



UNIVERSITY OF GOTHENBURG

Tool assisted traffic accident causation analysis

An action design research approach

Bachelor of Science Thesis in Software Engineering and Management

ADAM DEBBICHE

ANDERS TREPTOW

YUWEN HE

University of Gothenburg
Chalmers University of Technology
Department of Computer Science and Engineering
Göteborg, Sweden, June 2012

The Author grants to Chalmers University of Technology and University of Gothenburg the non-exclusive right to publish the Work electronically and in a non-commercial purpose make it accessible on the Internet.

The Author warrants that he/she is the author to the Work, and warrants that the Work does not contain text, pictures or other material that violates copyright law.

The Author shall, when transferring the rights of the Work to a third party (for example a publisher or a company), acknowledge the third party about this agreement. If the Author has signed a copyright agreement with a third party regarding the Work, the Author warrants hereby that he/she has obtained any necessary permission from this third party to let Chalmers University of Technology and University of Gothenburg store the Work electronically and make it accessible on the Internet.

Tool assisted traffic accident causation analysis

An action design research approach

Adam Debbiche

Anders Treptow

Yuwen He

©Adam Debbiche, June 2012

©Anders Treptow, June 2012

©Yuwen He, June 2012

Examiner: Helena Holmström Olsson

University of Gothenburg

Chalmers University of Technology

Department of Computer Science and Engineering

SE-412 96 Göteborg

Sweden

Telephone + 46 (0)31-772 1000

Department of Computer Science and Engineering

Göteborg, Sweden June 2012

ABSTRACT

The field of accident causation analysis deals with the analysis of data gathered after traffic accidents. The goal is to develop new techniques to prevent future accidents and save more human lives. This paper, through an action design research approach at SAFER, provides a tool that helps in identifying causation patterns from accident data presented in the form of charts. The paper examines different analysis techniques of accident causation data, as well as show how action design research was used in this case. The paper also examines the effects of ADR on the organization as well as the implication of adopting user involvement.

1- INTRODUCTION

The amount of data stored today is growing at a high rate and there is no sign of this slowing down anytime soon. It is therefore important to find technological solutions that allow the exploitation of large sets of data. Data mining, i.e. the extraction and discovery of previously unknown yet possibly valuable information from large sets of data, is a new field that is increasingly being used today and has emerged as a major research domain (Nirkhi, 2010). Within data mining, different techniques are used to analyze large sets of data. Nirkhi (2010) argues that artificial neural networks, decision trees and genetic algorithms are among the most popular approaches. Some domains like finances rely more on neural networks to analyze data due to their ability to discover patterns and predict future behavior which assists companies in strategic planning (Zhang and Zhou, 2004). Similarly, domains where graph theory is a common occurrence require data mining approaches related to structural pattern discovery in graphs (Wang et al., 2002). This implies that the characteristics and goals of the domain are applicable when assessing the feasibility of specific implementations of data analysis tools. In the field of accident causation analysis, there exist different techniques of analyzing data gathered at accident scenes. Thus, the feasibility of accident causation analysis tool depends on the analysis approach used and the said domain's characteristics and goals.

To this end, our study focuses on a specific domain of data mining: traffic accident causation data analysis. Our collaborating organization (SAFER) has a history of manually transforming and

visualizing chains of events during their accident analysis work. According to Kotter (1995), introducing change in an organization is a long process. Hence, moving from well-established manual processes to automated data mining is sometimes a daunting task in practice. Our focus in this paper is specifically oriented towards the design and development of an automated data analysis tool. Within this specific setting, our study focuses on root cause analysis of actual traffic accidents i.e. pre-crash scenarios. Currently, there exists a formal method for retrieving, classifying and analyzing the data collected at crash sites, and this method is referred to as DREAM (Ljung, 2002). This method suggests the development of charts by looking at multiple viewpoints on the causes of an accident. Each accident can produce multiple charts.

At present, the practitioners have access to a database with a large number of charts from different accident cases. They can choose to combine any number of charts (from the same or different accidents) which are then aggregated and presented in the form of a graphical representation known as an aggregated DREAM chart. This is a tedious process in terms of human resources as much time is spent on composing these charts which are essentially necessary in order to properly analyze traffic accident causation. The specific problem is therefore investigating in this study how SAFER, as one specific example of an organization (within the domain of traffic accident analysis) faces challenges of analyzing large sets of data, and that they are currently following repetitive but solid manual analysis tasks. Based on this, our research objective is to: *assess the feasibility of an automated computer aided analysis tool for data analysis of traffic accident data through a prototype implementation.*

To approach this problem in a way that also illustrates how automated data mining tool support may be developed, an action design research (ADR) approach (Sein et al., 2011) will be used. The tool strives to automate the – today manual – chart aggregation process through a theory driven prototype that will automate the charts aggregation process which is considered a practical and theoretical contribution. We also make a

methodological contribution by being an early adopter of action design research (ADR).

The paper continues with a related literature section where different traffic accident analysis methods are described as well as our theory on how to build a DREAM based tool. After that we introduce our research method (ADR), explain why it was chosen and how it was used. In the data and discussion part, the results are presented and discussed. The paper ends with a conclusion and suggestions for further research.

2- RELATED LITERATURE

In this section, we present literature related to our field of study: traffic accident data causation analysis. In section 2.1, different techniques of accident data causation analysis previously developed are presented. In section 2.2, we write about how to build a good DREAM based tool through user involvement. These sections will, together with reflections related to our ADR research method, be central aspects of our discussion later in the paper, and served as guidance to the development work of our data analysis tool.

2.1- TRAFFIC ACCIDENT ANALYSIS METHODS AND TECHNIQUES

With regards to different analysis methods of accidents, Otte et al. (2009) presented a method known as Accident Causation Analysis with Seven Steps (ACASS). It allows analyzing and collecting causation factors of traffic accidents. According to the authors, ACASS can appropriately define the human errors of the actors involved in a traffic accident. The method contains a model that allows collecting the important information at an accident scene. An approach to interview people involved in an accident is introduced so that the human causation factors are obtained. This is achieved through an analysis system (in seven steps) which takes into account the chronological order from observation (recognizing the danger) to operation (responding to the danger). Additionally, ACASS groups the accident causation factors into three groups: human factors, technical factors from the vehicle and environment and infrastructure (see figure 1). Each group contains categories which in turn contain more specific criteria that specify the factor within the category (Otte et al., 2009). Also, ACASS allows data collected from a scene to be submitted to a database.

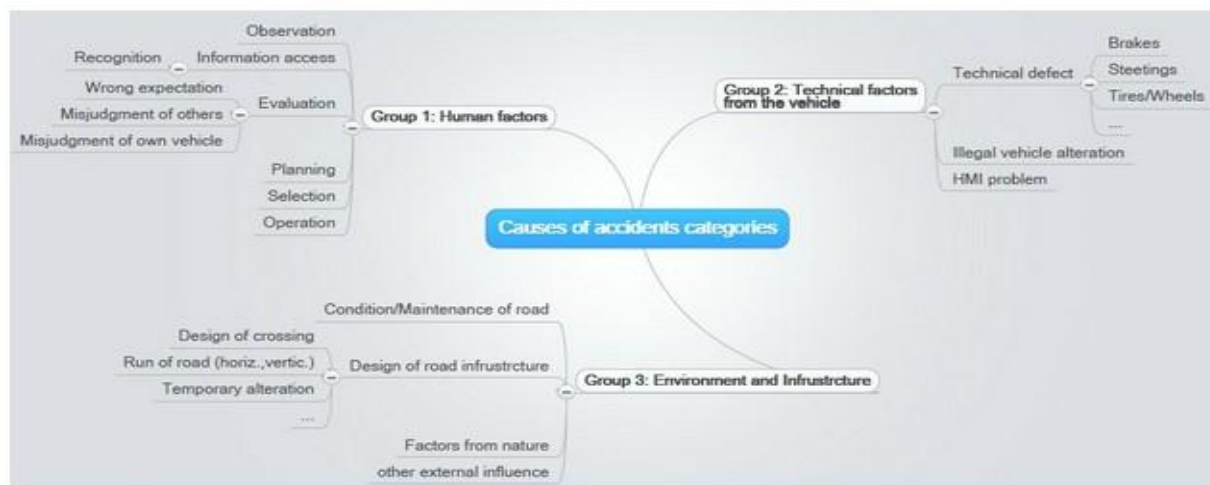


Figure 1. Structural-analytical view of causes of accidents in the human-vehicle-environment-model (Otte et al., 2009)

