheme on corporate environment protection (chapter 5). The paper concludes with a summary and an outlook on further research (chapter 6).

2. CEMIS and Excel – Related work

In scientific literature the use of Excel as a CEMIS is rarely discussed. Some authors include Excel as an example for a CEMIS in their definition, e.g. [13], while others explicitly exclude Excel from being a CEMIS, e.g. [6], because of its missing collaborative capability. In practice, Microsoft Excel is in fact widely used as a CEMIS because of its availability, ease of use and capabilities for data processing, analysis and visualisation. Approximately 70% of the industry is using programs for spreadsheet analysis, such as MS Excel, to handle environmental information [1] and to support the various tasks of corporate environmental protection.

CEMIS very often provides respective interfaces to import from or export data two Excel [13]. These interfaces only support one-way communication, which does not allows collaborative work. Data is either imported or exported (e.g. for reporting or visualization), but not both at the same time. It is not possible to edit and update exported data and save changes in the original system.

The file-based single-user solution Excel supports rudimentary collaborative work scenarios by using the shared-workbook feature and placing the file on a network location. But only the user who opens the file first can edit and save data. Other users can only work with a copy of the original file, what leads to several copies of the original file with different data and the need to manually bring changes together.

To be able to use Excel with multiple users collaboratively, different approaches are now created, but each with their own specific disadvantages. Although w*eb-based file server* (file server) (e.g. SharePoint) has shares and rights concepts, they are file-based, and changes are not synchronized (without synchronization). *Synchronized network drives* (e.g. DropBox) synchronize user changes, but are also file - based and do not allow the collaborative editing of the document (without editing function). *Repository -based* approaches (e.g. Kivo) support the ability work collaboratively on a document with common version control, but do not allow simultaneous work on an Excel document in real time.

In contrast, so-called *online spreadsheets* (eg Google Drive, EditGrid, Zoho Docs, Microsoft Office Web App) allow collaborative spreadsheet processing in a web application. These allow collaborative work, but they do not reach the scope and speed of Excel for data processing and analysis. In addition (apart from Office Web App), users cannot use the familiar Excel environment but must instead learn to use a different system. Microsoft's own online spreadsheet "Excel Web Access" supports collaborative work in real time, but provides only a part of the original functionality of Excel.

The Business Intelligence Suite "Jedox base" of the company Jedox upgrades Excel with a MOLAP server. This allows editing Excel-worksheets simultaneously by several users in real time. Furthermore, the commercial suite "Jedox Premium" offers web capability, support for mobile devices and the possibility of integrating SAP.

Except for Jedox all solutions just discussed have only a part of the instruments that feature a multiuser system, such as network capability, versioning, a rights and role concept and consistent data storage in a relational database.

3. Architecture of Quexolver

The multi-user capability of Excel is made possible by the Quexolver client-server architecture, which is described in this chapter.

3.1. Main components

The architecture of the framework consists of three key components: database, backend and Excel-add-in (Windows/Mac). The communication between Excel and the database is accomplished by means of web services. The implementation of web services is triggered by the respective user. The available web services (with consideration of the roles and right concept) are listed by the Quexolver Excel add-in to the user.

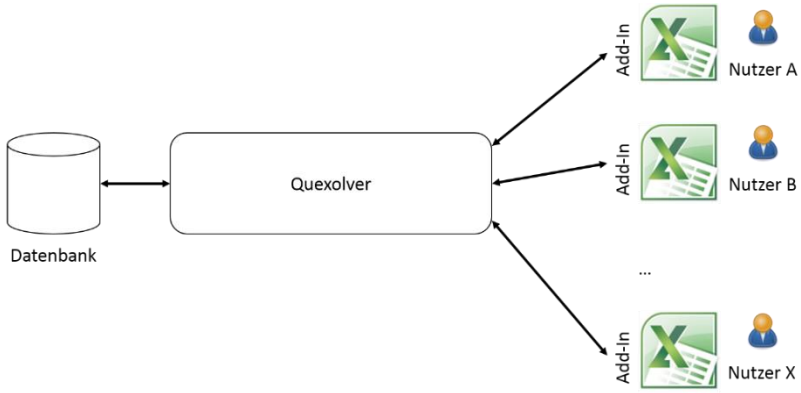

Figure 1: Excel multi user mode with Quexolver

The data in an Excel spreadsheet are stored in a central database. Communication between Excel and the database is carried out by web services. The execution of said web services is triggered by the respective user. Considering the respective role and rights, a Quexolver add-in (Windows / Mac) in Excel presents the executable web services to the user.

