Project Report

Data Management in Collaborative Interdisciplinary Research Projects—Conclusions from the Digitalization of Research in Sustainable Manufacturing

Wei Min Wang [1],*, Tobias Göpfert [2] and Rainer Stark [1,3]

[1] Department for Machine Tools and Factory Management, Technische Universität Berlin, Pascalstr. 8-9, 10587 Berlin, Germany; rainer.stark@tu-berlin.de

[2] Berlin School of Library and Information Science, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany; t_goepfert@gmx.net

[3] Department for Virtual Product Creation, Fraunhofer-Institute for Production Systems and Design Technology, Pascalstr. 8-9, 10587 Berlin, Germany; rainer.stark@ipk.fraunhofer.de

* Correspondence: w.wang@tu-berlin.de; Tel.: +49-30-39006-272

Abstract: As research topics become increasingly complex, large scale interdisciplinary research projects are commonly established to foster cross-disciplinary cooperation and to utilize potential synergies. In the case of the Collaborative Research Center (CRC) 1026, 19 individual projects from different disciplines are brought together to investigate perspectives and solutions for sustainable manufacturing. Beside overheads regarding the coordination of activities and communication, such interdisciplinary projects are also facing challengs regarding data management. For exchange and combination of research results, data from individual projects have to be stored systematically, categorized, and linked according to the logical interrelations of the involved disciplinary knowledge domains. In the CRC 1026, the project for information infrastructure observed and analysed collaboration practices and developed IT-supported solutions to facilitate and foster research collaboration. Data management measures in this period were mainly focused on building a shared conceptual framework, and the organization of task related data. For the former aspect, an ontology basesd apporach was developed and prototypically implemented. For the latter aspect, a message board integrated task management system was developed and applied.

Keywords: collaborative research; interdisciplinary research; data management; ontology; semantic matching

1. Introduction

Modern science is characterized by increasingly complex research fields that require the combination of knowledge from different disciplines [1]. Research on sustainable manufacturing, for instance, depends on conceptual integration across domains such as product development, manufacturing technologies, sustainability engineering, mathematics, education, and economy. As a consequence, large scale interdisciplinary research projects have become a regular instrument of universities and funding agencies to address complex research topics [2,3]. In Germany, for example, the German Science Foundation (DFG) funds so-called Collaborative Research Centers (CRCs) to involve multiple institutions with different disciplinary backgrounds for collaborative research.

Among policy makers and funding agencies it is widely assumed that interdisciplinary research collaboration yields great potentials for cross-disciplinary fertilization and allows for

exploitation of synergies. Hence, it is expected to generate more innovative solutions than classical disciplinary research. The complexity of the collaboration process and the entailed overhead are often underestimated by contrast [4]. The integration of different disciplinary perspectives requires an intense learning process to overcome cognitive distances between different disciplines [5]. This process is usually time-consuming and demands extensive mediation and coordination by experienced personnel [6]. At the beginning a shared vision has to be constructed which accommodates the different perspectives and determines their contributions for reaching a common goal. Based on such thematic framework, knowledge can be purposefully exchanged, combined, and restructured to form a unifying conceptual framework defining the collaborative research field. The conceptual framework in turn provides guidelines for individual and collaborative research activities [5]. Additional overhead for interdisciplinary collaboration can occur in terms of time, financial issues, or management efforts resulting from collaboration related phenomena on organizational, cultural, and individual levels [4,7]. On the organizational level, geographic distances may hinder collaboration, as time has to be spent for travel from one location to another. Institutional organization structure can affect autonomy of teams and their ability to collaborate in an efficient and self-determined way. Working culture—as a shared set of attitudes, values, and beliefs—may influence behavior of researchers in terms of openness in communication or willingness to share knowledge. On the individual level, personality and cultural background may determine personal motivation for collaboration and influence perception and evaluation of collaboration incentives. Environmental conditions such as funding schemes and reward systems may also interfere with the willingness of institutions and individuals to engage in collaborative research activities. Thus, interdisciplinary research means to orchestrate unique personalities from multiple institutions, diverse working cultures, and organizational structures in a shared environment, and steering their efforts through an evolutionary process towards a common goal [5,8,9].

One aspect that is gaining importance in context of interdisciplinary research collaboration is the management of research data. This development is partly owed to the challenges resulting from the complexity of interdisciplinary research, and partly to the rapidly increasing amount of research data [10]. Due to disciplinary differences, there is no consistent understanding of what can be regarded as research data [11]. Following a definition of the DFG, research data can be any kind of "digital and electronic data that are generated in the course of a scientific endeavor, e.g., through literature review, experimentation, measurements, surveys or interviews". This definition includes primary data—*i.e.*, data resulting directly from a data collection activity (e.g., measurement protocols of the energy consumption of particular machine)—as well as secondary data, which are derived from those (e.g., efficiency ratings of the machine based on the measurement protocols).

The goal of research data management is to systematically collect and archive data in order to make them available for further usage [12]. From a traditional perspective, data management was seen as an activity that only becomes relevant in a very late stage of a creative process. Consequently, it is focused on more finalized artefacts such as research reports and publications. This notion of data management led to deficient measures to preserve data which are generated and processed in the course of research activities, and even to the loss of primary data [13]. Still, there is a rising awareness for the importance of primary data to ensure reusability and traceability of research results, whereby the management of research data has become an important topic within national and international research communities [10].

Data management in the context of interdisciplinary research collaboration is facing similar challenges as the collaboration process itself. Despite the increasing need for researchers to access data from other disciplines, data management is still mainly disciplinarily characterized [10]. Each discipline has its own metadata sets; disciplinary institutions may have their own organizational and technical systems, and even their own understandings of what actually are "research data" [14]. Requirements regarding data management are mostly unique to the collaboration setup and evolve over the different phases of the project. During the initial phase of interdisciplinary projects it is

more important to provide less formalized tools and methods allowing for easy sharing of data and stimulating internal knowledge exchanges. In later phases, when common goals have been defined and collaboration practices constituted, means for consolidation and standardized structuring of generated data are needed (e.g., standard metadata scheme). Data relevant for, and resulting from, the collaboration process have not only to be stored and distributed systematically, but also categorized and linked according to the semantic interrelations of the involved disciplinary knowledge domains [15]. Hence, data management in interdisciplinary collaborations means to overcome domain specific boundaries on different levels, and to constantly raise requirements and provide appropriate measures to support researcher in all phases of the collaboration process. To address the challenges related to the management of research data and distributed research activities, funding agencies included various kinds of mechanisms in their funding policies [10,12]. Within the CRCs, the DFG funds dedicated service projects for information infrastructure (INF) to develop IT-supported solutions to facilitate research collaboration and to fulfill data management tasks [12].

In this article, experiences and results regarding the activities of the INF project of the DFG funded CRC 1026—Sustainable Manufacturing will be presented and discussed. During the first funding period of the CRC 1026 (from 2012 to 2015) the INF project developed methodological approaches and IT-tools to meet specific requirements of the CRC 1026. In the course of this funding period, INF observed and analyzed collaboration practices to constantly adapt solutions to the evolving requirements of a nascent research field. In section two, the characteristics of the CRC 1026 and its specific requirements and challenges will be introduced, and respective requirements for IT-support derived. In section three, selected results and findings of the INF project will be presented. In section four, the results and findings will be discussed.

2. Characteristics of the CRC 1026—Sustainable Manufacturing

2.1. Mission and Structure of the CRC 1026

The core mission of the CRC 1026 is to develop solutions for sustainable global value creation by combining knowledge from different disciplines. Sustainable manufacturing provides a great lever to realize more sustainable development of industry in the future, and has become a major topic in society and politics. One reason why the manufacturing sector is especially interesting in the context of sustainable development is its role as an essential part of the industry sector, which in turn is a major stakeholder in many areas of human living. On the European level, for instance, the industry sector employs 17% of the workforce [16], represents 26% of the final energy consumption [17], and emits 28.5% of the greenhouse gases (GHG) [18]. Manufacturing not only influences sustainability aspects through its direct impacts (e.g., resource consumption, labor conditions, GHG emissions, *etc.*), but also indirectly by determining the resource consumption of products over their entire lifecycle [19,20]. Furthermore, manufacturing is also a crucial sector for developing countries [21].