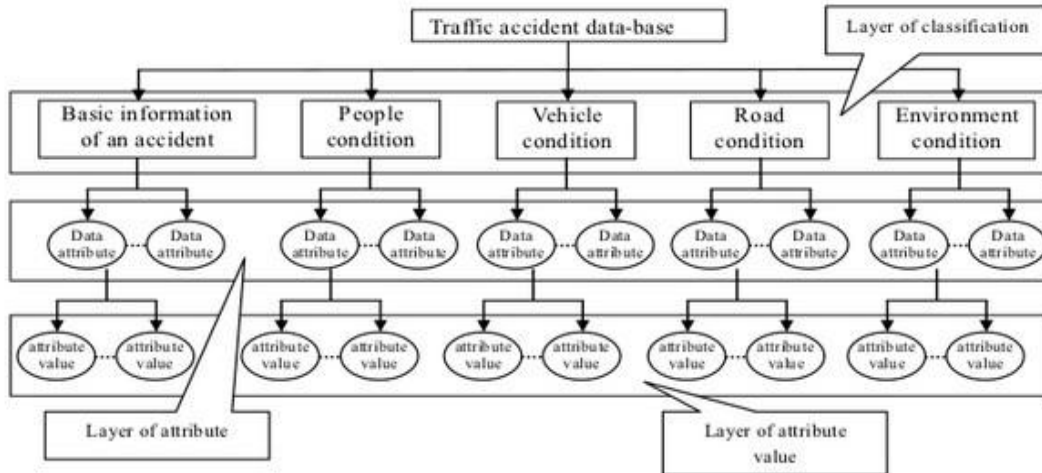


Figure 2. Diagram of Traffic Accident Data Structure (Xi et al., 2010)

Xi et al. (2010) have developed another accident causation analysis method based on traffic accident information system (in China). Since the method is based on accident data recorded in a database, the authors focus on the characteristics and classification of traffic accident data. They argue that each accident record includes multiple data attributes and that a data attribute is organized according to five aspects: basic information of an accident, information of relevant people, vehicle information, road information and environment information. Consequently, Xi et al. (2010) identify two layers of data attributes in the traffic accident database as seen in figure 2.

In their work, they suggest a method that provides quantitative analysis for the contribution of accident analysis data taken from a database. Two formulas are used; the first one (1) calculates the importance of four attributes including people, vehicle, road and environment (Layer 1). The second (2) formula calculates the importance classification of attribute (Layer 2). The result of both formulas (1) and (2) is always between 1 and 4 (1: unimportant, 2: general, 3: important, 4: very important). These results in every classification and attribute being assigned an importance value (1-4). The authors state that the result of the method can only be an important foundation to formulate improved strategy for traffic safety (Xi et al., 2010).

Another accident causation method is DREAM, a method first developed by Ljung in 2002 (Ljung et al., 2007). As other analysis methods, it organizes accident data using a classified schema of contributing factors of accidents in a systematic way. DREAM is the adaptation of the Cognitive Reliability and Error Analysis Method (CREAM) (Hollnagel, 1998) with the aim to suit the road traffic domain. The original goal of DREAM was to identify traffic situations for which the development of technical solutions had the potential to prevent future accidents (Warner et al., 2008). It was thus used to guide the analysis process within different types of technical solutions targeting different areas of accident avoidance. Nowadays, the focus of DREAM is however mostly to identify interactive systems for risk avoidance (Warner et al., 2008).

After accident investigators collect data from a scene of an accident (through interviews and observations). DREAM is initially used to develop an accident model from the data collected at the scene consisting of the human, vehicle and traffic environment (technology) and the organization (Ljung, 2002). Once the accident model is created, the practitioner uses DREAM's classification scheme to begin drawing the chart. A DREAM chart is composed of an observable effect - known as a phenotype and contributing factors to the observable effect (genotypes) according to Ljung (2002). The DREAM manual offers a list of all the possible phenotypes and

genotypes and how they are linked. A chart is created for each actor involved in a car accident. The goal of DREAM is thus “...to make it possible to systematically classify and store accident causation information which has been gathered through in-depth investigations by providing a structured way of sorting the causes behind the accident into a set of formally defined categories of contributing factors”. (Warner et al. 2008, p.7)

The latest version of DREAM is version 3.0 at the time of writing with a newer version planned for release during the first or second quarter of 2012.

With regards to the three different methods presented in this section, Table 1 shows the different characteristics of each method:

Table 1. Characteristics of each method reviewed in this paper

<i>Methods</i>	<i>Characteristics</i>
Accident Causation Analysis with Seven Steps (Otte et al., 2009)	<ul style="list-style-type: none"> ● Proposes an approach to interview people with the goal to extract human causation errors ● Data collected can be entered in a database ● Divides accident causation factors into three different categories
Accident Causation Analysis Method Based On Traffic Accident Information System (Xi et al., 2010)	<ul style="list-style-type: none"> ● Based on traffic accident information system (databases) ● Focuses on quantitative analysis (statistical)
Driving Reliability and Error Analysis Method (Ljung, 2002)	<ul style="list-style-type: none"> ● Focuses on identifying interactive systems for risk avoidance ● Visualization of accident schema in the form of charts ● Aggregation of multiple charts to discover patterns that cause certain types of accidents ● An organizer of explanations - not a provider

Given the alternatives presented here, we decided to focus on DREAM: According to unpublished internal reports (Björklund et al., 2007), SAFER conducted comparison studies to determine which method should be used in two of its projects. The goal of the first project was to investigate which pre-crash method would be suitable in the Investigation Network and Traffic Accident Techniques (INTACT) project at Chalmers. This led to the exclusion of some methods from the start. Each method was first evaluated by one of the group members involved in the study then discussed within a group. The discussion was based on a set of guiding principles identified in the beginning. At the end of the project, the team presented their recommendations. It was suggested that DREAM should be used along with another accident analysis method called Sequentially Timed Events Plotting (STEP) as they have a great potential of complementing each other and offer a clear description of events and factors leading to an accident. Moreover, DREAM was found to be compatible with the guiding principles:

- It offers case and aggregated analysis
- Has a theoretically described accident model and a clearly described analysis method
- No guilt. The goal is not to determine who committed traffic violations
- Several concurrent levels of analysis
- Predefined accident factors
- Counter measures: the goal of DREAM is to develop counter measures in order to prevent accidents
- Can be implemented in a database
- Interview with witnesses and drivers is an important part of the data collection procedure

In the second project, SAFER also conducted a study (SAFER, 2011) along with other partner organizations to determine which of the following methods were suitable in their “Road Safety Data, Collection, Transfer and Analysis” (DaCoTA) project: DREAM, ACASS and HFF. During six months the different methods were compared by first setting up a coding exercise where each participant in the study coded five examples cases once with each method. Next, each coder filled in a questionnaire to evaluate

their experience with each method. After the coding exercise and questionnaire, all SAFER partners were asked to identify their favorite coding system. The results of the coding exercise showed that DREAM had higher conformity (65%). The questionnaire showed that DREAM had the highest conformity as well as the most explanatory manual. The results of stating preferred method showed that most partners preferred DREAM while others wanted to see some elements from ACASS and HFF included in DREAM. The internal report not only concludes that the results are in favor of DREAM but also that the method is supported by the European Commission. However, the report also noted that some changes to DREAM should be made.