3.2. Interface to the database - Hibernate and DAO

Quexolver, however, is not just an implementation of web services. The interface between the web services programmed in Java and the database is realized through a data access object *(DAO)*[12]. By using DAO it is possible to separate the logic of the program from data storage. This means that the used database system (DBS) is exchangeable. The recent experiences (see chapter 4) with Quexolver are achieved using the open source DBS PostgreSQL.

If a certain web service is triggered by the user (e.g. "show me the values of all environmental indicators from the year 2012"), a request is sent to the database. The database system provides the query results as a table. For the further processing of the data by the Java based web service, the result table has to be converted into objects (Plain Old Java Objects - POJOs). This conversion is done automatically in Quexolver by Hibernate. Hibernate is a persistence framework program written in Java which converts and returns objects in data tables and vice versa (Object-relational mapping, ORM) realized [2].

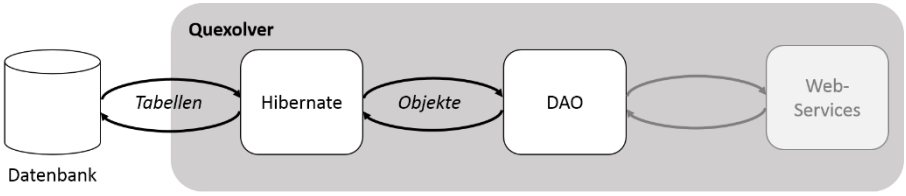

Figure 2: Interfaces between database and web services (Hibernate and DAO)

Hibernate does not only realizes the mapping of the data tables to objects. It also serves as a so-called persistence layer and implements the persistence of processed objects. Data to be stored from Excel into the database is persistent from the moment on once it is processed by Hibernate.

3.3. Data processing between Excel and DAO

At first, apart from the multi-user mode, the list of objects in the DAO interface must be associated with information about concerning at which point they shall be presented in an Excel worksheet. The objects are enhanced with the corresponding cell information[5]. The so-called converter, another component of Quexolver, is responsible for this operation. Each web service interacts with a converter. In the direction of the data flow from the database to Excel, cell information is added to the objects as described. In the opposite direction of data flow from Excel to the database (data storage), the cell information is removed from the objects before they are stored for reuse.

Another Quexolver component is necessary due to the simultaneous processing of data by multiple users. It is the so-called merger. The merger compares the local copies of data in the Excel spreadsheet with the objects stored in the database. There are two ways of changing data. First, the user may want to save new data that he has entered via Excel in the database. The merger checks, if data have been added, modified or deleted and refers them back to the web service, which can now save the updated dataset. This ensures that not all data is stored, but only the added, modified or deleted data.

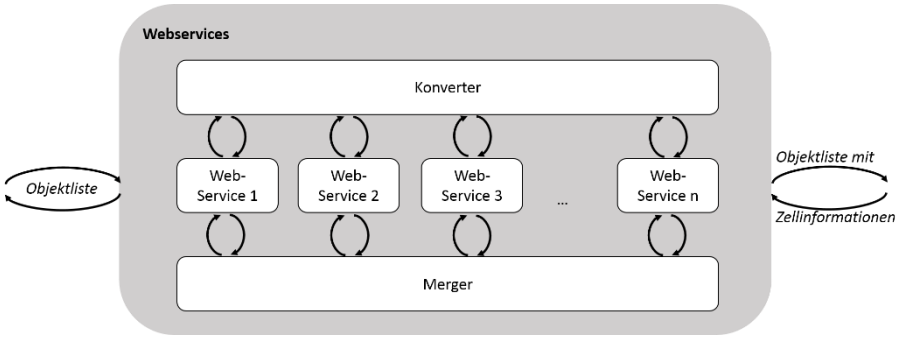

Figure 3: The use of converter and merger by web services

The second option is the change of data by another user, who works at the same time on the same data. This is realized by Optimistic Locking (Optimistic Concurrency) [10] realized. It is preferably used in applications, where the data is mostly accessed in read-only mode. The user who makes the first changes to the data record is privileged. If another user changes the same record, he will receive a notification of the change of data.

4. Configuration and use of the multi-user operation

For the single user, there are almost no changes in the daily work routine caused by the use of Quexolver, as he can use the well-known user interface of Excel and its full capabilities as usual.

[5] A cell is a place for data in an Excel worksheet.

Additionally, there is the new advantage due to the possibility of simultaneous editing of the same worksheets by multiple users. Time-consuming, error-prone data consolidation is thus avoided. The following describes the procedure for simultaneous editing of the same data by multiple users and how to adapt Quexolver to the respective business requirements.