The complexity of this research field stems from the fact that aspects of sustainable development and manufacturing have to be combined. On the one hand, there is the concept of sustainable development. Its meaning as defined by the Brundtland Commission in 1987 has been modified ever since by various national and international organizations to be adjusted to changes of surrounding conditions or specific purposes [22,23]. However, most of the current definitions of sustainable development embody the interplay of the three dimensions of economy, environment, and society, as well as its resulting implications over time [24]. Manufacturing on the other hand, comprises the design and operation of physical processes (e.g., machining), as well as corresponding overhead processes (e.g., generation of operating resources) and organizational process throughout the entire product lifecycle (e.g., product development, distribution, product use, and end-of-life). Hence, considering sustainable development in the context of manufacturing means to engage in value creation processes along highly complex networks with regard to economic, environmental, and societal aspects and dynamic changes over time.

From the CRC1026's perspective, sustainable manufacturing is defined as the creation of manufactured products that, in fulfilling their functionality over their entire lifecycle, cause a sustainable impact on the environment (nature and human) while delivering economic value. The involved researchers from the four major disciplines of manufacturing, environmental engineering, mathematics, and economics are organized in 17 individual projects and three project areas (A, B and C). The project area A investigates strategic aspects of sustainable manufacturing to provide a broad systemic reference frame for effectively implementing strategies of sustainable value creation. Project Area B is technology oriented, and focuses on manufacturing research and development of an appropriate methodology to integrate elements of manufacturing technology in sustainable value creation concepts. In project area C, the two perspectives of the project areas A and B are merged. The main research goal is to develop methods and tools for learning and teaching and thereby enhance the productivity in conveying the challenge of sustainability to a global audience. Additionally, two cross-sectional projects—*i.e.*, INF and Public Awareness (PA)—support the research activities in terms of providing information infrastructure and a means of public communication and transfer of knowledge produced in the CRC 1026.

2.2. Expected Outcomes and Data Formats

Based on the setup of the CRC, it was assumed that research activities in and across the individual projects will result in heterogeneous outcomes, such as software tools, methods, models, algorithms, and raw data. Besides the publication of these results in scientific papers, other forms of documentation and presentation had to be considered, such as measurement protocols, data tables, code packages, or diverse types of models (e.g., business models, 3D models, finite element method (FEM) simulation models) based on different modeling languages (e.g., Unified Modeling Language (UML), Business Process Model and Notation (BPMN)) and produced by different kinds of tools (e.g., Computer aided design (CAD), process modeler). To evaluate general tendencies of prospective results and their respective formats, INF conducted brief interviews with researchers from twelve individual projects at the beginning of the CRC in 2012. These interviews were based on a structured questionnaire consisting of four open questions, in which the researchers were asked to describe their expected results and in which way they will be available. Multiple answers were allowed.

In the project area A, results were mainly expected to be in the form of reports, tables, and databases as surrounding field scenarios, sustainability indicator sets, lifecycle assessment methods, mathematical, and macroeconomic causal models are investigated. In the case of the mathematics related projects, optimization algorithms and respective software were also expected as results. In the technology focused project area B, raw data from the measurement of machining processes (e.g., drilling, welding) and simulation data (e.g., from FEM simulations) were anticipated, as well as methods and IT-tools supporting sustainable product design. Furthermore, the interviewed researchers also expected that digital models (e.g., CAD models) of machines and machine parts will be generated in that project area, as well as physical prototypes (e.g., of machine components). In project area C, results from game theoretical experiments, business assessments, and case studies on enterprise sustainability performance were presumed to be processed and documented in reports and papers. Furthermore, software tools for tracking, analyzing, and assessing the ergonomics of working processes with accompanying physical demonstrators should be developed. In this regard, visual data (e.g., video sequences) from motion tracking and capturing tools were also anticipated.

With respect to the expected formats of the results, Figure 1 shows that most researchers planned to document their results in the form of publications, whereas raw or primary data were only mentioned twice. Even in the technology oriented B area, only one project mentioned that there would be measurement data to be considered as a research result.
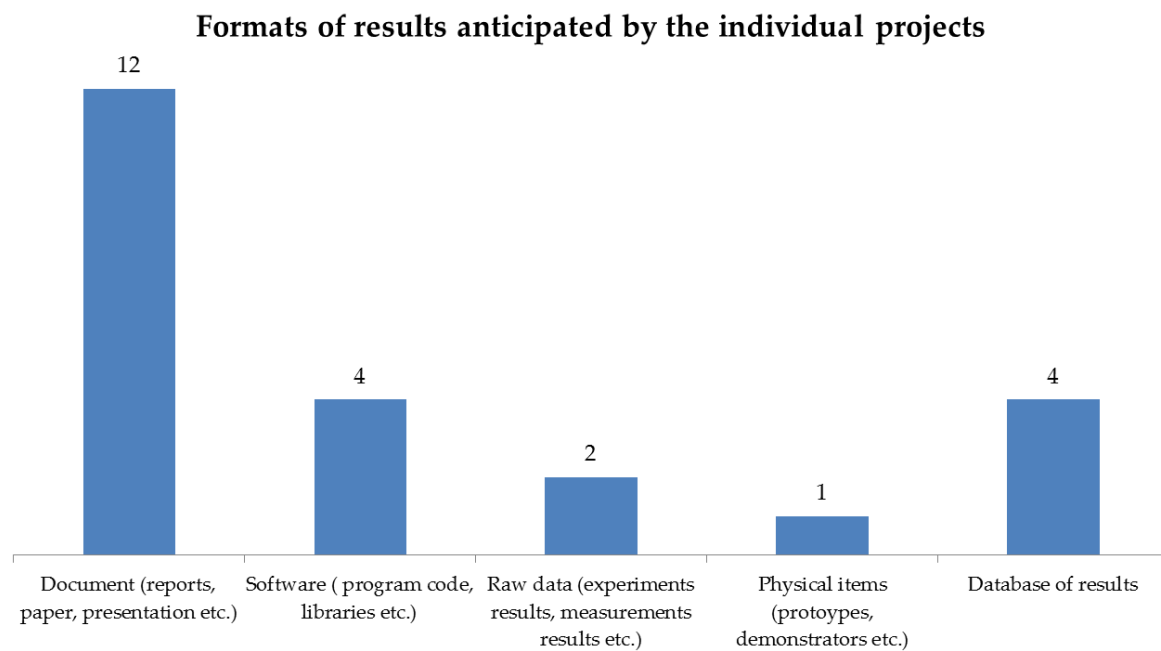
Formats of results anticipated by the individual projects

Figure 1. Excerpt of survey results: most results were planned to be published as scientific papers.