2.2- USER INVOLVEMENT AS A QUALITY ASSURANCE AND DEVELOPMENT STRATEGY

As mentioned above, the prototype to be developed is based on the DREAM method. With this in mind, it is necessary to ensure that the development approach conforms to user needs. Previous to our research, the collaborating organization lacked a clear understanding of the exact needs they had or the potential of this implementation. Therefore, relying heavily on user involvement to guarantee the appropriateness of the implementation seems reasonable. User involvement is a popular issue that is currently discussed in the software industry. The reasons for this include the negative feedback from customers about products during/after development, dissatisfaction with the software, cost issues and the instability of marketing (Majid et al., 2010). Majid et al. (2010) place strong emphasis on this as they argue that unsuccessful software products are always based on the unacceptable and faulty design. In our case, the communication with practitioners at SAFER is frequent enough to allow user involvement in a way that is likely to have a positive effect on the design and implementation of the prototype, in particular for capturing requirements and gathering feedback in the iterative development phases.

Das (2007) points out that a measure of a successful software product is the degree of the design fulfilling the customer's requirements. He therefore suggests user involvement to be adopted in the software requirement engineering area. It can be used to help developers identify stakeholders and their needs, and

documenting the specifications. Relying on user involvement thus has a positive effect on the success of software development and user satisfaction (Das, 2007).

However, several questions remain, including how and why user involvement works in practice. Majid et al. (2010) have conducted a survey on user involvement in software development life cycles. They investigated to which extent users' involvement should be in the development cycle. Their initial literature studies state that due to the user interaction including information and technology exchange, each phase in software development must pay attention to user involvement to ensure quality (Majid et al., 2010). The result of the survey showed that the requirement analysis stage shares the highest percentage of user involvement, 77.42%, followed by testing and deployment stage with 64.52%. The involvement percentages of project selection and planning stage, as well as system design stage were less than the first two ones, 54.84% and 35.48% respectively. The development stage came in last with only 16.13% in total. Also, the result showed that the involvement of users focused more on the functional requirements rather than non-functional requirements. Thus, they drew the conclusion that the degree of user involvement varies at each stage of the development life cycle and software engineers should focus on real users' need in the overall software lifecycle (Majid et al., 2010).

Heiskari and Lehtola (2009) present a case study in a company producing software solutions to investigate the state of user involvement in practice. They point out that there are several risks and challenges when having users involvement in the development process. For example in agile methods, users are encouraged to participate with developers, but the main focus of agile methods is to deliver a product instead of being user-centered (Heiskari and Lehtola, 2009). Thus, the goal of the case study is to provide effective and efficient way to adopt user involvement by understanding how different departments which have different functions in the organization involve users in practice. Semi-structured interviews were conducted with various people in different departments and recordings of interviews were translated into textual descriptions. The authors

present several challenges found in practice such as little information about the user, integrating user knowledge into the existing processes, understanding the big picture before going into details and very little interaction between the end users. They argue that the current state in companies is that users are involved in different departments in several ways, but it is difficult to make sure whether users influence the actual development process or the product (Heiskari and Lehtola, 2009). They conclude that the main principle of user involvement is to gain a thoughtful understanding of user needs and fulfill those requirements in an effective and efficient way during development, not necessarily to have users participate with developers (Heiskari and Lehtola, 2009).

With this theory of user involvement, this study will be conducted using a method called action design research. The outcome of adopting user involvement will be reflected upon in the discussion.

3- ACTION DESIGN RESEARCH (ADR)

3.1- WHAT IS ACTION DESIGN RESEARCH?

Action design research (ADR) (Sein et al., 2011) is relatively new and has its roots from both design research (Hevner et al., 2004) and action research (Susman and Evered, 1978). When defining action design research, it is important to consider both design research and action research: Design research involves developing an ensemble of IT artifacts to solve a practical problem, the design of the artifact is in this case the focus of the research process and the organizational intervention is considered secondary. However, in action research, the researcher is often part of the team in the organization where the research project is taking place as opposed to having a more observational role (Sein et al., 2011). Action design research tries to bring the best of these two methods and bridge the gap between research and practice.

According to Sein et al., (2011), current design research methods pay little to no attention to the shaping of the artifact by the organizational context. Also, current design research methods assign the evaluation to a separate phase after the building of the artifact. Sein et al. (2011, p. 37) write that "...they

value technological rigor at the cost of organizational relevance, and fail to recognize that the artifact emerges from interaction with the organizational context even when its initial design is guided by the researchers' intent.". Although there has been earlier attempts to combine organizational intervention into design research methods (Iivari, 2007), they still separate the different stages (intervention, building and evaluation).

To this end, action design research is a method that seeks to generate design knowledge by building an innovative IT artifact with the organizational context from which it emerges constantly in mind. Table 2 demonstrates the different characteristics of action research, design research and action design research:

Table 2. Difference between AR, DR and ADR

Action research	<ul style="list-style-type: none"> ● Researcher is tasked to solve an immediate problem in an organization through intervention ● Involves theory generation (Sein et al, 2011) ● Tries to link theory with practice
Design research	<ul style="list-style-type: none"> ● Seeks to develop an IT artifact to address a class of problems ● Development is followed by evaluation. "build and then evaluate" ● Organizational intervention is secondary
Action design research	<ul style="list-style-type: none"> ● Recognition of the organizational setting from which the need of an IT artifact is born. ● The stages of building, intervention and evaluation are inseparable. ● Aims at building innovative artifacts in an organizational context and learning from the intervention while solving a problem (Sein et al, 2011)

3.2- RESEARCH SETTING

The study was conducted in close collaboration with SAFER which is a joint research unit between the Swedish automotive industry, academia, and authorities where these partners cooperate within the field of vehicle and traffic safety. We had daily access to the practitioners involved with accidents causation data analysis and so could interact with

them when needed. The practitioners were directly dealing with the problem to be solved: automating the process of aggregating DREAM charts through the development of a prototype. Meetings to discuss the functional and nonfunctional requirements were held. Potential users were also involved through demos of the prototype in order to get feedback and suggestions. Section 3.4 explains more in detail about the development and the interaction with SAFER in regards to action design research.

Another important aspect to mention here is the difference between researchers, practitioners and investigators. While the people we collaborate with are researchers at SAFER, conducting investigations at the scene of an accident is part of their research. Therefore, investigators and practitioners refer to the same group of people (researchers at SAFER). In this paper, they are called practitioners and the term researchers refer to the authors of this study.

3.3- MOTIVATION FOR USING ADR

Action design research was selected as research method given that two things were explicit from the start of the study. First, SAFER were looking for a prototype implementation of a tool designed to assist in traffic accident analysis. Second, SAFER wanted to be involved in the decision making of the development at all stages of the process. These two main reasons were later further supported by the fact that an iterative development process was adopted. ADR in itself is based on an iterative approach which makes it a good fit for the development process.

Consequently, as these three attributes of ADR match what SAFER wanted from the study; action design research was identified as a highly suitable candidate. Action research (AR) (Susman and Evered, 1978) was considered based on the collaborative element that is central there also. However, Olsson (2011) argues that action research is iterative more in terms of whole cycles of research and not as much within each cycle, and also not as design artifact centric. Therefore ADR represents a more suitable choice. While ADR is a newly formed research method, the fact that it is informed by both the highly established but strictly design oriented design research (DR) (Hevner et al, 2004), and ARs collaborative elements, meant that relying on ADR also allowed this research

to contribute as an early adopter of the novel research method.

