Grounded Design and GIScience - A framework for informing the design of geographical information systems and spatial data infrastructures

Alexander Kmoch Corresp., 1, 2 , **Evelyn Uuemaa** 1 , **Hermann Klug** 3

1 Department of Geography, Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia

2 School of Engineering, Computer & Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand

3 Interfaculty Department for Geoinformatics (Z_GIS), Universität Salzburg, Salzburg, Austria

Corresponding Author: Alexander Kmoch
Email address: alexander.kmoch@ut.ee

Geographical Information Science (GIScience), also Geographical Information Science and Systems, is a multi-faceted research discipline and comprises a wide variety of topics. Investigation into data management and interoperability of geographical data and environmental data sets for scientific analysis, visualisation and modelling is an important driver of the Information Science aspect of GIScience, that underpins comprehensive Geographical Information Systems (GIS) and Spatial Data Infrastructure (SDI) research and development. In this article we present the 'Grounded Design' method, a fusion of Design Science Research (DSR) and Grounded Theory (GT), and how they can act as guiding principles to link GIScience, Computer Science and Earth Sciences into a converging GI systems development framework. We explain how this bottom-up research framework can yield holistic and integrated perspectives when designing GIS and SDI systems and software. This would allow GIScience academics, GIS and SDI practitioners alike to reliably draw from interdisciplinary knowledge to consistently design and innovate GI systems.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

Grounded Design and GIScience - A framework for informing the design of Geographical Information Systems and Spatial Data Infrastructures

Alexander Kmoch[1,2], Evelyn Uuemaa[1], and Hermann Klug[3]

[1]**Department of Geography, Institute of Ecology and Earth Sciences, University of Tartu, Estonia**
[2]**School of Engineering, Computer & Mathematical Sciences, Auckland University of Technology, New Zealand**
[3]**Interfaculty Department for Geoinformatics (Z_GIS), University of Salzburg, Austria**

Corresponding author:
Alexander Kmoch[1]

Email address: alexander.kmoch@ut.ee

ABSTRACT

Geographical Information Science (GIScience), also Geographical Information Science and Systems, is a multi-faceted research discipline and comprises a wide variety of topics. Investigation into data management and interoperability of geographical data and environmental data sets for scientific analysis, visualisation and modelling is an important driver of the Information Science aspect of GIScience, that underpins comprehensive Geographical Information Systems (GIS) and Spatial Data Infrastructure (SDI) research and development. In this article we present the 'Grounded Design' method, a fusion of Design Science Research (DSR) and Grounded Theory (GT), and how they can act as guiding principles to link GIScience, Computer Science and Earth Sciences into a converging GI systems development framework. We explain how this bottom-up research framework can yield holistic and integrated perspectives when designing GIS and SDI systems and software. This would allow GIScience academics, GIS and SDI practitioners alike to reliably draw from interdisciplinary knowledge to consistently design and innovate GI systems.

1 INTRODUCTION

Geographical Information Science (GIScience), also Geographical Information Science and Systems, is a multi-faceted research discipline and comprises a wide variety of topics (Goodchild, 1992; Mark, 2002). Investigation into data management and interoperability of geographical and environmental data sets for scientific analysis, visualisation and modelling is a particular challenge of the Information Science aspect of GIScience (Yang et al., 2010). Methods and techniques around the semantic web and linked data, data visualisation, distributed computing and knowledge engineering are typical research topics in pure Information Science and Systems research. While such research is mainly being conducted in the domain of business economics and Computer Science studies, the interdisciplinary Information Science aspects are an essential research aspect to integrate Information Science not only with the wider field of Geography, but even further with the domain of Earth Sciences or Geosciences (Blaschke et al., 2011; Gahegan et al., 2009).

One of the big goals of GIScience is to improve the use and reuse of (geographical) data in scientific analysis, visualisation and modelling, and the formal methods thereof, which are required to become more interdisciplinary and interconnected to solve the challenges of our time (Goodchild, 1992; Mark, 2002). While GIScience has been disputed for being an independent scientific discipline for a period around its emergence (Pickles, 1997), its interdisciplinary nature is of more relevance than ever (Goodchild, 2006, 2009; Yang et al., 2010). This article contributes to the field of GIScience, in particular on the interdisciplinary information science aspects, where methods and techniques around spatial and semantic data interoperability for visualisation, distributed geoprocessing and knowledge engineering for Geographical Information Systems (GIS) and Spatial Data Infrastructure (SDI) research and development are essential (Blaschke et al., 2011; Goodchild, 2009). GIScience can lead interdisciplinary research that encompasses knowledge from the Earth Sciences and Computer Science (Figure 1).

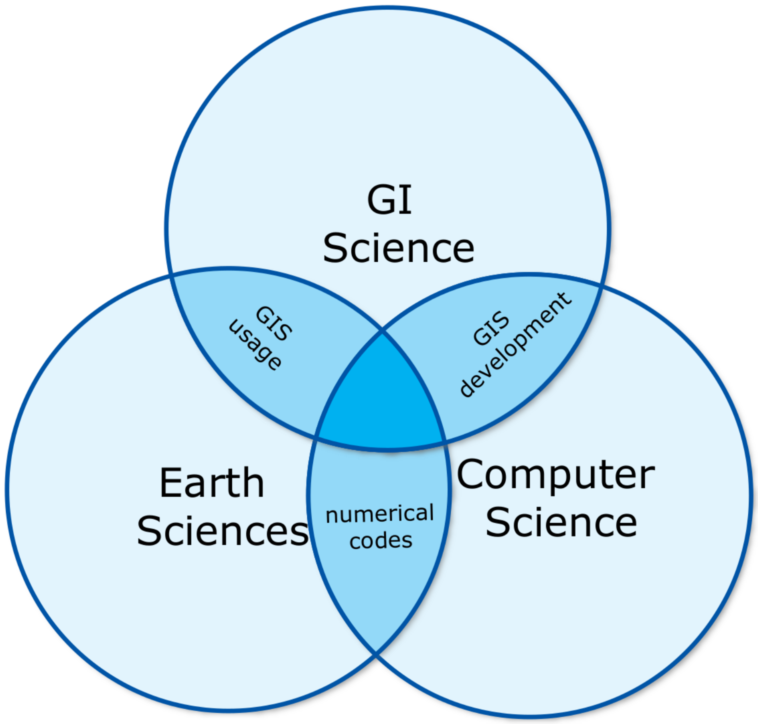

Figure 1. GIScience, as the holistic connecting discipline lays claim to the central cross section, draws theory from Earth Sciences and Computer Science to lead advances in interdisciplinary systems development research

To advance theory in GIScience existing gaps need to be revealed and the current understanding of phenomena needs to be improved. The context of the presented research framework is 1) the ever-increasing collection of specifically geoscientific and environmental data in general, in a wide variety of formats, 2) the evolving data formats, data and database models and formalised standards and 3) their application in data collection, storage, harmonisation, discovery, access and transfer in the Earth Sciences. A central assumption for this research framework is that method and techniques of geoscientists, GIS practitioners, information modellers and computer scientists are not sufficiently coalesced to improve efficiency and effectiveness of interdisciplinary and transdisciplinary data handling (Bandaragoda et al., 2006; Tress et al., 2003).