4.1. Quexolver in multi-user mode

Each user logs in to the system with his name and password. The login screen is part of the Quexolver add-ins for Excel and visible in the right column within the Excel user interface. In Quexolver a role is defined for each user. Each role is assigned with certain rights, which grant access only to selected web services. These web services appear similar to a list of links on a web page at the right side of the Excel interface (see Figure 4). The user can execute the selected web service by simply clicking on it.

The execution of a web services causes the appearance of one or more worksheets with data in Excel. The user can now enter additional data, change or delete it, or perform analyses and evaluations, thereby using familiar Excel environment.

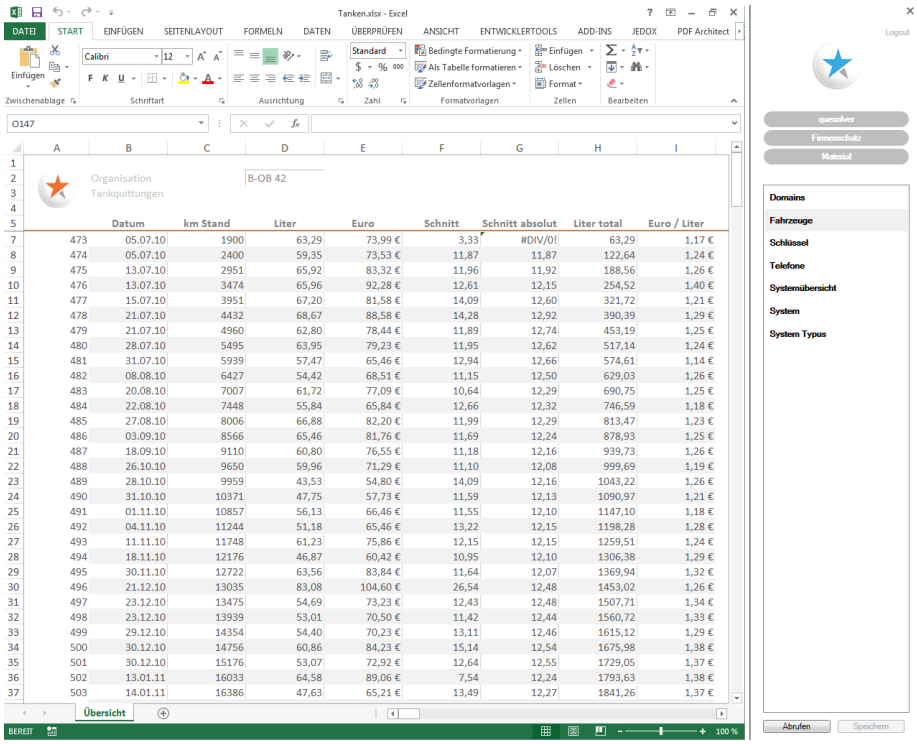

Figure 4: User Interface (Quexolver add-in right column)

If data is entered, changed, or deleted by user A, these changes appear in real-time in the worksheet of any user currently working on the same data user B. By saving changes, these are persistent and active for all other users.

The Merger ensures that effective changes are only incremental. This means that only the changes (inserts, updates, deletes) are stored and never all the work. This provides the possibility of a simultaneous processing of the same worksheets by multiple users without losing any data due to overwriting. For example, if data in the same worksheet from multiple data sources (that means by multiple users) is inserted into the database, the entire data is stored successively through the users's saving. Changes are visible even in real-time for all users during processing.

The use of Quexolver is characterized by two features: first, the user remains in the familiar environment. The addition of the login screen and the list of usable web services that are executed by a simple mouse click to the well-known Excel interface are intuitive to use and do not require any training. Second, the concurrent recording of data by multiple users is possible without requiring additional work processes. Instead, it eliminates even the manually merging of same Excel files. This is, unfortunately, in larger institutions with multiple data sources is currently still common and carries a high error rate during data acquisition.

4.2. Configuration of Quexolver

The configuration of Quexolver for a new application is possible in various ways. Usually, the starting point is an existing Excel file, which is either copied and sent to different users to be filled with data and later manually consolidated via copy and paste into a single file or sent from user to user to collect the data. To configure Quexolver, this Excel template is used to create a so-called entity-relationship model of the central database. By means of this model and generative software development techniques the Hibernate persistence layer and database are automatically generated.

Now the data in the database is mapped to the objects (Plain Old Java Object - POJO) and can be forwarded by the generic DAO interface to the web service for further processing (Data Access Object , see chapter 2.2).In the converter, another mapping of data to the cells, in which they should appear in the Excel worksheet is stored. Moreover, these web services are implemented, which accomplish the processing of the data. Therefore, typical web services are used. This includes querying a full data table from the database into an Excel worksheet, storing of changed or added data, as well as filtering and combining different data attributes in preparation for further data analysis.