2.3. Definition of User Requirements for IT Support

To create novel solutions for sustainable manufacturing, research results have to be exchanged within and across the individual projects in all three project areas, and knowledge must be combined and restructured. To define requirements for an appropriate IT-support for such collaboration, exploratory talks and personal interviews were conducted in 2010, during the preparation phase of the CRC 1026. The exploratory talks were led between researchers from the individual projects in common workshops, and aimed primarily at identifying a first set of potential collaboration fields and interfaces (see Figure 2). Based on these prospective interfaces, fundamental requirements for respective IT-support could be derived. For example, the capability to handle diverse types of models is necessary as multiple projects are planning to engage in collaborative modelling processes. Moreover, a web-based content management system is needed to enable the exchange of data for collaborative activities of geographically distributed research partners. Thereon, INF prepared a set of 40 distinct features suitable for supporting collaborative research. Following the exploratory talks, semi-structured personal interviews were conducted with 16 designated project leaders in August 2010 to collect more specific user requirements for the development of appropriate IT-tools. In addition to these interviews, the project leaders were asked to fill out an online survey to evaluate and prioritize the list of features identified by INF. For this survey an own-hosted LimeSurvey system (limesurvey.org) was utilized. In the questionnaire the participants were ask to evaluate the relevance of the 40 identified features on a four-point scale, where "4" equals "very important" and "1" equals "not important at all". An option for "I don't know" was also available and assigned with the numeric value of "0".Ten of the prospective 16 project leaders participated in this survey, and the features were prioritized according to their averaged ratings (see Table 1).
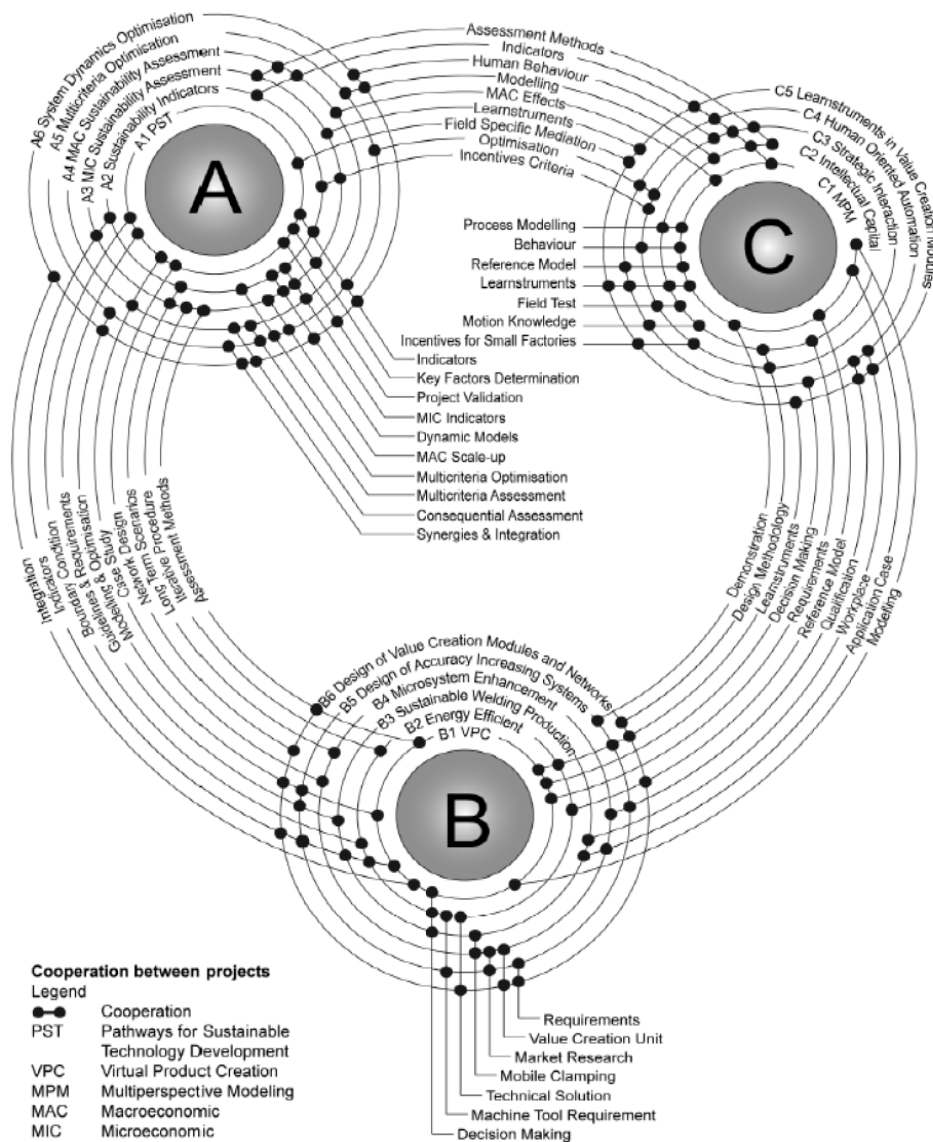
Figure 2. Potential interfaces between individual projects of the CRC 102.

Table 1. Functional features identified by INF and ranked according to user feedback.

No.	Average Rating	Feature Name	Description
1	3.24	Documents management	Online workspace for sharing documents
2	3.23	Wiki	Common wiki for a shared knowledge base
3	3.18	Mailing lists	E-Mail distribution list for automated notifications
4	3.18	Offline Access	"Dropbox"-like functionality for offline synchronizing of files
5	3.07	Ontology	Tool to create formalized representation of shared knowledge
6	3.03	Document comparison	Feature to compare different document versions
7	3.03	Web content management	Feature to present content on a website
8	2.93	Yellow Pages	Register of CRC researchers and their skills
9	2.87	Literature management	Feature to manage publications
10	2.87	Document versioning	Automated version control feature
11	2.85	Mind mapping	Feature allowing for collaborative creation of mind maps

Table 1. *Cont*.

No.	Average Rating	Feature Name	Description
12	2.8	Semantic Web	Capabilities for enriching data on the platform with semantic information
13	2.73	Tagging	Possibility of adding keywords to files, discussion threads *etc.*
14	2.67	Polls	Simple and short surveys
15	2.58	Mobile access	Access to the platform via mobile devices
16	2.57	Task management	Management of CRC internal tasks
17	2.57	Access control	Control of access privileges to platform contents
18	2.57	Calendar	Shared calendar
19	2.51	Management Dashboard	Feature to summarize project progress for principal investigators
20	2.5	Online surveys	Feature allowing for conduction of online surveys
21	2.5	Process modelling	Feature supporting collaborative modelling of process
22	2.49	Chat (Text, Voice, Video)	Instant communication features
23	2.43	RSS-Feeds/Newsletter	Automated user notification of content updates and changes
24	2.41	Blog	Feature allowing users to share information to a wide audience
25	2.4	Workflow Management	IT support for CRC internal workflows
26	2.33	Single Sign on	Access all IT services from one single source and with one user account
27	2.3	Reporting	Features to support reporting of the project's progress
28	2.3	Virtual Reality	Features allowing for immersive experience of research results
29	2.23	Document digitalization	Features supporting digitalization of analogue documents
30	2.23	Personal Dashboard	Customizable personal landing page, which summarizes news relevant to a specific user
31	2.23	Plugins/Viewer	Platform add-ons which allow for viewing of different data types, e.g., CAD models
32	2.22	Augmented Reality	*cf.* "Virtual Reality"
33	2.17	Shared database	Common data base to store content created in the CRC
34	2.1	Forum	Shared discussion space for the CRC
35	2.04	Tangible User Interfaces	Devices and methods allowing for interaction with digital contents in a tangible way, e.g., by using physical objects
36	2.03	Formula editors	Web-based editors supporting mathematical formulas
37	2.03	Location Awareness	Feature indicating the presence of researchers in the shared work space
38	2.03	Shared Screen	Web conferencing-like feature allowing for sharing of user's screen
39	2	Enterprise Service Bus	Central feature, which combines and coordinates all available IT-services
40	1.94	Cloud Computing Access	Ubiquitous access to the shared workspace and its services

From the combined results of the exploratory talks and interviews, four major use cases were derived, each clustering a set of specific needs regarding one aspect of collaborative research activities and serving as frameworks for the development and composition of the envisioned CRC collaboration platform.