GIScience as a research discipline comprises a large variety of research topics, with some of them being classically of quantitative and empirical nature, by for example determining statistical and spatial relationships of geographical features. Other topics are not always as clearly objective or empirical in nature, e.g. the GIS software design and development as part of GIScience theory development and knowledge contribution.

The researcher needs to explain his research perspective to clarify the application of methods and modes of scientific inquiry. This is particularly important, if qualitative methods are employed and possible bias is involved (Creswell, 2003; Holliday, 2010; Saunders et al., 2012). However, researchers find themselves regularly in a situation where they pose the hypotheses based on existing data instead of using traditional empirical scientific method which requires posing hypotheses based on previous studies and then testing them based on your own data (Figure 2).

Thus, to avoid disguising the conduct of a study under the cloak of the classic scientific method, an explicit approach has been chosen that better represents the needs of framework development in the

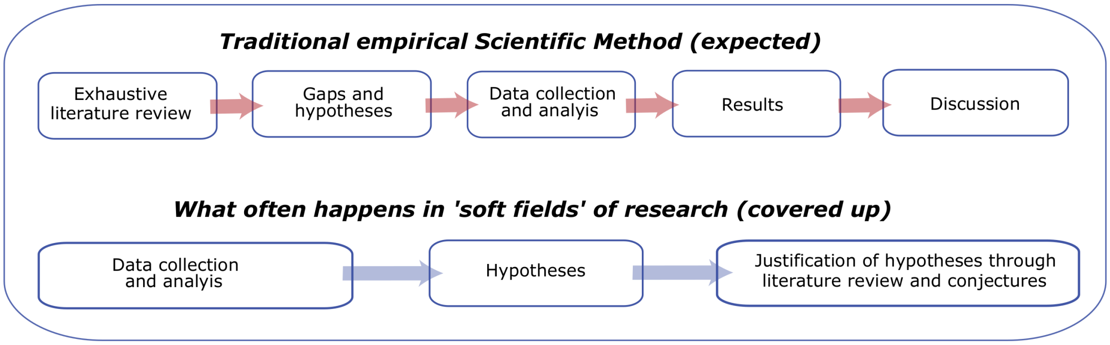

Figure 2. Discrepancies in modes of research typically adopted in GIScience, showing the challenging irony of the research approaches required to be conducted in a strictly positivist community, adopted from Walsh (2014)

GIScience and Systems discipline.

Classical philosophy of science is differentiating between objective and subjective epistemological stances, which define the researcher's relationship with the notion of what is knowledge and truth. An objective epistemology with a positivist, empirical paradigm embodies the Scientific Method and assumes an external reality. Thus, there is only one truth and the generated knowledge is independent of the researcher (Goles and Hirschheim, 2000; Gray, 2014). A subjective epistemological stance assumes that knowledge is not value free and thus the reality is dependent on the researcher, human perceptions and the social context. Therefore bias should be articulated accordingly (Bhaskar, 1998; Dobson, 2002; Mingers, 2004). Interpretivist and constructivist paradigms are commonly described as subjective perspectives (Feast and Melles, 2010; Hirschheim, 1985). In contrast to traditional approaches, philosophical pragmatists deny the correspondence notion of truth completely and propose that truth essentially is what works in practice, thus in Information Science and Systems research progress is achieved when existing technologies are replaced by more effective ones (Goles and Hirschheim, 2000; Porra et al., 2014).

Critical Realism is another, more mediating and increasingly accepted philosophy of science (Bhaskar, 1998; Dobson, 2002). Critical Realism acts as a bridge between positivist and constructivist/interpretivist stances and guides the researcher in addressing bias in qualitative and mixed methods research (Scott, 2007; Zachariadis et al., 2010). Critical Realism reflects on the underlying knowledge generating process, the conflicts in society and takes on an emancipatory role (Easterbrook et al., 2008).

Although most researchers focus on one particular research method, some researchers have suggested combining one or more research methods via means of triangulation (Couclelis, 2009; Gray, 2014; Sikolia et al., 2013). This paper is proposing to integrate several qualitative methods of inquiry, i.e. Design Science Research (DSR), Grounded Theory (GT) and Case Studies (Beck et al., 2013; Glaser and Strauss, 1967; Yin, 1994). GT advises that the researcherFLs background has an impact on his or her ability to let patterns emerge from the data, and that preconceived ideas can influence his or her ability thereof. Here the reflecting perspective of Critical Realism is supporting the researcherFLs discipline in handling bias and avoiding preconceived ideas. The constant critical reflection is in tune with letting categories emerge while constantly comparing theory with new data (Birks et al., 2013; Thornberg, 2012). From the DSR perspective it is important to address perceived challenges that a new design should overcome (Beck et al., 2013; Gregor, 2002; Gregory, 2010). Here the researcher needs to take an emancipatory stance how the emerging theory can inform the design iterations. The three methods are applied jointly in triangulation, which comprises the foundation of the presented 'Grounded Design' research method(Figure 3). The following sections will describe each method in detail.

2 GROUNDED THEORY - BACKGROUND

GT was first published by the two sociologists Barney Glaser and Anselm Strauss (Glaser and Strauss, 1967). Glaser describes it being a full research methodology. Its core emphasis is on **discovery of theory from data**. Since then it has been used in a variety of disciplines including Information Systems and Software Engineering (Urquhart et al., 2010). As opposed to an empirical positivist approach of testing