An extension to the usual work with Excel is the combination of data entries with documents. This is useful, for example, if invoice data is stored with the corresponding invoice document. The documents are also saved in a database. Clicking on a date that is linked to a document corresponds to a web service that loads the corresponding documents from the database and opens them.

5. Application of Quexolver

Quexolver has already been used prototypically in different professional environments. An example of this application in the operational environmental protection is described in the following.

5.1. Case study: Extension of an existing Quexolver solution for the integration of environmental aspects

The existing ERP system of a company in the service industry should be extended to collect the relevant environmental information due to the introduction of a voluntary environmental management system. Since the existing solution is already a Quexolver solution, its data model must be extended. A particularly relevant indicator for the environmental performance of the company's environmental aspects is the consumed fuel by business trips using the company's own motor vehicle and the resultant CO_2 emissions. In the first step, the derived environmental indicator "annual mileage" has been implemented.

To include the environmental information, the existing data model has been extended. For this, the required attributes have been added to the respective entities. On the functional level, the input of the data was implemented for specific roles, which led to consistent data management. This ensured, on the user level, that only users with the appropriate qualifications could enter and change master and environmental data. The environmental data can be viewed by all users. For the company’s employees, this shall lead to an increased awareness about the environmental impact.

The environmental information was processed along with the master and vehicle-specific data, accumulated for a year and, furthermore, over all years. The comprehensibility of the information from the aggregated annual values to the individual data of the relevant vehicle is ensured via drill-down. In terms of data consistency, missing or invalid data is communicated to the user at any level and is also comprehensible via drill-down to the (environmental) master data.

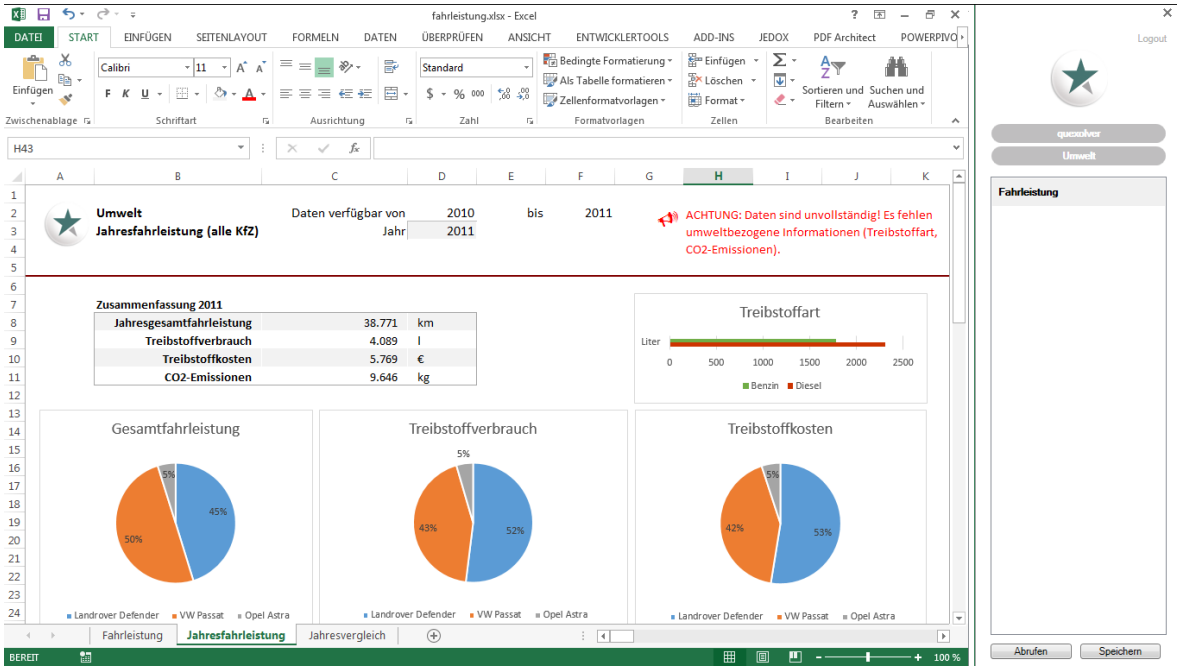

Figure 5: Excel front-end for visualization of the environmental indicator "annual mileage"

6. Summary and Outlook

Despite intensive research so far, continuous adoption and diffusion of CEMIS could not be achieved in practice [11]. Therefore, more recent research does not deal with universal CEMIS solutions and the general availability of CEMIS, but often with the analysis of existing CEMIS solutions in organizations [5, 14]. Thus, the specific design challenges and design factors can be more focused. In the context of corporate environmental protection, the research object of this paper, Microsoft Excel was the most frequently used standard software Microsoft Excel. This makes it possible to analyse and design the specific design requirements and factors of a widely adopted and diffused artifact.