Requirements management refers to enabling researchers to discuss with each other about collaboration processes and to find mutual agreements regarding the exchange of research results. For this use case, IT technologies for web-content management, e-mail distribution, chat, document management, and online surveys should be used. A shared collaboration platform should allow researchers to define their requirements regarding requested research results from other projects by using web content management technology. E-mail distribution lists will enable the delegation to retrieve automated notification if requirements change or other tasks occur. Document management

will be used in order to save and exchange data and trace their evolution. Online surveys or simple polls should be used to collect requirements. Furthermore, all these technologies can also be used by individual projects in order to define requirements for themselves, and can thus be used for systematic project internal goal-setting.

Collaboration describes general requirements to be fulfilled to support coordinated research activities. The internal communication should be supported with unified and fully integrated messaging applications (voice, video, and text chat) and with location awareness. Contact lists, project webpages, and contextual information should support the coordination process and help people with finding the right peers and getting into contact with them. In order to store data and synchronize it with one's personal working environment, a web-based document management application should be implemented, offering features like versioning control and upload and download of content. To enable live cooperation (e.g., for distributed design reviews), a desktop sharing solution should be introduced.

Control project progress addresses specific needs of the CRC management and the principal investigators (PI) to track the progress of the entire CRC as well as individual projects. Management dashboards should be developed which summarize and display indicators relevant to the progress of the entire CRC, specific collaboration activities (e.g., demonstrators), or individual projects. A task management system should be applied to support the coordination of work packages and the tracking of their progress. Web-content management technology should allow for linking relevant content (e.g., discussion threads or documents) to specific task and for pre-defining upload spaces of results. E-mail distribution and RSS should allow for automated notification about changes in a task's state or comments from task assignees. Message board or comment features should be connected to the task management feature to enable discussions regarding specific tasks.

Factors surrounding data management were directly considered by the use case *Content discovery.* It addresses the need for appropriate methods and tools to ensure that any specific content can be stored and easily found by all researchers. This use case was also considered to be of increasing importance as the amount of content (*i.e.*, research results such as text, figures, pictures, measured data, *etc.*) will constantly grow during the course of the CRC. Document management features should be used to systematically store content in document databases to allow for searching content not only by the name of a file or its location in a folder structure, but also by its metadata (e.g., topic, author, project, creation date, *etc.*). Social tagging should be used in order to classify the content and build up a common wording for topics, types of research results, and their relations with each other. A keyword, content-type, and context sensible content search that is supplemented with indexing technologies should be realized. In order to notify users about changes in content they will be able to subscribe to updates via e-mail or RSS-Feeds. For the control of access rights, a mandatory access control system should be provided which also assures that content is only visible to defined roles, specific users, or user groups.

Finally, these use cases and the initial list of the 40 prioritized features provided a criteria set for software solicitation.

3. Results from the Implementation in the CRC 1026

At the beginning of a new collaborative research project, such as the CRC 1026, it is crucial to construct a shared conceptual framework and to build up a common knowledge base in order to enable cross-disciplinary collaboration [5]. During these processes, collaboration practices on organizational and individual level have to evolve and to consolidate. Thus, the INF project focused on the development of methodological and technological support of communication and coordination activities within the CRC to facilitate this phase of exploration. Data management in this context especially requires the provision of shared workspaces for data exchange, the management of task related data, as well as the creation of a semantic foundation on a project wide level allowing for the linkage of disciplinary data. During the course of the CRC INF constantly monitored the usage statistics of the implemented tools and evaluated the user requirements through various channels (e.g.,

personal talks, user workshops, and online surveys) to adapt its development activities. Furthermore, INF participated in numerous collaborative activities (e.g., workshops, group discussion, *etc.*) to observe and analyze actual collaboration practices.

3.1. Foundation of the Information Infrastructure—The CRC Collaboration Platform

The backbone of the CRC's information infrastructure is a collaboration platform based on the open source portal solution "Liferay" (liferay.org). This java-based framework includes an extensive set of built-in-applications, called "portlets", which provide a rich collection of basic collaboration features. For instance, the web-based document management feature, with its integrated version control and extensive metadata settings, allows for the simple exchange of data between individual projects. A role based access control system defines access privileges for researchers according to their role in the project (*i.e.*, principal investigator, researcher, external scholar, or student assistant) and their special assignments (e.g., managing director). Coordination tools, like shared contact lists and discussion boards, provide contact information and communication channels for individual exchange. The wiki provides a common ground for sharing disciplinary knowledge and building a shared understanding of essential concepts.

To adapt these default features to the identified use cases, modifications had to be carried out on different levels. In case of the shared calendar, for instance, adjustments were quite simple as only a reworked user interface and an Outlook connector was developed and applied. In other cases, such as the task management system described in Section 3.2, the functional logic of the Liferay system had to be modified with its own portlet developed. Another example for more sophisticated customization is the shared workspace on the collaboration platform, which is also called the "documents and media" section. To facilitate the usage of this shared workspace, it was re-designed to resemble a traditional folder structure. Besides shared folders, which are open to all members of the CRC, there should also be private folders with limited access for each individual project and for the management bodies. Furthermore, users should be able to create dedicated collaboration folders on-demand and set access privileges to those arbitrarily. To allow users to modify access privileges to that extent, the default access control logic of Liferay had to be modified, and a new document management portlet developed. The user interface of the shared workspace was further modified to distinguish between folders in which the respective user's project is directly involved in, and those which are open to all members of the CRC (see Figure 3). The former are displayed prominently in the upper section of the interface, while the latter are displayed in the lower section, as it is assumed that direct involvement is related to a higher interest for the user. This distinction is made based on access roles applied on the respective folders.

To further adapt Liferay's functional repertoire to meet the CRC's requirements, supplemental open source tools were integrated to fulfil functional requirements and to provide auxiliary services. For hosting of web applications, Apache™ Tomcat and Apache™ HTTP server are used (apache.org). MySQL (mysql.com) is used as the database management system. By adding a customized search engine based on Apache Lucene™ (apache.org), INF enabled users to search through stored contents according to their type, name, and text body. The web analytics tool Piwik (piwik.org) was implemented to track activities on the website and to enable analysis of page traffic, including: country of origin, the average time spent on the website, bounce-rates, and downloads. The free email server hMailServer (hmailserver.com) allows the Liferay system to process emails and enables automated notifications by the system. For web conferencing the open source system, BigBlueButton (bigbluebutton.org) was implemented. To utilize all the capabilities of BigBlueButton, supplemental service tools were implemented, such as: Red5 (red5.org) media streaming server, FreeSwitch (freeswitch.org) audio connector for voice over IP (VoIP), and Redis (redis.io) key value store for configuration management. Furthermore, Libre Office (libreoffice.org) was implemented for document processing. Finally, Limesurvey (limesurvey.org) was utilized to provide online survey capabilities. For load balancing the infrastructure, components are distributed among two virtual machines.
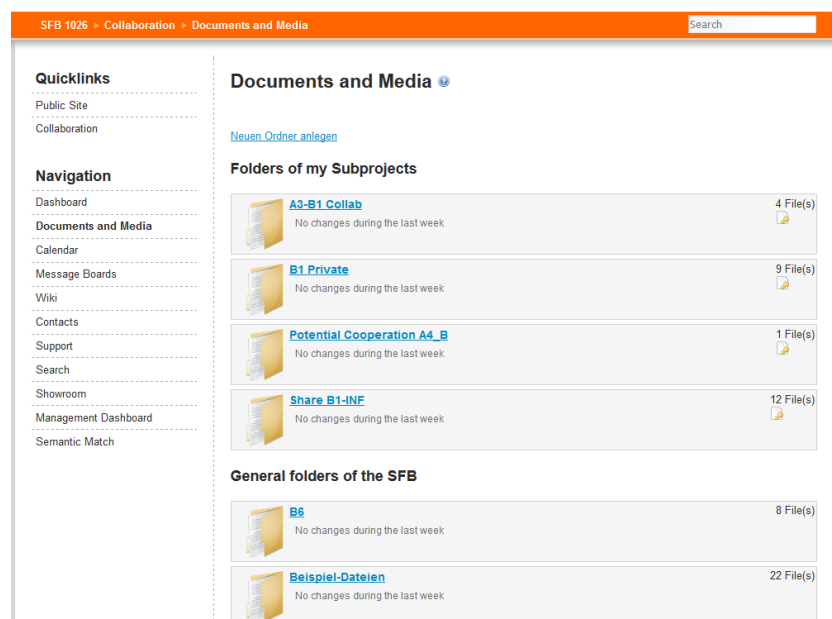
Figure 3. User interface of the "documents and media" section of the collaboration platform.