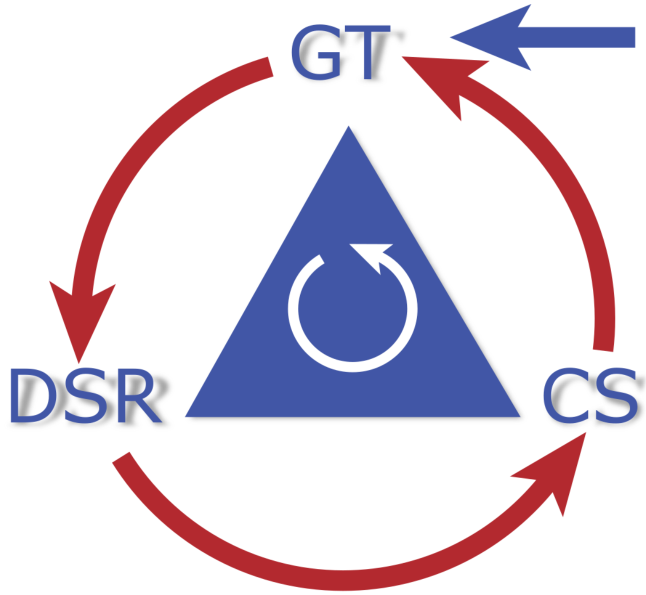

Figure 3. Triangulation of Methods: GT, DSR, Case Studies (CS), the blue arrow indicates the starting point - GT establishing initial theory from available data, improvement via DSR and verification over case studies, results as input for next improvement iteration

existing theory, GT presents a method for generating theory from data. This data can be collected from interviews, documents, observation and other sources as soon as it becomes available. Analysis involves the identification of concepts, sub-categories, and categories and how they relate to each other. The emerging categories and their relationships are then checked with existing literature in the field to explain how they relate to each other. While in Information Systems research GT is often used for coding their data, the prime result of this method is underlying theory (Urquhart et al., 2010). Figure 4 illustrates the stages, processes and in- and outputs of GT as explained in detail in the following sections.

One particular aspect that differentiates GT from many other research methods and methodologies is that it is explicitly emergent and does not aim to prove or falsify a hypothesis in the first place (Glaser, 2008; Glaser and Strauss, 1967). The aim is to understand the research situation, as Glaser states it, to discover the theory implicit in the data (compare with Figure 2). GT provides reconciliation where the field of geographical and environmental spatio-temporal data and software integration for earth sciences is complex and scattered. The in-depth analysis of the situation based on GT will reveal required patterns to inform a holistic design approach subsequently (Beck et al., 2013).

2.1 Research Situation and Minor Literature Review

In GT initial research questions are to be avoided and an in-depth gap analysis of the field under investigation might blind the researcherFLs ability to let theory emerge. But to frame the area of interest and based on first impressions from the field as well as to comply with academic guidelines for research publications, the researcher can start off with a minor literature review to be able to understand technical jargon and the current research situation within the data. Glaser instead suggests background reading in adjunct fields to provide the understanding to make sense of the data, while avoiding the most closely related literature. Consequently, the increasingly comprehensive literature review can be progressed as the literature becomes relevant to the emerging categories and theory, as otherwise early reading may constrain coding and memoing too much (Glaser and Strauss, 1967; Strauss and Corbin, 1994; Thornberg, 2012).

An important concern is how the research treats disagreement between the emerging theory and the literature. The researcherFLs throughout to fit the emerging theory to the data and to make sense of the actual situation. A means to tackle the assumption that the theory could be wrong is to stick critically with the constant comparative method and also to seek dis-confirming evidence. A study that applies Grounded Theory seeks to extend the theory so that it makes sense of both the data from the case study

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

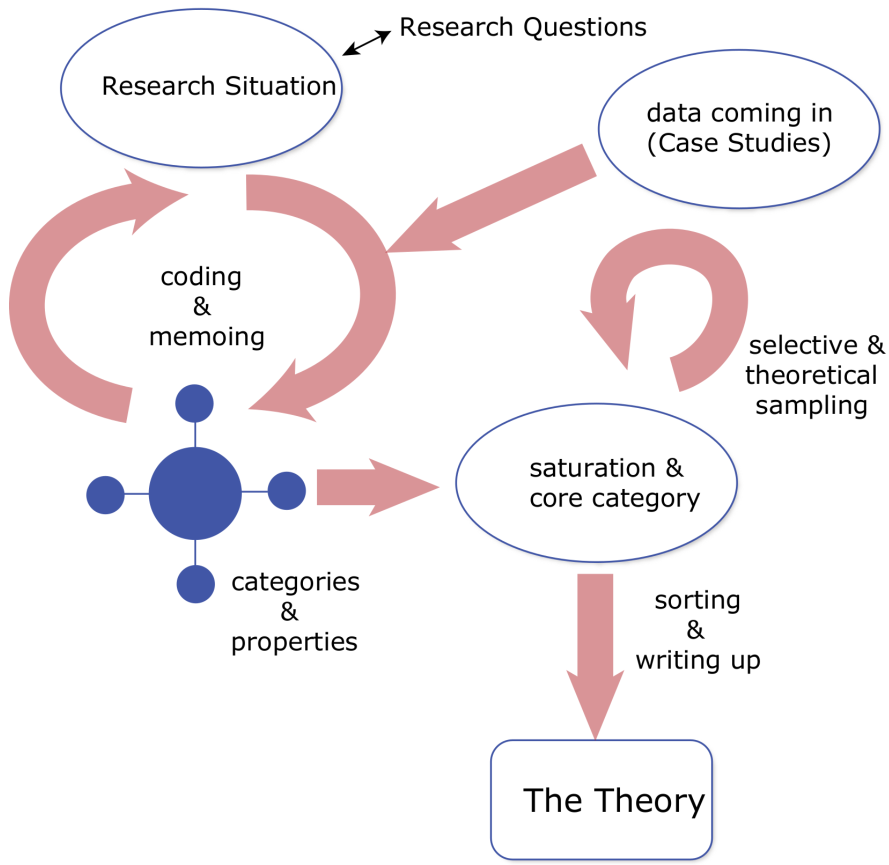

Figure 4. Grounded Theory Process, from Research Situation to resulting Theory (with data from Case Studies)

and the data from the literature (Birks et al., 2013; Sikolia et al., 2013).

2.2 Constant Comparison Method

The single most important and powerful tool in GT is the Constant Comparison Method (Glaser, 2008; Glaser and Strauss, 1967; Sikolia et al., 2013; Strauss and Corbin, 1994). It is a process by which codes, as well as later emerging categories arising out of each piece of data are constantly compared against the codes and categories from the same data, other observations and with the emerging theory to produce higher levels of abstraction (Figure 5). This will be then transcribed into patterns and constantly compared, what categories are suggested by that data and subsequently with the emerging theory in mind. That is the constant comparison method - initially comparing data with data, and later comparing data with theory.

2.3 Coding and Memoing

The analysis of the body of data starts with coding, i.e. excerpting key points which are then assigned a code. Initially the code might be a phrase that summaries the key point in two or three words. Over time similar code patterns emerge. Through the constant comparison codes that share meaning can be abstracted to categories. Glaser lists several abstract theoretical coding structures which can be used as a framework to describe how the categories relate to each other. This is called Theoretical Coding (Glaser, 2008; Glaser and Strauss, 1967).