The goal was then to develop a novel prototype shaped not only by our design principles but also the organization from which it emerges (SAFER). Therefore, it was natural to adopt ADR as research method because of its emphasis on organizational context which is considered a key characteristic.

Another reason for choosing ADR is the iterative process of evaluation of the artifact; In action research, the evaluation phase is done after the development. ADR on the other hand emphasizes that evaluating the IT artifact and the intervention in the organization should be done constantly as Sein et al. (2011) argue. The decision to involve the user in an iterative manner during each sprint (see Section 2.2) meant having a research method that stresses the importance of the interwoven activities of building, intervention and evaluation worked well with the quality assurance strategy adopted.

3.4- HOW WE USED ADR

This study first started with the problem formulation phase where the research problem was perceived. In our case, it was SAFER who perceived the need for the prototype. According to Sein et al. (2011), identifying and conceptualizing the research problem has to be done first. An initial meeting was held with the practitioners at SAFER where they explained the aggregation process used. We diagnosed a resource heavy and time consuming process as the problem with the current chart aggregation approach. According to SAFER, if they need to aggregate charts built using DREAM they have to dedicate a lot of time and manual labor to work using different tools such as Microsoft Excel and additional manual analytical work. The current approach was also lacking many features that SAFER wanted such as chart manipulation and visualization options. We concluded that fixing the current process of chart aggregation was not feasible without the introduction of a computer aided analysis tool. Based on this, the goal was to build a prototype tool that would not only assist with the charts aggregation process but also respond to the requests that the practitioners had through the implementation of various features that were identified.

Based on the first phase, the building, intervention and evaluation (BIE) stage was started by foreseeing an automated computer aided analysis tool that would

help SAFER with chart aggregation by increasing the speed and saving time through human resources reduction. According to Sein et al. (2011), this stage is where the building of the artifact, intervention in the organization and the evaluation take place concurrently. As mentioned in section 2, the approach SAFER uses depend on the DREAM method (Ljung, 2002) from the start when data is collected at the scene of an accident to the end when classification of data into charts occurs. Sein et al. (2011) identified a principle in ADR called theory-ingrained artifact which means that the artifact to be developed should be informed by theories. Based on this, the DREAM method itself must be used as a theoretical driver for the development of the prototype:

According to Sein et al. (2011), two types of theories are best suited for action design research as defined by Gregor (2006):

- Theory for explaining and predicting which implies understanding the cause and prediction while describing the theoretical constructs and the relationships between them.
- Theory for design and action which is concerned with how to build something. This type focuses primarily on the theoretical knowledge that is used in the development of software systems.

The DREAM method is used as a theoretical driver in our case since it explains how to proceed with the development of the prototype itself (in theory) in terms of implementation. This is in fact consistent with the Gregor's (2006) definition of theory for design and action. Therefore, using DREAM principles as theoretical knowledge in the development makes our theory a design and action one. Indeed, the prototype implementation will follow the same components of DREAM such as charts and aggregated charts as well as DREAM concepts like phenotypes and genotypes. Basing the prototype on the DREAM method would provide a tool that works in a way that is familiar to the practitioners since they were already working with DREAM manually. In addition to that, DREAM is widely used already by SAFER and its partners in Europe. However, no evidence of a computer tool

that implements DREAM was found. By making this tool available, any researcher that is familiar with DREAM would benefit from it in their work.

Once the decision had been made on what theoretical lens to rely on during the prototype development, we continued into the iterative BIE stage. Documents relevant to how the DREAM method works were collected. Understanding how SAFER used DREAM to organize the data was vital in order to base the prototype on it. At this point, the focus shifted to the iterative process of the BIE stage. Sein et al. (2011) argue that this phase determines the source of innovation which could result from the artifact design or the organizational intervention. The authors identify an IT-dominant BIE and an organization-dominant BIE. The IT dominant BIE is recommended if the goal is to create an innovative design.

An IT-dominant BIE was picked because our intervention in the organization is IT-centric. Furthermore, our intervention is low level (accident causation analysis) as opposed to an organizational wide one favored by an organization dominant BIE. Additionally, Sein et al. (2011) note that in an IT-dominant BIE the practitioners should first influence the design which is what we do at this stage as described earlier (continuous feedback). Second, the early versions of the design serve as lightweight interventions in a limited context (Sein et al., 2011). Indeed, our early iterations of the design were only limited in organizational intervention to the practitioners that were directly involved in accident causation analysis in the organization. Likely, only at a later stage will the more mature versions of the prototype be introduced to a wider set of practitioners for more refinement through use in the organizational setting and context. Figure 3 below shows the generic schema of an IT-dominant BIE.

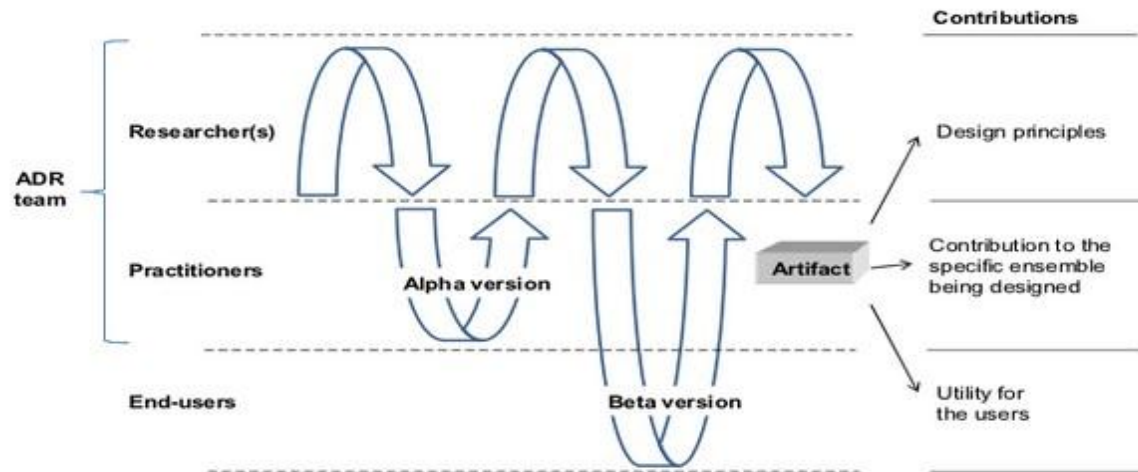


Figure 3. IT-Dominant BIE (Sein et al., 2011)

With the IT-dominant BIE acting as the design continuum at this stage, an initial design of the prototype was developed and then revised and shaped by SAFER before the implementation started. The process of shaping the prototype and developing it was then performed in an iterative process that continued throughout the design cycles that not only involved us but also the practitioners. The practitioners that could be seen as the final users too were continuously involved in each iteration where they provided feedback on the features that had just been implemented as well as guidance on things that needed to be changed and in what way. Live demos were constantly conducted to show how the prototype worked. We also focused on the principle of concurrent evaluation as opposed to it being a separate stage which is also another important principle of ADR. The head practitioner was heavily involved when the artifact was in the alpha stage of development. Subsequently, the prototype evolved (through organizational intervention) into a more mature artifact (beta version) which allowed it to be deployed to a wider organizational context. The objective of this wider evaluation is the continuous refinement of the tool (Sein et al., 2011).

The third step of ADR is reflection and learning, the objective of stage is to reflect on the design during the project and evaluate the adherence to principles (Sein et al., 2011). Section 4 reflects on the learning outcomes of this study and discusses the implications. The last stage of ADR is the formalization of learning,

at this point the goal is to move from the specific to the generic (Sein et al., 2011) and provide a set of design principles for a class of field problems. This study is the first one that uses ADR within the field of accident causation analysis. Therefore, further studies are needed in order to develop more concrete results that can provide mature design principles in this area.