By September 2015 the collaboration platform had about 20 GB of data, 5000 items (documents, media files, *etc.*), 52 message boards with about 250 posts, and 99 wiki pages. The moderate amount of data stems mainly from the fact that large data packages such as motion capturing results were not stored there, but locally in the infrastructure of the respective researcher. For collaboration between the individual projects only relevant aspects of these data were extracted and exchanged through the collaboration platform (e.g., as tables, reports, publications, *etc.*).

Since December 2013 the usage of the internal collaboration area was tracked dedicatedly using member specific variables in the web analytics tool. The recorded traffic data from December 2013 until December 2015 show a slight but constant growth in the platform activities with punctual peaks concurring with events such as plenary meetings or upcoming conferences. On average, there were 17 unique visits per week to the internal collaboration area (see Figure 4). A unique visit in this regard means that each distinct user is only counted once for page visits within 24 h.
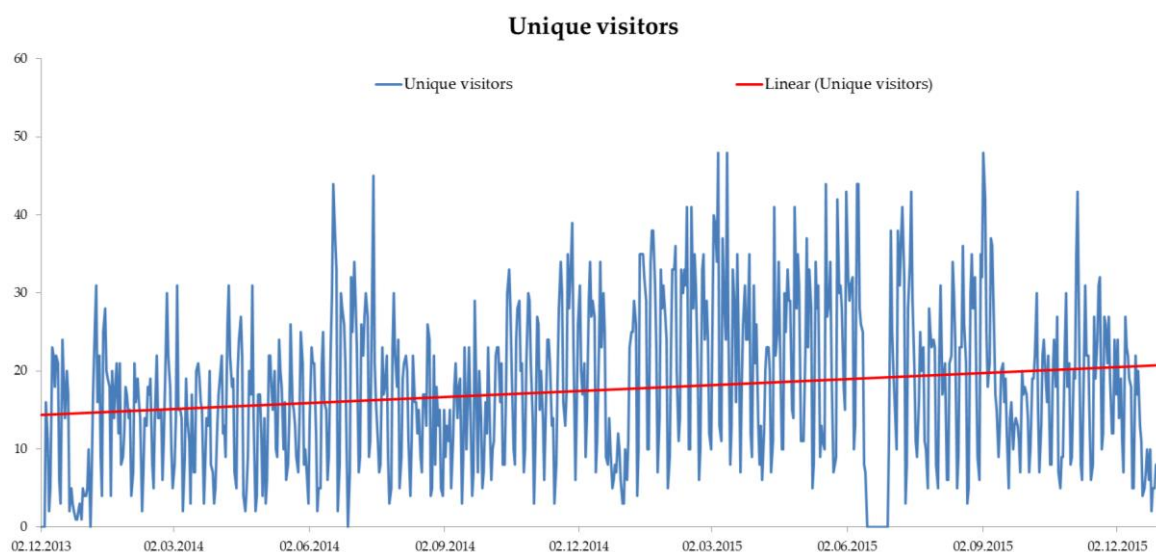

Figure 4. Usage statistic from the internal collaboration area from December 2013 until December 2015.

3.2. Management of Task Related Data within Large Groups

During the course of the CRC the formation of working groups and task forces became a usual form of collaboration. The management of tasks and task related data in such constellations is sometimes difficult, as their processing involves changing groups of persons. Keeping origination and evolutions of tasks transparent is essential to coordination of joint activities. Documentation in email correspondences or isolated meeting minutes often lead to high additional coordination efforts if, for example, changes of personnel occur. Hence, INF designed a task management system which is deeply integrated with the central group communication feature of the collaboration platform—the message boards. Each task has to be related to a message board thread in which its context is described and documented. The same thread provides a common space for discussions about the task, and makes eventual agreements about its processing transparent. Relevant documents, such as meeting minutes or used data and models, can be linked with it. Unlike e-mail-based communication, the structure of message board threads is comprehensible even to users joining the group in a later state. This approach fosters both the community building aspect and the transparency of collaboration activities within the CRC. By demanding task related data to be exchanged through the collaboration platform, the risk of data getting lost in the collaborative process is reduced. The availability and findability of relevant data is improved as they have to be stored in a shared environment and are independent from single persons (or respectively their E-mail in-boxes). The provided versioning control functionality of the collaboration platform reduces the risk of inconsistency resulting from redundant data storage (e.g., if several researchers are working on one paper). Furthermore, data linked to discussion threads are given a specific context and purpose, and hence, enriched with semantic meaning.

3.3. Ontology Based Data Management Concept and Prototype

To reach the overall goal of the CRC, research activities have to be coordinated and results combined. Hence, the creation of a shared understanding is also a substantial task of INF. To fulfill that task a common CRC ontology should be developed. Ontologies are formalized representations of a shared understanding of some domain of interest. An ontology necessarily embodies some kind of world view of a domain and can serve as a unifying framework to solve problems regarding communication between persons and/or organizations (e.g., by providing a normative model), interoperability (e.g., by enabling re-use and sharing of data/information/models between IT-systems), and systems engineering (e.g., by facilitating definition of requirements) [25]. In the CRC1026, an Ontology Working Group was established to define a collaborative development process and to coordinate development activities. Currently, the CRC ontology consists of ten sub-ontologies, each clustering concepts regarding specific aspects of sustainable manufacturing which are covered by individual projects (see Figure 5).
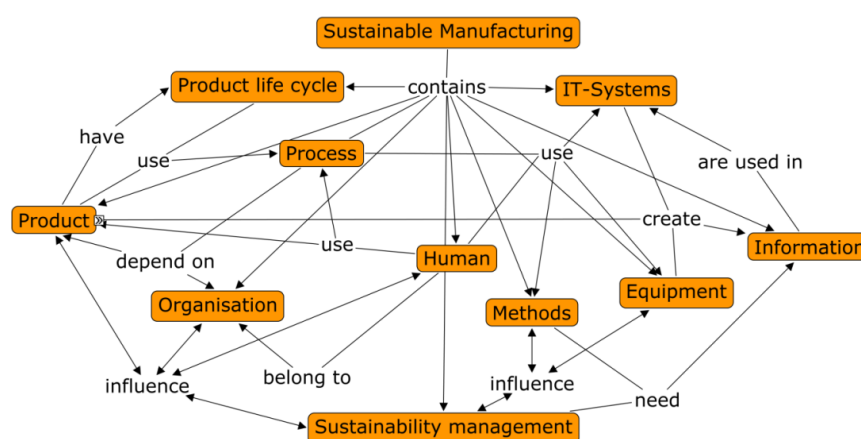

Figure 5. Overview of sub-ontologies developed.

These ontologies are modeled using standardized data-modelling frameworks, namely the Resource Description Framework Schema (RDF-S) and the Web Ontology Language (OWL), both providing mechanisms to formally describe ontologies (*i.e.*, groups of related concepts/resources and the relationships between these concepts/resources). Modeling tool Protégé 4.3 (protege.stanford.edu) is used. So far, the ontologies are only modeled in English, as it is the official language of the CRC 1026. Further languages were not considered for capacity reasons.