Categories and properties

A category is a theme or distinct context which embodies certain characteristics. It is interpreted in the light of the emerging theory. Properties in this case are basically sub-categories that split off finer distinctions in the categories where appropriate.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

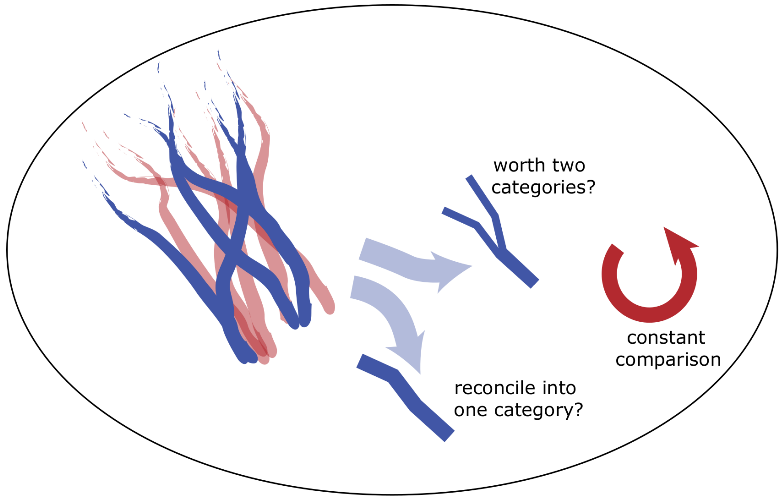

Figure 5. Applying the Constant Comparative Method on data: the result can mean absorbing the data into an existing category, or creating a new because of sufficient distinction.

Core category

Eventually several categories emerge as a result of data analysis and one category will be found to emerge with high frequency of mention. It is able to account for most variations in the data and connects meaningfully with many other categories which are emerging. This is likely to the core category. It advised to be not too eager to choose a core category or to choose too early in the data collection. However, when it becomes clear that a category is mentioned with high frequency and is well connected to other categories, it will be adopted this as the core category.

Memos

Memoing is an on-going process of writing conceptual notes throughout the GT process as whenever ideas occur to the researcher. The memos are intended to guide emergence of conceptual links between categories as the researcher notes down their relations to different categories. In GT it is assumed that the theory is concealed in the data and coding makes some of its components visible. Building on that memoing captures the relationships which link the categories to each other.

2.4 Sampling and Saturation

As categories emerge from data, the researcher seeks to add samples in a way that it further increases the diversity for the purpose of strengthening the emerging theory. This is done by defining properties of the categories and how those properties effectuate from category to category. Once the core category is established, the researcher then only codes for the core category and those categories that are closely related to the core. Collecting and interpreting more data about a particular category over time at some point does not add any new insight about that category, its properties, and its relationships. Then the category is said to have reached Theoretical Saturation and the researcher can then stop collecting data and cease coding for that category.

2.5 Sorting and Writing-Up

Once the data collection and coding are finished and categories saturated, the theoretical memos can be arranged on a conceptual level, which is called Sorting. Sorting allows forming an outline of the theory which aims to explain how the categories relate to the core category. As the theory starts to emerge, the researcher can conduct extensive literature review to see how the literature in the field relates to their emerging theory. Finally the resulting theory is written up.

In the presented research framework the process of coding, sampling and memoing informs the subsequent design process. Every iteration presents a new refined blueprint and is included into the theory and thus becomes subject to the constant comparison method. The generic design development principles are outlined in section 3.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

3 DESIGN SCIENCE RESEARCH - BACKGROUND

DSR is a particular FLlensFL of research methods in Information Systems (IS) research which fosters creation of new knowledge through the designing and implementation of innovative artefacts (Gregor and Hevner, 2013; Hevner and Chatterjee, 2010, p. 27). DSR proposes a scientific research framework for innovative software development and prescribes following criteria for the creation of new artefacts: 1) awareness, 2) suggestion, 3) development, 4) evaluation and 5) conclusion (Figure 6). These criteria are then executed along iterative stages to be improved, where the first stage merely represents the starting point for DSR theory development (Beck et al., 2013; March and Smith, 1995; Owen, 1998; Takeda et al., 1990; Vaishnavi and Kuechler, 2004).

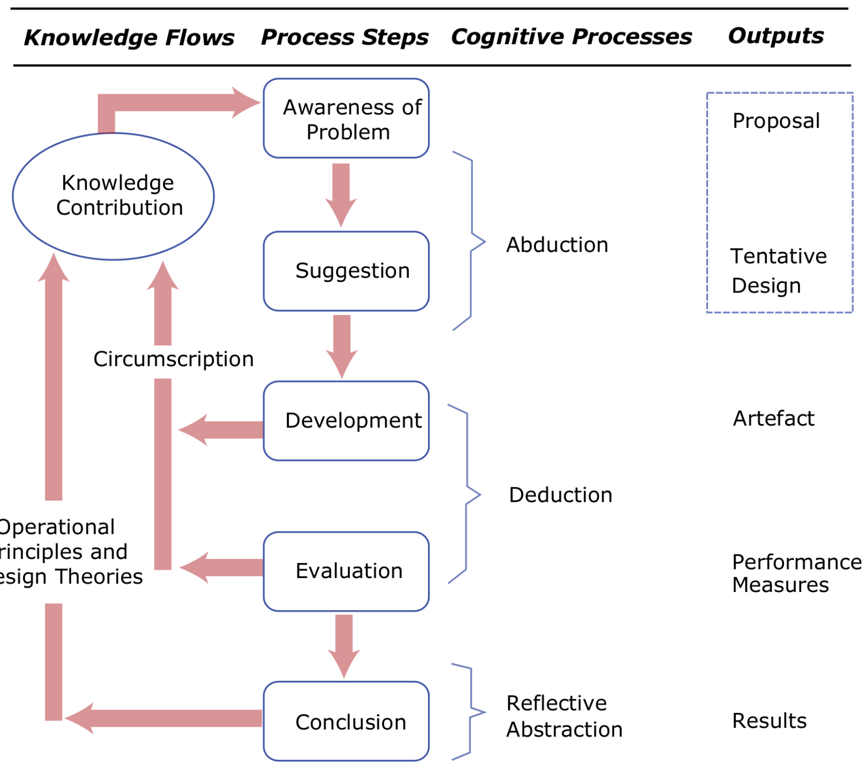

Figure 6. Design Science Research Process Model and Cognition throughout the cycle, synthesized from (Takeda et al., 1990; Vaishnavi and Kuechler, 2004; Gregor and Hevner, 2013)

3.1 Awareness

The starting point in DSR is awareness -the recognition of a problem which can be solved by using or developing new artefacts. On the one side GT and Case Study frame the awareness. On the other side, a researcher's experiences in Computer science and GIScience and the perceived discrepancies at the first encounters with geo-data management in the earth Sciences raises awareness for the need of improvement. Further considerations in the complexity of the problem are the different levels of conceptualisations:

- real world trying to be captured in data: e.g. geographical feature abstraction and data formats designed by humans to capture real world objects/features and besides standardised ways, i.e. ISO/OGC, there are different ways used to do that
- different types of problem analysis of how networked (geographical) information systems can use those data: a) how effective/efficient they are, and b) being used by humans with data in data formats, which are human artefacts themselves
- which data types and systems orchestrations are more comprehensive than others, and how well information systems that use those data can handle those data

- mentioned challenges are addressed in partially, but rarely as a whole/holistically for GIScience and Systems and Earth Sciences

3.2 Suggestion

Consequently from the awareness of the problem, suggestions are made, discussing what kind of artefact might solve the problem. Here GT informs design suggestion process. That is a particularly important aspect, because GT alone would replicate the current fuzzy state-of-the-art which is not a succinct design but a confused landscape of patchy realisations, addressing only subsets - with GI Science theory and integrated geoscientific (or environmental) modelling as the emerging desire, but no proven theory how to achieve that. GT informs the design process to improve the design theory, based on the case study data and literature and the revealed gaps in the existing theories - which results in design suggestions (Beck et al., 2013; Thornberg, 2012).

3.3 Development

Based on the suggestions, an artefact is designed and implemented in the development phase. The implementation can be pedestrian and prototypical, depending on the anticipated goal. Furthermore, the design as well its implemented artefact is supposed to be improved over multiple iterations.

3.4 Evaluation

In the evaluation phase of the design process it is checked whether the artefact solves the problem, and its strengths and weaknesses are analysed. Software interfaces and data models can be tested experimental through manual implementation if no reference benchmark is available or if the goal is not a quantitative performance improvement, but the general testing of applicability of a new design.

3.5 Conclusion

The GT process of developing theory is the main guiding input for the iterative improvement stages of the DSR method (Gregor, 2006; Gregory, 2010). Emerging categories and theoretical sampling of the similarities and the differences of information (implemented standards, addressed problems, orchestrated automated services) are consulted and reviewed with the framework design and thus the next stage of implementations (Owen, 1998; Takeda et al., 1990). Based on the FLSolution Space vs Problem Space MaturityFL matrix of Gregor and Hevner (2013) the proposed method acts around the interface between FLExaptationFL (non-trivial extension of known solutions for new problems) and FLInventionFL (new solutions for new problems) and critically distances itself from FLRoutine DesignFL (applying known solutions to known problems) (Figure 7).

In the conclusion stage results and future aspects such as open questions or plans for further development are compiled and discussed. This can well be the input for the next iteration or if the developed theory for the design becomes clear and eminent it should be written up and documented accordingly. Thus, the grounded theory is objectively refined into a novel design theory.

4 CASE STUDIES FOR DATA COLLECTION AND EVALUATION

GT provides the instruments to reason over all data as they become available. This creates a theory which is grounded in the data and which is the foundation of the design blueprints. The data is predominantly drawn from the context of the Earth Sciences - as the geographical case study areas on various scales, from local to global scale.

In Case Study research as qualitative strategy of inquiry the researcher explores events, processes and forces that generate and consume data. It is investigating phenomena in real-life settings looking for supporting empirical evidence for a well-formulated theoretical model (Yin, 1994). This implies that the a priori testing defined hypothesis is constructed from the existing literature. Also clear research questions, case selecting criteria, information about data collection methods, and clear explanation of the data analysis process need to be provided for rigorous research (Dubé and Paré, 2003).

Some of the presented characteristics of Case Study research in Information Systems contradict GT methodology, in particular the a priori paradigm. However, case studies can be used to generalise theory from. Thus Case Study research is appropriate to frame the method triangulation for this research method (Couclelis, 2009; Creswell, 2003; Creswell and Miller, 2000; Scott, 2007). For GT the described case

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

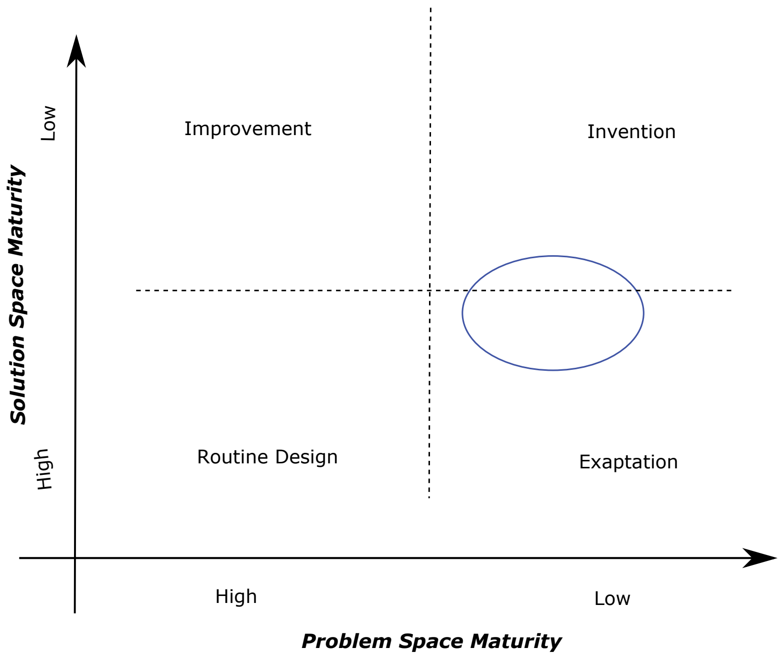

Figure 7. The DSR Knowledge Contribution Framework (Gregor and Hevner, 2013), the ellipsoid indicates where the potential of advancing theory lies for research based on the presented Grounded Design research framework

study is one of main sources of data, as well as emerging technological and theoretical developments regarding environmental data management.

Figure 8 illustrates how theory is drawn from, but also constantly reconciled (compared) with the data. An explorative literature review can also serve as data source, particular when sampling for emerging categories. Progress is achieved in design science when existing technologies are replaced by more effective ones. Design Science suggests substantive tests in the sense of natural science research. Not only must an artefact be evaluated, but the evaluation criteria themselves must be determined for the artefact in a particular environment. The aim is to compare literature to the emerging theory in the same way that data is compared to the emerging theory. In an emergent study, it is likely to be unknown at the beginning which literature will turn out to be relevant later. Thus, the constant comparison of GT remains an essential core process.

In DSR, progress is achieved when existing technologies are replaced by more effective ones. DSR suggests substantive tests in the sense of natural science research. Not only must an artefact be evaluated, but the evaluation criteria themselves must be determined for the artefact in a particular environment.