3.5- DATA COLLECTION.

The different phases of our study included multiple sources of data such as meetings, related literature papers, qualitative interviews and live demos. The data collection mostly covered topics such as the theory behind the DREAM method, the functional requirements of the tool and the needs of SAFER. Table 3 below summarizes our data collection procedure during every phase of ADR (Henfridsson and Olsson, 2007):

Table 3. Summary of the data collection procedure

Stage 1: Problem formulation
The problem formulation started with an initial meeting where SAFER explained the practical problem. We collected documents and research papers about DREAM. Data sources:
<ul style="list-style-type: none"> ● Meetings ● SAFER documents (research papers relevant to the problem) ● Literature related to accident causation analysis

Stage 2: Building, Intervention, and Evaluation
<p>This stage was done in the form of sprints. We held meetings continuously with a senior practitioner to refine the design. The prototype was also demonstrated numerous times to gather feedback. Interviews were also conducted to gather requirement related data.</p> <p>Data sources:</p> <ul style="list-style-type: none"> ● Design meetings ● Demos ● Interviews
Stage 3: Reflection and Learning
<p>The goal of stage three is to analyze intervention results and evaluate adherence to principles (Sein et al, 2011). The prototype was tested with the head practitioner at SAFER to make sure it follows DREAM's theoretical principles. This was mostly done through live demos where the practitioners used the prototype. A lunch seminar at SAFER was also arranged where we presented the tool to a wider set of users.</p> <p>Data sources:</p> <ul style="list-style-type: none"> ● Demos ● Lunch seminar
Stage 4: Formalization of Learning
<p>This phase is characterized by abstracting the learning outcomes into a class of problems. In our case. Hence, more research is needed in order to establish abstract design principles.</p>

4- DATA AND DISCUSSION

In this section, we discuss the implications of our study at SAFER. Also, the collected data is used to illustrate what is discussed here. The data is presented in the form of episodes where actual data or events are described. In section 4.1, the practical

implications of our study are discussed, mainly the tool itself as well as the practical organizational interventions. Next, in section 4.2, the implications of using user involvement and DREAM are presented. Design principles related to the last stage of ADR are also presented. In section 4.3, the use of ADR is discussed and reflected upon.

4.1- PRACTICAL IMPLICATION

4.1.1- THE ARTIFACT

When creating the artifact, much emphasis was put on the graphical aspect of representing the data in such a way that it would give as much as an overall layout as possible. The graphical representation (also known as information visualization) had to be fitted according to what kind of data and what information the user is looking to extract from that data in terms of by searching or by coincidence i.e. the representation has to be done in such a way that if a human knows what he or she is looking for they should be able to spot it. The user should also be able by just browsing the representation, extract information that he or she may not have been looking for but existing none the less. In this case, the data in correlation with the analysis method (DREAM) is depicted as a series of chain-of-events method (Sandin, 2008).

Sandin (2008) presented multiple types of information visualization for chain-of-events methods with figure 4 displaying some of them. During the early stages of research it was evident that the choice of how to visualize information varied between users of the DREAM method. Even though this was the case, seeing as how the DREAM method had such a

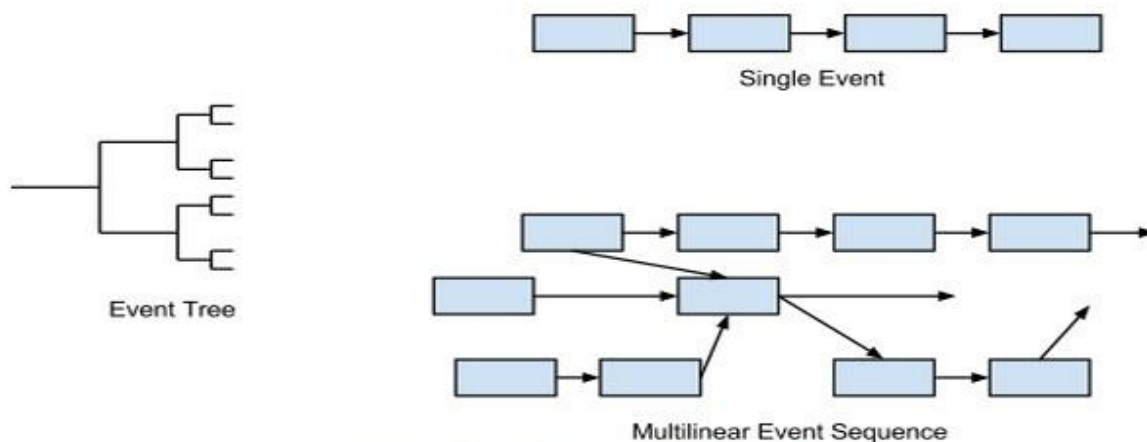


Figure 4. Adopted from Sandin, 2008

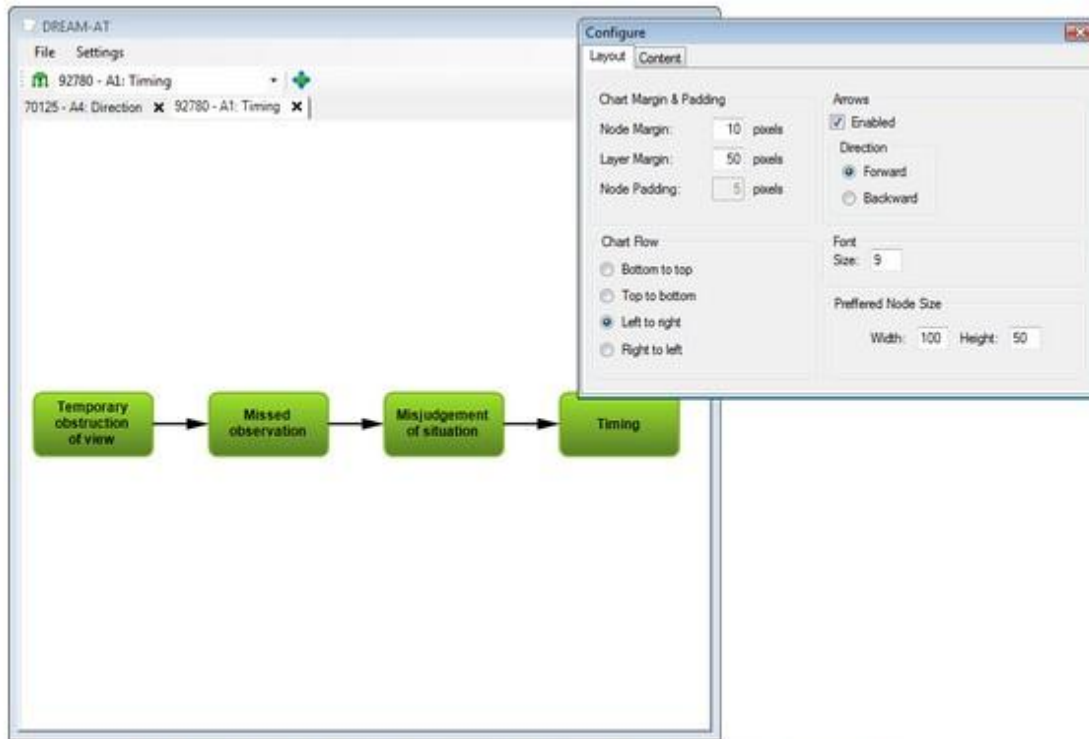


Figure 5. The DREAM-Aggregation Tool prototype running on Windows

consistent and general way of grouping causes and effects leading to consequences, in accident causation terms all data was always most readable drawn as either single event or multi linear event sequence.