From the data management perspective, ontologies can facilitate (automated) data exchange and data integration across heterogeneous IT-systems, institutions, and disciplines by serving as inter-lingua [26,27]. In the context of the CRC, the aspect of supporting data discovery and organization with semantic technologies became increasingly relevant in the second half of the project's course, when collaborations got more intense and the amount of data stored on the platform grew constantly. In dedicated workshops with platform users (in which INF involved researchers from other projects to evaluate the development of platform tools), it was discussed that in an interdisciplinary context, sometimes search features were not sufficient for finding contents. In some cases, the users may not even know what exactly they were looking for when they intend to browse the collaboration platform for "useful" data. In this regard, the organization of data in a classical hierarchical folder structure was seen as unsuitable for large scale collaborations. Such structures tend to get increasingly confusing during the course of a project as folders and subfolders are created and named according to personal habits and organizational logics. Hence, an ontology-based concept to support interdisciplinary data management on the collaboration platform arose from these user workshops.

The idea behind that approach is that a shared vocabulary with defined semantic relations between enclosed terms, as represented by the CRC ontology, can be used as a framework to organize data stored in a shared environment. Matching the vocabulary with data objects (e.g., files, documents, or wiki pages) and/or their metadata (e.g., author, title, data type) allow those to be organized not only by physical attributes (e.g., location in the folder structure) but also by their semantic relations within the environment. Furthermore, the integration of semantic relations between data objects can uncover connections which otherwise might stay unnoticed.

The prototypical implementation on the collaboration platform should allow for researchers to look for specific concepts, and to explore the connections to other concepts within the CRC starting from their own project domain. While exploring the network of concepts, data objects related to those concepts are displayed. Hence, this approach will support the networking of disciplinary knowledge by facilitating the discovery of interfaces and relevant contents. For the prototype, only files, wiki pages, and message board threads on the platform were considered, as they represent the majority of the existing assets. Other assets such as tasks, calendar events, and blog entries can be included at a later stage. For the matching of concepts a semantic matching algorithm was applied, which was originally developed to connect ontological concepts to large corpora of scientific publications, thus generating refined data for bibliometric cluster analysis [28]. The following paragraphs describe the basic structure of the algorithm, and its practical application in the collaboration platform.

The basic idea of the semantic matching algorithm is to identify matches between concepts from an existing ontology with elements in large amounts of data objects based on text comparison. For the prototype, it means that matches between concepts from the CRC ontologies with texts from platform assets should be discovered and displayed. This requires the ontological concepts of the ontology models to be accessible in natural language. Thus, each concept in the CRC ontology models was annotated with a natural language name using the "rdfs:label" property from the standard RDF-S syntax.

Furthermore, the assets also need to contain natural language text in which the concepts can be identified by the algorithm. This text can be part of a title (e.g., a files' name), actual content (e.g., wiki page content), or keywords (e.g., tags). The text does not need to encompass some sort of pre-defined terminology, but should generally be allocated in a knowledge domain similar to the corresponding ontology.

The algorithm can generally be structured into three main parts. The first part discovers individual occurrences of concepts in these assets. This is realized by straightforward string comparisons identifying the abovementioned natural language labels of concepts in text components of assets. In order to further facilitate concept identification, both asset texts and concept labels are lemmatized. If a concept's label is identified somewhere in an asset's text, this is called an "individual occurrence" of the concept. As a label could potentially be identified in different parts of the asset, a concept can be observed to individually occur several times per asset as well.

The algorithm's second part cumulates these individual occurrences into overall scores for each combination of concept and asset. Consequently, a concept occurring often in an asset is scored highly, while a concept without any occurrence receives a score of zero. In order to calculate these scores, commonly used information retrieval methods are adapted for the semantic matching purpose. More specifically, the application of sublinear scaling, as well as a modified version of maximum term frequency normalization, enables the calculation of comparable scores between zero and one. As these calculations only use syntactic information so far, this part is also referred to as syntactic scoring, and correspondingly produces syntactic scores. Furthermore, the combination of a concept and an asset with syntactic scores greater than zero are hereafter called syntactic matches.

The third part of the algorithm handles the ontology's semantic information, determining scores for semantic cohesion between concept and asset. To calculate such a semantic score for a concept, the presence of ontologically connected concepts, henceforth referred to as 'semantic partners', are taken into account. Inspired by other methods of semantic search and concept similarity calculation [29–35], these partners include:

- Semantic children: Sub-concepts of the evaluated concept
- Semantic siblings: Sub-concepts of a super-concept of the evaluated concept
- Semantic neighbors: Concepts with any other direct semantic connection to the evaluated concept

Based on the cumulated syntactic scores of these semantic partners and related to their potential maximum presence, the semantic score represents the degree of semantic cohesion. It ranges from zero, with no semantic partner at all, to one, for all possible semantic partners with maximum syntactic score. Consequently, it is possible to calculate semantic scores for concepts independent from individual occurrences of these concepts in a given asset. The semantic score can therefore be regarded as a measurement of implicit concept occurrence. Semantic and syntactic scores are finally combined in equal parts to generate a total score for all combinations of information assets and concepts.

This prototypical implementation of a semantic matching mechanism enables navigation throughout concepts and assets based on their connections (matches). For each asset, occurring concepts are listed and can be selected, while for each concept all assets containing this concept are displayed. With further development of this semantic matching approach, various additional applications are thinkable. To extend the prototype to include other areas of the collaboration platform and other types of assets, the identified concept matches could simply be used like tags or keywords. For example, an ontology based tag cloud would provide users with an intuitive approach to explore the semantic relations of assets, while simultaneously providing a general overview of the platform's main communication and information trends. Furthermore, based on various approaches on semantic search [30,31,33,35,36], it would be possible to extend and improve the platform's search algorithm, using the semantic information of the concept/asset matches to optimize recall or precision. On the one hand, recall could be enhanced by automatically expanding search queries with ontological information and matching this with concept matches and scores to retrieve assets not found by traditional information retrieval methods. On the other hand, the concept scores could be used to refine relevance scores, thus providing more relevant query results and improving overall precision. Last but not least, the prototype application could be augmented with further functionality and improved presentation, thus enabling users to dynamically navigate through the network of concepts and assets. This could

present the opportunity to comprehensively analyze this network's interrelatedness [37], and even enrich the underlying ontology and matching results with individual comments and annotations.

4. Discussion of the Solution Deployment

Large scale interdisciplinary projects have high demands on coordination and communication, especially at the beginning. To create conditions so that information and knowledge can be produced in a self-organized way is a decisive management task within interdisciplinary projects [38]. Its success largely depends on identification and application of proper measures to stimulate collaboration, and for building group identity and mutual trust [5,39,40].

In the first funding period of the CRC 1026, the INF project developed solutions supporting task related data management and semantic networking of disciplinary domains. The presented task management system improves the collection and traceability of task related data, and proved to be especially useful during the preparation of the proposal for the second funding period, where multiple tasks and deliverables had to be coordinated among all individual projects simultaneously. On the other hand, the reservation must be made that, especially for simple tasks, the effort to perform the obligatory steps for creating a task sometimes exceeds the benefits from the researcher's point of view. Thus, the success of such a task management system highly depends on the commitment of involved institutions and the motivation of researchers. In the former case, peculiarities of single institutions, such as naming conventions, have to be adapted and unified to benefit from the versioning control system. Individual concerns (e.g., about safety and robustness of the applied system) have to be relieved by according measures (e.g., regular data backups) and involving users in extensive pre-launch tests. This latter aspect can be countered by improving the usability on a technical level, and by providing incentives on an organizational level (e.g., by the integration of nontraditional systems of valuing contributions) [39].