Excel is a standard part of the Microsoft Office software used in most businesses and institutions. It is used in various applications as software for both the resource and project planning. The many advantages of Excel are offset by the disadvantages associated with the lack of multi-user capability. The experience with the presented tool Quexolver has shown, that addition of the multi-user capability and maintaining the familiar Excel environment at the same time is a promising approach to ensure a rapid and easy integration of the new tool in everyday work processes.

In particular, for processes, where many decentralized users collect data from various sources, the introduction of Quexolver leads to a simplification of these applications. Due to better automation of the consolidation and aggregation of data from various sources, the error rate is reduced significantly by Quexolver. The high flexibility of Excel as well as the service orientation of Quexolver, allows one to simplify and automate very specific technical processes.

Work is in progress to generalize the configuration of Quexolver for new technical applications and, thus, to simplify the framework. In another case study at the Federal Environment Agency of

Germany, further experience is gained with Quexolver in the scope of environmental management in order to incorporate further improvements of the framework.

7. Acknowledgements

This paper was created as part of the project "Quexolver Green". The project is funded by the European Union (European Social Fund - grant no 2013000471).

8. Bibliography

[1] Allam, N., & Juncker, H.; CEMIS Software available on the German Market. *EnviroInfo 2011: Innovations in Sharing Environmental Observations and Information.* Shaker Verlag, Aachen.

[2] Bauer, C., & King, G.; *Java-Persitence mit Hibernate.* Hanser-Verlag, 2007.

[3] Butler, T., & McGovern, D.; A conceptual model and IS framework for the design and adoption of environmental compliance management systems. *Information Systems Frontiers, 14*(2), S. 221-235, 2011.

[4] Gräuler, M., & Teuteberg, F.; Greenwashing in Online Marketing – Investigating Trust-Building Factors Influencing Greenwashing Detection. *Proceedings der Multikonferenz Wirtschaftsinformatik (MKWI) 2014.* Paderborn, 2014.

[5] Ijab, M., Molla, A., & Cooper, V.; Green Information Systems (Green IS) Practice in Organisation: Tracing its Emergence and Recurrent Use. *18th Americas Conference on Information Systems (AMCIS).* Seattle: Association for Information Systems (AIS), 2012.

[6] Juncker, H., & Lang, C.; Problems of Practical CEMIS (Corporate Environment Information Management Systems) Usage. *Information Technologies in Environmental Engineering: New Trends and Challenges* (S. 153-163). Springer, Berlin u.a.,: Paulina Golinska, Marek Fertsch, Jorge Marx Gómez, 2011.

[7] Junker, H.; Die Beliebigkeit betrieblicher Umweltinformationssysteme. *EnviroInfo 2010 – Integration of Environmental Information in Europe*, S. 232-247.

[8] Junker, H.; BUIS für den produktionsintegrierten Umweltschutz – Wunsch oder Wirklichkeit? *IT-gestütztes Ressourcen- und Energiemanagement, Konferenzband zu den 5. BUIS-Tagen , Berlin-Heidelberg*, S. 243, 2013.

[9] Junker, H., Marx Gomez, J., & Lang, C.; Betriebliche Umweltinformationssysteme. *Multikonferenz Wirtschaftsinformatik (MKWI) 2010*, (S. 1045-1062). Göttingen.

[10] Kung, H., & Robinson, J.; On optimistic methods for concurrency control. *ACM Transactions on Database Systems (TODS), 6*(2), 213-226, 1981.

[11] Melville, N.; Information systems innovation for environmental sustainability. *MIS Quarterly, 34*(1), S. 1-21, 2010.

[12] Oracle; *Core J2EE Patterns - Data Access Object*. Von http://www.oracle.com/technetwork/java/dataaccessobject-138824.html abgerufen, 13.06.2014.

[13] Perl, E.; *Implementierung von Umweltinformationssystemen - Industrieller Umweltschutz und die Kommunikation von Umweltinformationen in Unternehmen und in Netzwerken.* Wiesbaden: Deutscher Universitätsverlag, 2006.

[14] vom Brocke, J., & Seidel, S.; Environmental Sustainability in Design Science Research: Direct and Indirect Effects of Design Artifacts. *7th International Conference DESRIST 2012. Lecture Notes in Computer Science*, S. 294-308. Las Vegas: Springer.