With this in mind, the tool was designed following the structure of the aggregated charts which in turn follow the multi linear event sequence method of presenting the diagram. This because most practitioners at SAFER used this principle of explaining the flow of the sequence (from left to right) with the most occurring variation being reading the sequence backwards from top to bottom. The tool was therefore designed to follow the multi linear event sequence giving the possibility to change flow of the sequence As seen in figure 5 the user is given multiple choices as to how not only the diagram is rendered but also the layout of it such as: margins, size of text, direction of arrows, possibility to view without arrows, etc. This gives each user the possibility to view the charts in a manner that makes it as readable as possible for each individual.

Prior to data aggregation, each chart is presented in a drop-down list. Once clicked, a chart is then displayed in a tab page shown in the main display

area. By displaying them in tab pages it is easy to cycle between chosen charts and quickly get an overhead of the difference rather than to display each chart one at a time. This design decision also relates to the aggregated charts as once these are created, are added in the drop-down list.

The actual aggregation is done by selecting charts to be part of the aggregation and giving the aggregated chart a name. This chart is then added in the drop-down list and displayed on the screen (see figure 6). It is then possible to view the number of occurrences for connections in the chart as well as filtering away when there are less than a certain number of connections between nodes in the chart.

With the help of the tool developed, the user can simply import DREAM data and aggregate thousands of charts. The aggregation is done automatically. This saves a lot of time previously dedicated to this task (weeks). Before, several people that know DREAM had to spend a lot of time on aggregating charts. The chances of errors were high and many tools were involved. DREAM-AT should make this easier. We tested the method of aggregating the charts manually and in our case aggregating 3 charts took as long as half a day, while the tool does the same aggregation in seconds.

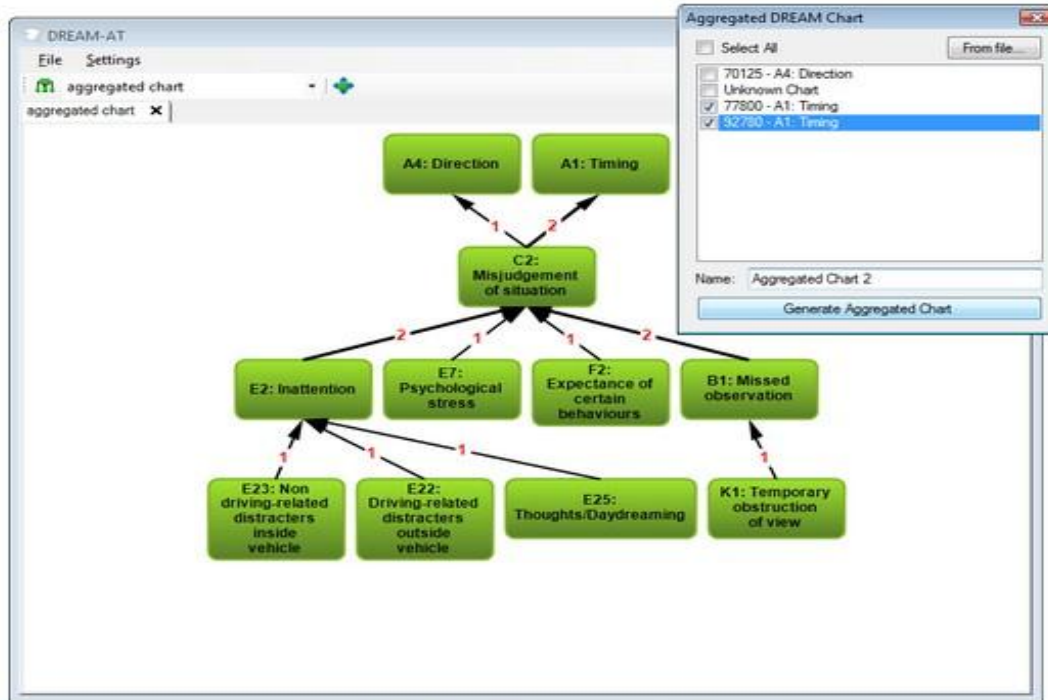


Figure 6. A generated aggregated chart

4.1.2 - ADR EFFECTS ON THE ORGANIZATION

First, it is important to reflect on our research method and how it helped us solve a practical problem in regards to the goals of ADR. Sein et al. (2011) argue that action design research aims to address a problematic situation while building an innovative artifact in an organization and learning from the intervention. In our case, we built a prototype that automates charts aggregation during the organizational intervention. Furthermore, both the developers (us) and SAFER benefited from this collaboration. In fact, during our intervention, we brought change to some SAFER practices in a way that improved their work as illustrated in the examples below:

Episode 1:

The first organizational intervention was caused by the comma separated files (CSV) exported from a database. It was suggested by the researchers to have a function that simply exports all DREAM related tables instead of exporting each file separately. SAFER participants took it into consideration and delivered the idea to the developers of the database. At the end, the new feature was added and the current database system (DaCoTa) allows exporting all the DREAM related tables instead of selecting them manually.

Episode 2:

The CSV files use the comma (,) as a separator for the fields. Some practitioners use commas in the text when they enter data into the database. This resulted in corrupted CSV files when exporting because a CSV also interprets the commas that are part of a field (text) as a separator. We suggested that the separator of the CSV files used by the system to be changed to tabulation instead of a comma. SAFER participants agreed on this and were willing to update this. Doing so made the prototype error free when importing CSV files. Practitioners that enter data into the database could also use commas freely without corrupting the exporting process.

Episode 3:

Another example of organizational intervention relates to one of the features implemented towards the end of our study. SAFER participants required one function to filter all the DREAM charts after they had been loaded into the prototype. Initially, the goal was to use an extra CSV file that contains additional data about accidents. The extra file was to be imported from a new accident case management system that SAFER is developing. However, since the new system is still being developed, we needed to agree on how the prototype would work with it. After meetings and demos, it was decided that the prototype would rely on an extra filtering file (CSV) with two fields used to identify charts previously loaded and filtered.

The above examples show how researchers work together with an organization to improve the work practices and learn from each other which is one of the goals of ADR. It is worth mentioning that the second episode also benefits from ADR via frequent communication and knowledge sharing principle.

4.2- THEORETICAL IMPLICATION

4.2.1- EFFECTS OF USER INVOLVEMENT

In this section, the effects of user involvement are first presented in the form of examples (episodes) to show the results of user involvement in the organization. Next we discuss the theoretical implications of user involvement in this study.

As mentioned in section 2.2, our collaborating organization lacked an understanding of their needs in regards to the prototype in the beginning. We therefore opted for a development approach that favored user needs. This was a good decision because we could on multiple occasions extract more detailed user requirements. Indeed, some features were not clear enough until we showed one of the users (practitioners at SAFER) how it was developed and how we perceived the feature from our perspective.

Episode 1:

During a meeting, the goal was to determine how the tool would communicate with other systems already in place at SAFER. The practitioners proposed an initial design. When the implementation of this design started, we continuously involved the

practitioners to the point where everyone realized the initial design suggested was not feasible due to time constraints and the risk of duplicate requirements. The design was eventually revised into a feasible requirement. Figure 7 shows the initial design on a whiteboard.

Episode 2:

Another example of the tight collaboration based on user involvement between researchers and practitioners is the negotiation of function implementation details. SAFER practitioners preferred to have the possibility to filter out DREAM charts already loaded into the prototype with an extra imported file that would enable interfacing with other systems. This feature was unclear in the beginning and would take a long time to implement since we had to study how other systems worked. But after the discussion of the time left for researchers with one of the user and how the prototype could be changed, we eventually managed to adjust the feature in a way that not only made it respond to user needs but also feasible within the timeframe left.