The ontology based data management approach proved to be promising with regard to interdisciplinary data management. It provides a conceptual basis for collaboration and for categorization of data by displaying logical interrelations between individual projects. Moreover, the application of semantic technologies and ontologies is becoming increasingly relevant to the management of research data [26,41,42]. Hence, more general data management applications similar to the prototype described in Section 3.3 are imaginable, provided that data are appropriately annotated and described by metadata. For example, the semantic matching algorithm could be applied to any kind of data— just like for files of the collaboration platform—given that title, keywords, and description of respective data exist in natural language. Thus, tag augmentation, semantic search expansion, or network navigation (as described above and in Section 3.3), would also be possible in other data management structures. Considering data lifecycle management, this could be applied when preserving data (e.g., by not only creating metadata and documentation, but also subsequently determining concept matches and semantic interrelations). Accessibility and re-usability could then be improved by enhancing corresponding search algorithms and providing advanced navigation possibilities for facilitating the network of semantic connections. If semantic matching is additionally applied to the documents and publications resulting from data analysis, semantic similarities could be determined between data and documents that otherwise would have no connections. However, we believe that semantic matching can provide the basic method enabling the further functionalities to be built on. Especially in an interdisciplinary collaboration context, semantic matching applications can help to reveal hidden relations between data from different disciplinary repositories and foster cross-disciplinary fertilization.

However, the time consuming development process of both the ontology itself and the matching algorithm allowed only a very late deployment of the tool. Hence, extensive user tests could not be conducted.

From the deployment perspective, the next steps for the development of the semantic matching prototype will comprise the design of a comprehensive user interface with an appropriate visualization

form for semantic interrelations and further definition of usage scenarios. By now, two general scenarios have been identified. The "targeted search" scenario focuses on cases where researchers are purposefully looking for specific data within the CRC collaboration platform. The second scenario considers more exploratory cases where researchers are just "looking for connections" to other individual projects. Both scenarios demand specific mechanisms and features to be integrated into the semantic matching tool. On the theoretical side, the existing ontologies have to be further developed in order to connect the CRC specific knowledge model to existing information models in industrial (e.g., ISO 10303-239 PLCS, www.plcs.org) and public context (e.g., DBpedia). In the context of considering the domain of sustainable manufacturing, a consistent metadata framework is needed, such as the CERA2 framework for climate research data.

5. Conclusions

The experiences from the CRC 1026 show that the integration of data management in interdisciplinary projects is still a challenging aspect. Despite a growing awareness for data management, there is still a predominantly traditional perception of data management as a concluding activity of research endeavors in many scientific domains. Consequently, data management is still often seen as being synonymous to publication management. A confusingly high number of copyright models, embargo clauses, and license models applied by the different publishers in turn cause researchers to shy away from using centralized publication management systems or institutional repositories as they fear encountering copyright issues. Another hurdle for data management is the reluctance regarding the provision of raw data generated in the course of research activities. This partly results from confidentiality concerns (e.g., when data are relevant to pending patent applications or are used in current doctoral thesis). In other cases, collected data were rated as being "too project specific" to be of use for other research projects. Finally, tasks related to systematic collection and provision of data are sometimes perceived as overhead, which do not—or at least not directly—serve one's professional advancement and are not rewarded (by reputation, for example).

To face these challenges, researchers should be sensitized to data management aspects and their benefits. Especially at the beginning of interdisciplinary projects, common guidelines should be created regarding what "research data" are and how these should be handled in a collaborative environment. Furthermore, researchers should be made aware of the benefits of data management efforts by practical examples. For instance, it could be demonstrated that by providing primary data on their publications the researchers not only support the scientific advancement of their respective domain, but also strengthen the scientific validity of their work, and hence improve its quality. This in turn would add to their reputation. For INF projects, one substantial task in supporting data management should be the development and provision of services facilitating the process of data management. On the one hand, this can be realized by IT-tools which, for example, can allow for the extraction of metadata from items stored in collaborative environments in standardized formats (e.g., BibTex). On the other hand, INF projects can provide preparatory services (e.g., pre-processing of collected data and trainings regarding usage of existing repository services). In this respect, the expertise and information services of university libraries or external data management experts should also be utilized.

The information infrastructure developed by INF is suitable to collect and store data during the course of the CRC. However, to ensure long-term preservation and availability of research data, more persistent infrastructures such as institutional repository services from university libraries should be utilized. In case of the CRC 1026 the repository service, "DepositOnce" will be used to preserve research results. This repository is developed and maintained by the Service Center for Research Data and Publications (SZF) of TU Berlin. Each data set stored in "DepositOnce" will be provided with a digital object identifier (DOI) so it can be referenced to in further research works and publications. As the SZF is a joint service center from the university library, the central IT-service center, and the research department of TU Berlin, the persistency of the repository system is ensured.

Author Contributions: Wei Min Wang is a research assistant at the institute for Industrial Information Technology at TU Berlin and project leader of the INF project. He is mainly responsible for the planning and conceptual design of IT-Tools. Furthermore, he prepared and conducted interviews and surveys among the CRC researchers for requirements regarding monitoring and engineering. Tobias Göpfert was a student assistant at TU Berlin and a member of INF. He assisted in the design of IT-Tools and was mainly responsible for their technical development and implementation. He also developed the semantic matching algorithm applied in the CRC prototype for semantic data management. Prof. Dr.-Ing. Rainer Stark is Chair of the Institute for Industrial Information Technology at TU Berlin and director of the Fraunhofer department of Virtual Product Creation. He is also the speaker of the CRC 1026. He supervised the INF project and gave important input for the design of tools and methodological approaches.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CRC	Collaborative Research Center
DFG	Deutsche Forschungsgemeinschaft (German Science Foundation)
INF	Service Project for Information Infrastructure
GHG	Greenhouse gas
RDF-S	Resource Description Framework Schema
OWL	Web-Ontology Language
SZF	Servicezentrum für Forschungsdaten und –publikationen (Service Center for Research Data and Publications)

References

1. Haythornthwaite, C. Learning and knowledge networks in interdisciplinary collaborations. *J. Am. Soc. Inf. Sci. Technol.* **2006**. [CrossRef]
2. Raasch, C.; Lee, V.; Spaeth, S.; Herstatt, C. The rise and fall of interdisciplinary research: The case of open source innovation. *Res. Policy* **2013**. [CrossRef]
3. Corley, E.A.; Boardman, P.C.; Bozeman, B. Design and the management of multi-institutional research collaborations: Theoretical implications from two case studies. *Res. Policy* **2006**. [CrossRef]
4. Katz, J.S.; Martin, B.R. What is research collaboration? *Res. Policy* **1997**, *26*, 1–18. [CrossRef]
5. Pennington, D.D. Cross-Disciplinary Collaboration and Learning. Available online: http://www.ecologyandsociety.org/vol13/iss2/art8/ (accessed on 6 January 2016).
6. Williams, P. The Competent Boundary Spanner. *Public Adm.* **2002**. [CrossRef]
7. Patel, H.; Pettitt, M.; Wilson, J.R. Factors of collaborative working: A framework for a collaboration model. *Appl. Ergon.* **2012**. [CrossRef] [PubMed]
8. Van Rijnsoever, F.J.; Hessels, L.K. Factors associated with disciplinary and interdisciplinary research collaboration. *Res. Policy* **2011**. [CrossRef]
9. Nooteboom, B. Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance. *J. Manag. Gov.* **2000**. [CrossRef]
10. Büttner, S.; Hobohm, H.; Müller, L. Research Data Management. In *Handbuch Forschungsdatenmanagement*; Büttner, S., Hobohm, H., Eds.; Bock + Herchen: Bad Honnef, Germany, 2011; pp. 13–24.
11. Dallmeier-Tiessen, S. Strategien bei der Veröffentlichung von Forschungsdaten. In *Handbuch Forschungsdatenmanagement*; Büttner, S., Hobohm, H., Eds.; Bock + Herchen: Bad Honnef, Germany, 2011; pp. 169–190.
12. Effertz, E. The funder's perspective: Data management in coordinated programmes of the German Research Foundation (DFG). In Proceedings of the Data Management Workshop, Cologne, Germany, 29–30 October 2009.