By their nature, single case studies and instantiations of a design do not meet the requirement of FLgeneralityFL that is defined for research. Typically in qualitative research some type of FLopinion per personFL data is collected, e.g. from interviews or surveys. The body of data through the presented research framework are data sets and technologies from the case studies, formal standards developed by national and international expert groups correlated with the scientific, peer-reviewed literature. Those data are human artefacts. However, all those data are generated in a either consensus oriented way or peer-reviewed embody expert knowledge on different levels, and the case studies are referred to the data sets in the Earth Sciences domains in their different available data encodings.

Finally, the DSR process flow has been extended with GT and Case Studies (Figure 9). This results in the explicit triangulated research framework that will be applied throughout this article. Rigorous principles of good qualitative research are a reproducible and transparent explanation of methods, reliable submission to the chosen methods as well making appropriate claims. These criteria allow judging the quality of the subjective research as compared with the positivists' labels: credibility (internal validity), transferability (external validity), dependability or consistency (reliability) and confirmability (objectivity) (Creswell and Miller, 2000; Creswell, 2003; Sikolia et al., 2013).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

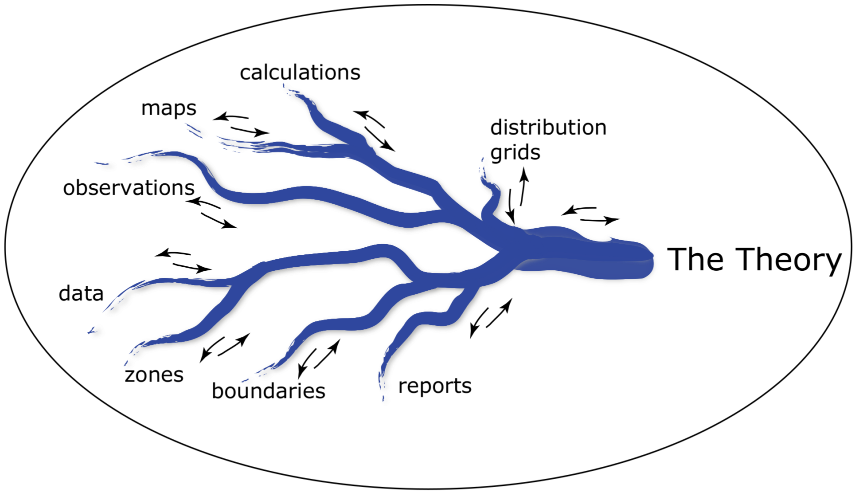

Figure 8. Theoretical rooting of GT in Data from Case Studies. The body of data through the presented research framework are data sets and technologies from the case studies, formal standards developed by national and international expert groups correlated with the scientific, peer-reviewed literature.

5 DISCUSSION

GT by itself is not intended to provide an immediate solution to a problem. Its capability is to yield better understanding of a certain research context. Therefore, it needs sufficient variety as well as depths of source data. Although the constant comparative method of GT uncovers the emerging, recurring patterns of the source data, this is not necessarily represented in statistical relations and distribution in the resulting grounded theory. The constant comparative method is a tool to abstract away the variety of the data and to capture its essence (given there is valuable essence for theory) - it is not a statistical exercise.

Design Science in return is aiming to generate new or improve theory thus creating new knowledge through moving forward. Community participation can provide essential input for the grounded theory development. Published research serves also as essential contribution of the design science feedback loop to validate and improve the overall design theory along the way.

The departure point for the presented method is existing theory and artefacts. Flaws and issues of the current situation are identified from the formed grounded theory and subsequently, reconciled with deep literature reviews. Thus, the need for improvement of some kind will then be pointed out and comprises the initial context for the Design Science iterations. From there, an executed research exercise develops its problem-solving dimension. It does so by implementing new artefacts and comparing their characteristics with the existing ones.

GT allows for, even prescribes, a deferring and delaying of theorising. This forces the researcher to work through the data and avoid premature judgements and investigation into causality. The actual theory is put together during the write-up phase from the categories. Thus, submission to GT forces the researcher to be open for patterns in data without immediately theorising reasons.

It is a human trait to try to find meaning in everything, thus every bit of information wants to be put into context and thus, be theorised. This can distract and cloud the researcher's ability to let the data flow openly and impart his ability to abstract concepts effectively. By avoiding theorising in the early stages of the research progress, new path ways are opened up from the emerging concepts, but not from the theory that might constantly being formed in the researcher's mind.

Through constant awareness the researcher strives to avoid finding reasons why a concept emerged or why it related to another concept, because this would limit the abstracting and comprehensive capability of GT. Consequently, the researcher stays 'uninvolved' with the data and theory until the method of coding and constant comparison from an abstracting exercise yielded satiated categories.

In the eventual synergistic triangulation of GT, DSR and Case Studies, GT provides the context, given that the data for GT sufficient. The case studies provide a well-defined field of input for GT. Eventually, the researcher theorises and reasons over the results and thus can yield a more comprehensive and less biased theory.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

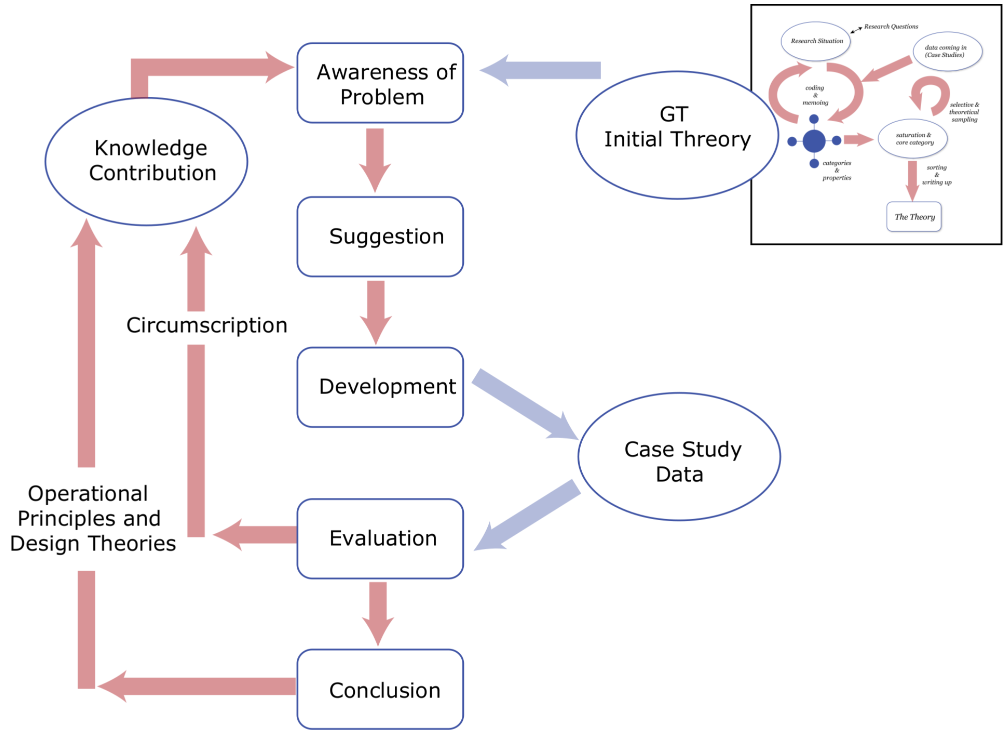

Figure 9. Explicit Design Science Research extended process flow with Grounded Theory and Case Studies

From this foundation the design process with its inferential and abductive clues make the improvement steps more clear and even reproducible. And it is not until the step towards the DSR process that the research tries to solve a gap or improve a situation. The next step is then to establish the design theory in and from produced DSR artefacts. From there the DSR process is also concerned about the implementability, practicability and efficiency.