The outcome of such close collaboration led us to some realization on our end as to how decisions in terms of design should be carried out between the researchers and the practitioners. Not only did the influence from the practitioners alter the way we conducted our work but also as to how we perceived the development process.

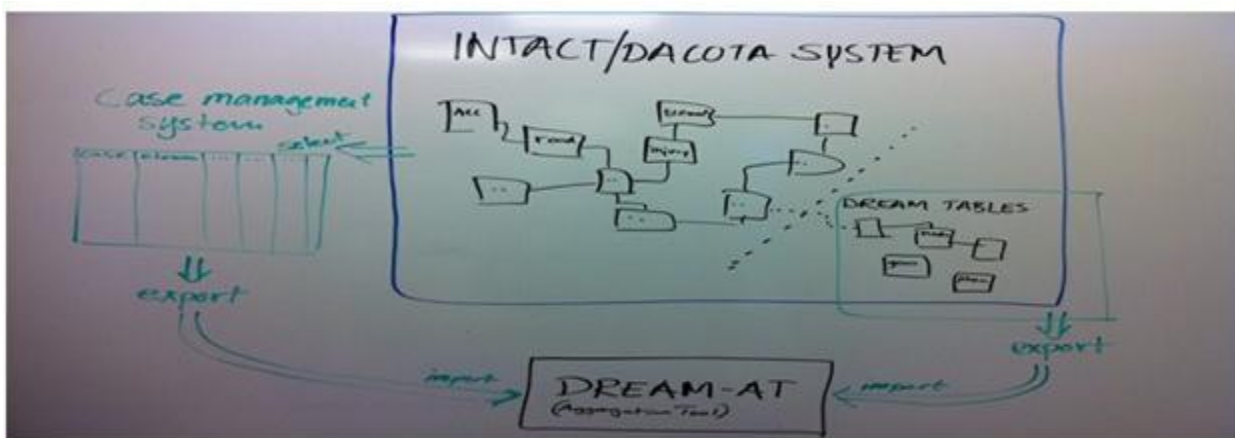


Figure 7. Initial design of one of features as designed in collaboration with the practitioners

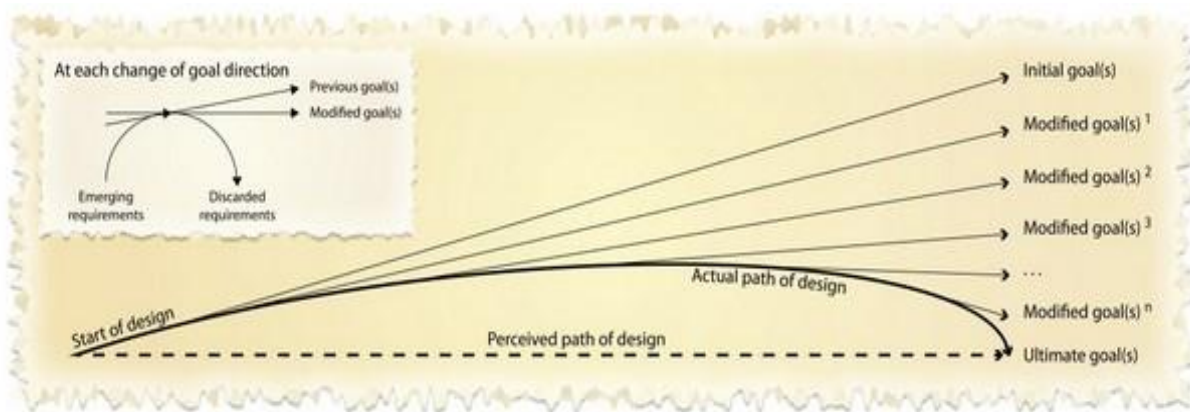


Figure 8. Implications of emergent design. Adapted from Olsson (2011) based on Gasson (1997)

The core concern of the artifact development was to relate to and understand the context and area of the problem; i.e. identifying design aspects for the prototype based upon the understanding of the method used by the practitioners (DREAM). Doing so therefore becomes a process during ADR seeing how the understanding of the goals existing for the artifact are gradually evolved as the understanding of the design improves as seen in figure 8 (Gasson, 1997). In our case the process is best defined as the means of creating a mutual pool of knowledge between the practitioners and the researchers in order to detect implications of the emergent design from both sides.

An example of such understanding can be found in the first episode where the presently used method was explained. Once they had done that we then explain to them how we had perceived their explanation and how we expected to work in accordance to this. The first stages of the implementation of design was therefore to bring knowledge between the practitioners and the researchers as close together as possible in order to detect the most evident design decisions, more than the important design decisions because up-front, it is very hard to evaluate what design decisions are important. The degree of importance of certain design decisions is better to come later.

The way of creating a shared pool of knowledge can be explained via figure 9 where the left circle represents knowledge possessed by the researchers in regards to software engineering, computer science, etc. whereas the right circle represents the knowledge possessed by the practitioners on the DREAM method and the process of analyzing around that method. The closer these two circles can be drawn

the more eminent the emerged design can be, i.e. that the more the two parties share their understanding and interpretation of the knowledge the more can a shared set of knowledge emerge and make evident design decision.

During development, researchers mostly base new design decisions on problems or ideas based on their expertise in the area of software engineering such as user interaction, graphical interface, etc. In retrospect, this could be because of the importance of actually getting a design idea physically done in order to properly evaluate it. At first glance, this process seems to lay importance on the side of the design researchers seeing as the emergent design can be produced without constant input and discussion between the two. This in turn could explain to some extent the poor user involvement during development phases (Majid et al., 2010). We found during development that while it is important for us to drive the development forward, it is important to share as much drive as possible with the practitioners and try to find the tricks to get the users involved as much as possible so that they feeling that they own the project as much as the researchers do, rather than just checking off now and again to see how it's going.

However, the design decisions that had the most impact, were not as radical in terms of decisions made that had a great sense of usefulness for the tool in comparison to the design decisions made once a practitioner actually saw what had been produced and therefore understood the design possibilities which had been previously unknown to them. Triggered by the realization of the importance that development of

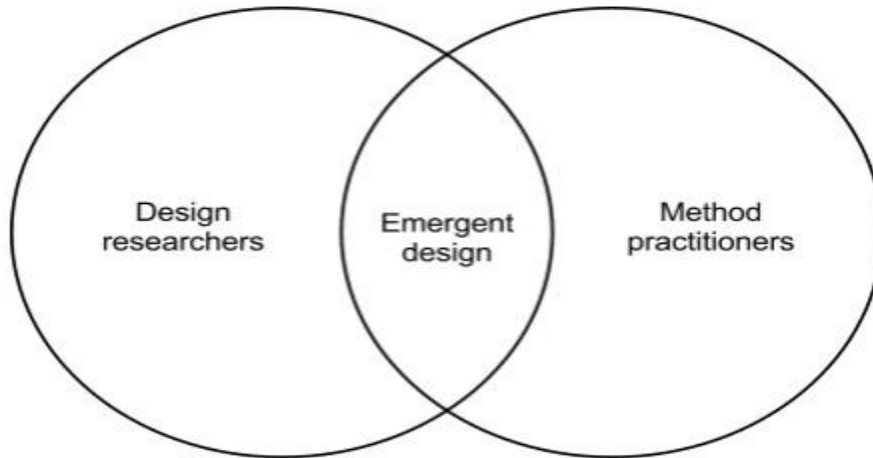


Figure 9. visualization of the interpretation of knowledge for the emergent design

a shared understanding, we started demonstrating all features immediately after reaching a demonstrable state. One feature was even demonstrated 15 minutes after it had been implemented in a demonstrable state. Because that during the demonstrations we got more than just feedback from the practitioners, we actually had a conversation about the future of the tool which drove the emerging discussions to guide the direction of the design. Subsequently, this exemplifies how the notion of learning-by-doing and reflection-in-action helps capture the understanding of actual needs during design (Gasson 1997; Olsson 2011). In addition, when we started doing this the practitioners also became more engaged in the non-functional requirements such as to fit the way they as individuals would use the tool in accordance to DREAM to analyze the data. It also became apparent that the practitioner that had been part of the design discussions from the start also had the most relevant design suggestions in terms of what was feasible to develop in the time period, difficulty of implementation, etc. This meant that the practitioner in question knew enough about the researchers' context so that the practitioner could propose design decisions that would not only benefit the artifact but also the means to develop it. By experiencing this we would like to emphasize importance of getting people involved early since it gets harder for people that get involved later to get a sense of understanding of what the emergent design decisions are and can be.