13. Rümpel, S. Der Lebenszyklus von Forschungsdaten. In *Handbuch Forschungsdatenmanagement*; Büttner, S., Hobohm, H., Eds.; Bock + Herchen: Bad Honnef, Germany, 2011; pp. 25–34.

14. Huschka, D.; Oellers, C.; Ott, N. Datenmanagement und Data Sharing: Erfahrungen in den Sozial- und Wirtschaftswissenschaften. In *Handbuch Forschungsdatenmanagement*; Büttner, S., Hobohm, H., Eds.; Bock + Herchen: Bad Honnef, Germany, 2011; pp. 35–48.

15. Wolter, L.; Wang, W.M.; Rainer, S. Information Systems to Support Collaboration in Large Research Projects. *IJMO* **2014**. [CrossRef]

16. Parent-Thirion, A.; Vermeylen, G.; van Houten, G.; Lyly-Yrjänäinen, M.; Biletta, I.; Cabrita, J. 5th European Working Conditions Survey. Available online: http://www.eurofound.europa.eu/surveys/2010/fifth-european-working-conditions-survey-2010 (accessed on 8 January 2016).

17. Lapillonnee, B.; Pollier, K.; Samci, N. Energy Efficiency Trends in the EU—Lessons from the ODYSSEE MURE Project. Available online: http://www.odyssee-mure.eu/publications/br/synthesis-energy-efficiency-trends-policies.pdf (accessed on 14 January 2016).

18. European Commission - Directorate-Genaral for Energy (op. 2013): EU energy in figures. Statistical pocketbook 2013. Luxembourg. Available online: https://ec.europa.eu/energy/sites/ener/files/documents/2013_pocketbook.pdf (accessed on 18 January 2016).

19. Duflou, J.R.; Sutherland, J.W.; Dornfeld, D.; Herrmann, C.; Jeswiet, J.; Kara, S.; Hauschild, M.; Kellens, K. Towards energy and resource efficient manufacturing: A processes and systems approach. *CIRP Ann. Manuf. Technol.* **2012**. [CrossRef]

20. Diallo, Y.; Etienne, A.; Mehran, F. *Global Child Labour Trends 2008 to 2012*; ILO: Geneva, Switzerland, 2013.

21. Steckel, J.C.; Edenhofer, O.; Jakob, M. Drivers for the renaissance of coal. *Proc. Natl. Acad. Sci. USA* **2015**. [CrossRef] [PubMed]

22. Edmund, A.S. Geschichte der Nachhaltigkeit: Vom Werden und Wirken Eines Beliebten BEGRIFFES. 2012. Available online: http://www.nachhaltigkeit.info/media/1326279587phpeJPyvC.pdf (accessed on 21 April 2014).

23. Wissenschaftliche Dienste des Deutschen Bundestages. Nachhaltigkeit: Wissenschaftliche Dienste des Deutschen Bundestages. 2004. Available online: http://webarchiv.bundestag.de/archive/2008/0506/wissen/analysen/2004/2004_04_06.pdf (accessed on 21 April 2014).

24. Giddings, B.; Hopwood, B.; O'Brien, G. Environment, economy and society: fitting them together into sustainable development. *Sustain. Dev.* **2002**. [CrossRef]

25. Uschold, M.; Grüninger, M. Ontologies: Principles, methods and applications. *Knowl. Eng. Rev.* **1996**, *11*, 93–136. [CrossRef]

26. Neher, G.; Ritschel, B. Semantische Vernetzung von Forschungsdaten. In *Handbuch Forschungsdatenmanagement*; Büttner, S., Hobohm, H., Eds.; Bock + Herchen: Bad Honnef, Germany, 2011; pp. 169–190.

27. Wang, W.M.; Pförtner, A.; Lindow, K.; Hayka, H.; Stark, R. Using ontology to support scientific interdisciplinary collaboration within joint sustainability research projects. In Proceedings of the 11th Global Conference on Sustainable Manufacturing: Innovative Solutions, Berlin, Germany, 23–15 September 2013; pp. 612–617.

28. Göpfert, T. Matching Ontological Concepts with Documents: A Semantic Matching Approach in Context of Bibliometric Community Detection. Mater's Thesis, Humboldt–Universität zu Berlin, 2016, in press.

29. Giunchiglia, F.; Kharkevich, U.; Zaihrayeu, I. Concept Search. *Lect. Notes Comput. Sci.* **2009**, *5554*, 429–444.

30. Navigli, R.; Velardi, P. An Analysis of Ontology-based Query Expansion Strategies. Cavtat-Dubrovnik. In Proceedings of the Conference on Machine Learning (ECML 2003), Dubrovnik, Croatia, 22–26 September 2003.

31. Khan, L.; McLeod, D.; Hovy, E. Retrieval effectiveness of an ontology-based model for information selection. *VLDB J.* **2004**, *13*, 71–85. [CrossRef]

32. Anyanwu, K.; Sheth, A. ρ-Queries: Enabling Querying for Semantic Associations on the Semantic Web. In Proceedings of the 12th International Conference on World Wide Web, Budapest, Hungary, 20–24 May 2003.

33. Zhou, Q.; Wang, C.; Xiong, M.; Wang, H.; Yu, Y. SPARK: Adapting Keyword Query to Semantic Search. *Lect. Notes Comput. Sci.* **2007**, *4825*, 694–707.

34. Rocha, C.; Schwabe, D.; de Aragao, M.P. A Hybrid Approach for Searching in the Semantic Web. In Proceedings of the 13th International Conference on World Wide Web, New York, NY, USA, 17–22 May 2004.

35. Cheng, G.; Ge, W.; Qu, Y. Falcons: Searching and Browsing Entities on the Semantic Web. In Proceedings of the 17th International World Wide Web Conference, Beijing, China, 21–25 April 2008.

36. Fernandez, M.; Lopez, V.; Sabou, M.; Uren, V.; Vallet, D.; Motta, E.; Castells, P. Semantic Search Meets the Web. In Proceedings of the 2008 IEEE International Conference on Semantic Computing, Santa Clara, CA, USA, 4–7 August 2008; pp. 253–260.

37. Wang, W.M.; Wolter, L.; Lindow, K.; Stark, R. Graphical Visualization of Sustainable Manufacturing Aspects for Knowledge Transfer to Public Audience. *Proced. CIRP* **2015**. [CrossRef]

38. Fuchs-Kittowski, K.; Heinrich, L.J.; Rolf, A. Information entsteht in Organisationen—In Kreativen Unternehmen—Wissenschaftstheoretische und methodologische Konsequenzen für die Wirtschaftsinformatik. In *Wirtschaftsinformatik und Wissenschaftstheorie*; Springer: Berlin, Germany, 1999; pp. 329–361.

39. Wenger, E.C.; Snyder, W.M. Communities of practice: The organizational frontier. *Harv. Bus. Rev.* **2000**, *78*, 139–146.

40. Pickett, S.T.; Burch, W.R., Jr.; Grove, J.M. Interdisciplinary research: Maintaining the constructive impulse in a culture of criticism. *Ecosystems* **1999**, *2*, 302–307. [CrossRef]

41. Cox, A.; Milsted, A.J.; Gutteridge, C.J. Linked Data Technology Making Autodiscovery of Data a Reality. In Proceedings of the 2nd Data management Workshop, Cologne, Germany, 28–29 November 2014.

42. Decker, B.; Politze, M. Ontology Based Semantic Data Management for Pandisciplinary Research Projects. In Proceedings of the 2nd Data management Workshop, Cologne, Germany, 28–29 November 2014.