6 CONCLUSION

In retrospection, the developed research philosophy is a successful guiding principle for the context of GIScience engineering research and development.

Through the circumscription in DSR new knowledge is generated, and implementations serve as instantiation of a new idea or partial method which then incorporated into the greater design theory for the particular problem domain.

The integration, interoperability and eventually the convergence of design patterns are the main output. This is hardly quantifiable. Through the strong GT aspects this approach has traits of a meta study method which is rooted in and subsequently, evaluated with concrete case studies.

'Grounded Design' does not conform with the empirical scientific method of falsifying a hypothesis, but instead it provides a comprehensible, documented and reproducible pathway of unlocking potential in interdisciplinary knowledge gaps through Grounded Theory – which then can be solved with empirical Design Science research and iterative and quantifiable systems design and development.

6.1 Why Grounded Theory?

The core features of GT, i.e. the Constant Comparison Method and the coding from all available data to retrieve emerging patterns that solely stem from the body of data, make it a formidable mode of inquiry for the researcher to examine the outlined transdisciplinary problem domain and develop relevant, applicable and generalizable theory. Literature review is conducted for new codes and categories as they emerge and theory is constantly compared with new data and amended or re-articulated.

6.2 Why not only Grounded Theory?

Originally developed in and for the social sciences, most literature about GT refers to the data as the interviews with participants. However, in order to develop a theory for GIScience, Information Systems, Computer Science and the Geosciences the presented 'Grounded Design' research framework proposes to draw from existing geo-data, software systems, standards/best-practices and the literature. Hence this method claims to improve GIScience theory for interdisciplinary geodata management, processing and visualisation in distributed computer systems. Thus, GT is triangulated with DSR to provide the theory for the systems design and its iterations and with Case Study research as source of data and artefacts.

6.3 Why Design Science Research?

The first and foremost feature of DSR is to provide a rigorous and structured methodology to design an artefact, i.e. create and innovative piece of software that improves or supersedes existing implementations. DSR maintains that only through implementations, .i.e. realisations of a theory in a real-world artefact, can reveal, prove or falsify properties, behaviour and relationships between components that might or might not have been theorised beforehand. Furthermore, drawing from the experiences of the implementations the design will be improved through iterations of implementations and exploring its properties, behaviours and relationships.

6.4 Why not only Design Science Research?

DSR is a relatively young research discipline and the published literature is not as exhaustive. The arrangement of a multitude of interfaces and standards and formats (for example, ISO/W3C/OGC standards) cannot easily be measured in quantitative metrics, in particular if there are no gold benchmarks available to compare with. The new framework design implements a method which has not been explicitly described like this, thus a comparison is only possible on a descriptive and explanatory basis. Also, the FLmagicalFL spark of creativity, the idea generating process to arrive at the design blueprints and their improvement is not well understood. Here, GT provides both: qualitative means of evaluation; and a documented, abductive/inductive approach that is grounded in the actual (Case Study) data.

6.5 Why the Case Study method?

The case study notion espouses the mentioned two methods with a limited set of domain (the various Earth Sciences disciplines) data constrained in time and space. Case Study results and generated theory shall be generalizable and applicable to other case studies. Case Study in GIScience has a particular nuance of geographical case study areas and thus serves as an explicit body of data for GT and DSR.

6.6 Why not only the Case Study method?

Case Study research in a rigorous scientific inquiry requires fixed a priori defined research questions and quantitative measures for validation to prove or falsify the a priori hypothesis. However, in the presented 'Grounded Design' research framework the triangulation of GT and DSR with both having a posteriori theory generating paradigm, the Case Study method is not executed on its own. It serves as an explicit field of inquiry for GT and DSR, and by its nature, the generated theory from within the constraints of this Case Study are sought to be generalizable to other (in particular geographical) case studies or even globally.

We described an explicit research method how to design and develop GIS and SDI systems on the premiss that the various Earth Sciences disciplines' data and models need to be integrated through explicit GIScience principles.

Finally, in an analogy, the GT processes of sampling and constant comparison could be depicted in a cognitive conceptual space. After Tobler's First Law of Geography (Tobler, 1970), closer things are more related to each other than more distant things. Logically, similar concepts are closer to each other in our mentioned cognitive space. GT aims to produce theory through letting patterns emerge from the data. In GIScience alike, patterns arise from the data through geographic analysis which characterises the spatial relationships in the data. Therefore, an analogy could be made between the constant comparison method of GT and spatial analysis. In conclusion, GT is a fitting method for spatial thinking in conceptual space. This refined research method can now be applied to develop an advanced data infrastructure design theory on the premise that the various Earth Sciences disciplines' data and models can be integrated through explicit GIScience principles.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

FUNDING

This work was supported by the six year funding (07/2011–06/2017) of the Ministry of Business, Innovation, and Employment (MBIE), New Zealand, for the SMART Aquifer Characterisation Programme (SAC) under contract number C05X1102, the Marie Skłodowska-Curie Actions individual fellowships offered by the Horizon 2020 Programme under REA, grant agreement number 660391, and the Mobilitas Pluss Postdoctoral Researcher Grant (MOBJD) number MOBJD233 of the Estonian Research Council (ETAG).

REFERENCES

Bandaragoda, C., Tarboton, D., and Maidment, D. (2006). Hydrology's efforts toward the cyberfrontier. *Eos, Transactions American Geophysical Union*, 87(1):2–6.

Beck, R., Weber, S., and Gregory, R. W. (2013). Theory-generating design science research. *Information Systems Frontiers*, 15(4):637–651.

Bhaskar, R. (1998). Philosophy and Scientific Realism. In *Critical Realism: Essential Readings*, pages 16–47.