Prior to the development of the artifact, emergent design was not a recognized notion in the development process for neither the researchers nor the practitioners. However, during the usage of the BIE the importance of this notion became apparent. For example, when trying to create a dialogue in the ADR of the goals for the practitioners and the researchers. As pointed out by Olsson (2011), gradual refinement of goals is no new phenomena within systems development but during our research we detected just how much of an impact that emergent design had in relation to how ADR emphasizes close cooperation between us as researchers and the practitioners (Sein et al., 2011). This also supported the implications that design goals and problems are emerging aspects, recognizing that designers guide toward what makes sense in the specific context of activity (Olsson, 2011). Thus the importance of emergent design shows that researchers using ADR can't have too big expectations at the start of the research as to what is going to be implemented in a research perspective. In that case it will only become a consultancy job and in which case ADR will not become a natural method for testing hypothesis. The way the design, interaction and how you as researchers work create the possibility to create knowledge.

4.3- METHODOLOGICAL IMPLICATION

This study makes a methodological contribution by being an early adopter of action design research. Therefore, it is important that this paper reflects on

the experience of using ADR. First, the method allowed us to intervene in an organization in order to solve a practical problem. The result was a prototype that was continuously refined by the researchers initial design decisions and the organizational context from which it emerged. Second, by adopting ADR the practices of SAFER improved when it comes to accident causation analysis and how it is conducted.

Since action design research is a relatively new method (Sein et al, 2011), adopting it instead of action research or design research meant that no prior papers that used ADR were available as a reference. While it allowed us to solve a real world practical problem and improve the work practices, more work is needed before mature design principles can be developed for the potential class of problems within the field of accident causation analysis. Sein et al (2011) mention that once a beta version of the artifact is available, it should be refined further in a wider organizational context (stage three), this study did get to phase three but didn't get through all of it. As a result, more studies and research using action design research ought to be conducted if ADR is to be established as a credible research method.

5- CONCLUSION

This study set out to assess the feasibility of an automated computer aided analysis tool for traffic accident causation data through a prototype implementation. Through action design research, we have not only developed a theory-ingrained IT artifact to assist in traffic accident analysis but also presented a set of design principles that are likely to be important to researchers within data mining and more specifically accident causation data analysis. In terms of future research, we particularly suggest more studies that adopt action design research as a method to fully explore its positive and negative aspects, specifically, a study that succeeds in developing mature design principles. It would also be interesting to see a study that assesses the work practices of SAFER after the introduction of the tool. To this end, this research provides a first glimpse of what an ADR study could be, we encourage other researchers to conduct more ADR related studies in full organizational settings.

REFERENCES

- Björklund, G., Björkman, K., Flink, E., Karlsson, J., Viström, M., Andersson, A. L., Johansson, E., Gustafsson, H., Patten, C. 2007. INTACT Report: Pre-crash analysis method. SAFER internal report.
- Das, V.V. 2007. Involvement of users in Software Requirement Engineering. MES College of Engineering, Kuttippuram, Kerala, India. Published in IEEE.
- Gasson, S. 1997. Cooperative information system design how multi-domain information system design takes place in UK organisations. PhD thesis, University of Warwick, Warwick Business School.
- Gregor, S. 2006. The Nature of Theory in Information Systems. *MIS Quarterly*, 30(3), pp. 611-642.
- Heiskari, J., Lehtola, L. 2009. Investigating the State of User Involvement in Practice. Software Business and Engineering Institute, Helsinki University of Technology. Published in IEEE
- Henfridsson, O., Olsson, C.M., 2007. Context-Aware Application Design at SAAB Automobile: An Interpretational Perspective. *Journal of Information Technology Theory and Application*, 9(1), pp.25-42.
- Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004. Design Science in Information Systems Research, *MIS Quarterly*, 28(1), pp. 75-105.
- Hollnagel, E. (1998). Cognitive Reliability and Error Analysis Method: CREAM. Oxford, UK: Elsevier Science Ltd.
- Iivari, J. 2007. A Paradigmatic Analysis of Information Systems as a Design Science. *Scandinavian Journal of Information Systems*, 19(2), pp. 39-63.
- Kotter, J.P. 1995. Leading Change: Why Transformation Efforts Fail. *Harvard Business Review*, March/April.
- Ljung, M. 2002. DREAM: Driving Reliability and Error Analysis Method. (Master's thesis). Retrieved April 12, 2012, from University of Linköping's web site: <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-2033>
- Ljung, M., Fagerlind, H., Lövsund, P., Sandin, J. 2007. Accident investigations for active safety at chalmers: New demands require new methodologies. *Vehicle System Dynamics*, 45, pp.881-894.

Majid, R.A., Noor, N.L.M., Adnan, W.A.W., Mansor, S. 2010. A Survey on User Involvement in Software Development Life Cycle from Practitioner's Perspectives. Faculty of Computing & Mathematical Science, Universiti Teknologi MARA, Malaysia. Published in IEEE

Nirkhi, S. 2010. Potential use of Artificial Neural Network in Data Mining. Computer and Automation Engineering (ICCAE). IEEE Conference Publications, pp.339-343.

Olsson, C.M., 2011. Developing a mediation framework for context-aware applications: An exploratory action research approach. Ph. D. University of Limerick.

Otte, D., Pund, B., Jänsch, M. 2009. A New Approach of Accident Causation Analysis By Seven Steps (ACASS). International Technical Conference On The Enhanced Safety of Vehicles.

SAFER, 2011. DaCoTA Accident Causation Coding Systems Selection Process and Results. Internal WP2 report.

Sandin, J. 2008. Aggregating Case Studies of Vehicle crashes by means of causation charts. Chalmers university of technology.

Sein, M. K., Henfridsson, O., Puroo, S., Rossi, M., & Lindgren J. 2011. Action Design Research. *MIS Quarterly*, 35(1), pp.37-56.

Susman, G., and Evered, R. 1978. An Assessment of the Scientific Merits of Action Research, *Administrative Science Quarterly* (23), pp. 582-603.

Warner, H.W., Ljung, A.M., Sandin, E., Johansson, E., Björklund, G. 2008. Manual for DREAM 3.0, Driving Reliability and Error Analysis Method. Deliverable D5.6 of the EU FP6 project SafetyNet, TREN-04-FP6TR-SI2.395465/506723.

Wang, X., Wang, J.T.L., Shasha, D., Shapiro, B.A., Rigoutsos, I., Zhang, K. 2002. Finding Patterns in Three-Dimensional Graphs: Algorithms and Applications to Scientific Data Mining. *IEEE Transactions on Knowledge and Data Engineering*, 14(4), pp.739-749.

Xi, J., Wang, S., Chen, X., Zhu, R. 2010. Accident Cause Analysis Method based on Traffic Accident Information System. Published in IEEE.

Zhang, D., Zhou, L. 2004. Discovering Golden Nuggets: Data Mining in Financial Application. *IEEE Transactions on Systems and Cybernetics*, 34(4), pp.513-522.