Birks, D. F., Fernandez, W., Levina, N., and Nasirin, S. (2013). Grounded theory method in information systems research: its nature, diversity and opportunities. *European Journal of Information Systems*, 22(1):1–8.

Blaschke, T., Strobl, J., and Donert, K. (2011). Geographic Information Science: Building a Doctoral Programme Integrating Interdisciplinary Concepts and Methods. *Procedia - Social and Behavioral Sciences*, 21:139–146.

Couclelis, H. (2009). Ontology, Epistemology, Teleology: Triangulating Geographic Information Science. In *Research Trends in Geographic Information Science*, pages 3–15.

Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative. and Mixed Methods Approaches*. SAGE Publications, Nebraska, 2nd editio edition.

Creswell, J. W. and Miller, D. (2000). Determining validity in qualitative inquiry. *Theory into practice*, 39(July 2015):124–130.

Dobson, P. J. (2002). Critical realism and information systems research : why bother with philosophy ? *Information Research*, 7(2):1–16.

Dubé, L. and Paré, G. (2003). Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations. *MIS Quarterly*, 27(4):597–635.

Easterbrook, S., Singer, J., Storey, M.-A., and Damian, D. (2008). Selecting Empirical Methods for Software Engineering Research. In *Guide to Advanced Empirical Software Engineering*, pages 285–311. Springer London, London.

Feast, L. and Melles, G. (2010). Epistemological Positions in Design Research: A Brief Review of the Literature. In *Proceedings of the International Conference on Design Education (ConnectED)*, number November, pages 1–5.

Gahegan, M., Luo, J., Weaver, S. D., Pike, W., and Banchuen, T. (2009). Connecting GEON: Making sense of the myriad resources, researchers and concepts that comprise a geoscience cyberinfrastructure. *Computers & Geosciences*, 35(4):836–854.

Glaser, B. G. (2008). Conceptualization: On theory and theorizing using grounded theory. *International Journal of Qualitative Methods*, 1:23–38.

Glaser, B. G. and Strauss, A. L. (1967). The discovery of grounded theory. *International Journal of Qualitative Methods*, 5:1–10.

Goles, T. and Hirschheim, R. (2000). The paradigm is dead, the paradigm is dead... long live the paradigm: the legacy of Burrell and Morgan. *Omega*, 28(3):249–268.

Goodchild, M. F. (1992). Geographical information science. *International journal of geographical information systems*, 6(1):31–45.

Goodchild, M. F. (2006). GIScience Ten Years After Ground Truth. *Transactions in GIS*, 10(5):687–692.

Goodchild, M. F. (2009). Geographic information systems and science: today and tomorrow. *Procedia Earth and Planetary Science*, 15(1):1037–1043.

Gray, D. E. (2014). Theoretical perspectives and research methodologies. In *Doing research in the real world*, pages 16–38.

Gregor, S. (2002). Design Theory in Information Systems. *Australasian Journal of Information Systems*, (December):14–22.
Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly*, 30(3):611–642.
Gregor, S. and Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, 37(2):337–355.
Gregory, R. W. (2010). Design Science Research and the Grounded Theory Method : Characteristics, Differences , and Complementary Uses. *Proceedings of the 18th European Conference on Information Systems (ECIS)*, pages 111–127.
Hevner, A. and Chatterjee, S. (2010). *Design Research in Information Systems: Theory and Practice*, volume 22.
Hirschheim, R. a. (1985). Information Systems Epistemology: An Historical Perspective. *Research Methods in Information Systems*, pages 13–38.
Holliday, A. (2010). Analysing qualitative data. In Phakiti, B. and Paltridge, A., editors, *Continuum companion to research methods in applied linguistics*, chapter Analysing, pages 98–110. Continuum International Publishing Group, London.
March, S. T. and Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4):251–266.
Mark, D. M. (2002). Geographic Information Science : Defining the Field. *Foundations of Geographic Information Science*, pages 3–18.
Mingers, J. (2004). Real-izing information systems: critical realism as an underpinning philosophy for information systems. *Information and Organization*, 14(2):87–103.
Owen, C. (1998). Design research: building the knowledge base. *Design Studies*, 19(1):9–20.
Pickles, J. (1997). Tool or Science?GIS, Technoscience, and the Theoretical Turn. *Annals of the Association of American Geographers*, 87(2):363–372.
Porra, J., Hirschheim, R., and Parks, M. S. (2014). The Historical Research Method and Information Systems Research. *Journal of the Association for Information Systems*, 15(9):536–576.
Saunders, M., Lewis, P., and Thornhill, A. (2012). *Research Methods for Business Students*.
Scott, D. (2007). Resolving the quantitative- qualititive dilemma: A critical realist approach. *Journal of Research and Method in Education*, 30(1):3–17.
Sikolia, D., Biros, D., Mason, M., Weiser, M., Sikolia, D., and Weiser, M. (2013). Trustworthiness of grounded theory methodology research in information systems. In *Proceedings of the Eighth Midwest Association for Information Systems Conference*, volume 24, pages 1–5, Illinois. Association for information systems.
Strauss, A. L. and Corbin, J. M. (1994). Grounded theory methodology. In *Handbook of qualitative research*, pages 273–285.
Takeda, H., Veerkamp, P., Tomiyama, T., and Yoshikawa, H. (1990). Modeling Design Processes. *AI Magazine*, 11:37–48.
Thornberg, R. (2012). Informed Grounded Theory. *Scandinavian Journal of Educational Research*, 56(3):243–259.
Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240.
Tress, B., Tress, G., van Apeldoorn, R., and Fry, G. (2003). *Interdisciplinarity and Transdisciplinary Landscape Studies: Potentials and Limitations*. Wageningen.
Urquhart, C., Lehmann, H., and Myers, M. D. (2010). Putting the 'theory' back into grounded theory: Guidelines for grounded theory studies in information systems. *Information Systems Journal*, 20(4):357–381.
Vaishnavi, V. and Kuechler, B. (2004). Design Science Research in Information Systems Overview of Design Science Research. *Ais*, page 45.
Walsh, I. (2014). Using Grounded Theory to Avoid Research Misconduct in Management Science.
Yang, C., Raskin, R., Goodchild, M. F., and Gahegan, M. (2010). Geospatial Cyberinfrastructure: Past, present and future. *Computers, Environment and Urban Systems*, 34(4):264–277.
Yin, R. K. (1994). Case Study Research: Design and Methods. *Applied Social Research Methods Series*, 5:219.
Zachariadis, M., Scott, S., and Barrett, M. (2010). Exploring critical realism as the theoretical foundation of mixed-method research: Evidence from the economics of IS innovations. *Working Paper Series -*

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27822v1 | CC BY 4.0 Open Access | rec: 25 Jun 2019, publ: 25 Jun 2019

Cambridge Business School.