Revisiting the Concept of Risk Analysis within the Context of Geospatial Database Design: A Collaborative Framework

J. Grira, Y. Bédard, and S. Roche

Abstract—The aim of this research is to design a collaborative framework that integrates risk analysis activities into the geospatial database design (GDD) process. Risk analysis is rarely undertaken iteratively as part of the present GDD methods in conformance to requirement engineering (RE) guidelines and risk standards. Accordingly, when risk analysis is performed during the GDD, some foreseeable risks may be overlooked and not reach the output specifications especially when user intentions are not systematically collected. This may lead to ill-defined requirements and ultimately in higher risks of geospatial data misuse. The adopted approach consists of 1) reviewing risk analysis process within the scope of RE and GDD, 2) analyzing the challenges of risk analysis within the context of GDD, and 3) presenting the components of a risk-based collaborative framework that improves the collection of the intended/forbidden usages of the data and helps geo-IT experts to discover implicit requirements and risks.

Keywords—Collaborative risk analysis, intention of use, Geospatial database design, Geospatial data misuse.

I. INTRODUCTION

PRESENTLY, risk analysis efforts are not systematically integrated to geospatial database design methods in a way to be conformant to requirement engineering guidelines and ISO risk standards [2]. Risk analysis requires identifying potential risks related to the uses of a given dataset. Defining the intended and forbidden usages of the geospatial data is part of the risk identification step, prior to risk analysis. However, usage intentions are rarely considered as an input to the risk analysis process. This may lead to faulty risk analysis and misleads decision making.

Although, several ISO Standards [10],[11] and research [2],[14],[13],[12] pointed the need to consider the intentions of use during the design process, usage intentions are still ill-defined at the design stage and are not methodically collected and recorded.

The aim of this paper is to present a collaborative risk-based framework that aims at improving collection and description of intentions of use of geospatial data as well as analyzing the risks related to misusing that data.

In the next section, we first review the related work in the field of risk analysis and in requirements engineering (RE). In Section III, we discuss the challenges of risk analysis and geospatial database design. In Section IV, we present a framework that engages stakeholders in a collaborative process based on the Delphi method in a first step and on a Volunteered process in a second step. We expose in the same section the architecture of the framework and show how intentions of use and the underlying risks can be captured and recorded within a geospatial data repository. Next, we discuss the advantages and disadvantages of the proposed framework in Section V. Finally, we conclude and present further work in Section VI.

II. BACKGROUND

Related work lies on two major areas: risk analysis and requirement engineering.

A. Risk Analysis

Risk analysis is part of risk management efforts. Risk management refers generally to the following activities: risk identification, risk analysis, risk evaluation, and risk treatment [15]. Depending on the perspective, risk analysis is defined differently:

From a standard-based perspective, risk management is considered as being a set of “coordinated activities to direct and control an organization with regard to risk”[11]. Risk analysis is a “process to comprehend the nature of risk and to determine the level of risk”[11]. Risk identification is a prerequisite for risk analysis: it can involve historical data, informed and expert opinions, and stakeholder's needs.

Once identified, the risk is then analyzed according to its likelihood and its severity. Risk analysis involves consideration of the causes and sources of risk, their consequences, and the likelihood that those consequences can occur [16].

In standards dealing with public protection and safety, risk analysis is defined as “the systematic use of available information to identify hazards and to estimate the risk” [10]. This definition focuses on the preventive actions and protective measures that may be taken in order to reduce the risk arising from the use of a product or a service [8].

Different techniques can be used in risk analysis; the most relevant for our purpose are experts' opinions and validation of a common understanding of the stakeholder’s needs.

From a project management perspective, the addressed risks are usually those defined as included in the scope of the project. These risks are known as being “project risks” and include organisational, technical and external risks [18].

J. Grira, Y. Bédard and S. Roche are with Geomatics Science Department, Laval University, Quebec (QC), G1V0A6, Canada. (e-mail: joel.grira.1@ulaval.ca, yvan.bedard@scg.ulaval.ca and stephane.roche@scg.ulaval.ca)
According to best practices in project management [17], project risks are typically recorded in a structure known as "risk register". This register is usually intended to keep track of the constraints that may threaten the project objectives achievement.

B. Requirements Engineering

In the area of requirements engineering, "The success of a software system depends on how well it fits the needs of its users and its environment" (e.g. [3]). Requirements comprise these needs, and RE is the process by which the requirements are determined. Thus, RE is about defining the problem that the software is to solve (i.e. defining "what" the software should do), whereas other software engineering activities are about defining a proposed software solution. Especially in goal-oriented RE (GORE) methodologies, the focus is on modeling the “why” aspect of requirements in addition to the “what” aspect. Some goal-oriented RE methodologies, such as KAOS [20] and i* [19], address risky issues during requirements analysis by modeling the goals and the design decisions. In such methodologies, the concept of obstacles, considered as a risk that undermines strategic interests of stakeholders, and the concept of anti-goal, defined as an obstruction to the fulfillment of stakeholders’ goals, is introduced. Obstacles and anti-goals may be derived from a goal structure and the expert’s knowledge about the system. Literature about GORE confirms that RE artifacts (i.e. the output of the RE process) have to be understandable and usable by domain experts and other stakeholders, who may not be knowledgeable about computing. RE research has defined five types of requirements activities: elicitation, modeling, analysis, validation/verification, and management. For the purposes of the present work, a high-level overview of these activities is presented:

- **Requirements elicitation** comprises activities that enable the understanding of the purposes for building a system. Most of the research in requirements elicitation focuses on the techniques used to improve the stakeholders’ identification [21] and helping them to express their needs [22]. Some other research focus of the RE elicitation techniques used to capture requirements [23].

- **Requirements modeling** consists on expressing requirements in terms of precise and unambiguous models. This process helps to evoke details that were missed in the initial elicitation. The resulting models could communicate requirements to the design team. Modeling notations are the main research focus and differ by the specific details they elicit and record (e.g. data, functions, and properties) [24].

- **Requirements analysis** comprises techniques for evaluating the quality of recorded requirements. Some studies look for errors in requirements [25] or focus on anomalies in requirements [26]. These studies reveal misunderstandings or ambiguity about the requirements that usually call for further elicitation. Risk analysis [27] is part of the requirements analysis techniques that help IT designers to better understand the requirements, their interrelationships, and their consequences.

- **Requirements validation** ensures that stakeholders’ needs are accurately expressed. Thus, validation is typically a subjective evaluation of the requirements with respect to informally described or undocumented requirements. Accordingly, the validation task requires stakeholders to be involved in reviewing the requirements artefacts [28]. Research in this area focuses on improving the information provided to the stakeholder [29].

As seen above, in the non-spatial domain, the RE research community has made significant progress along many fronts. However, in the context of geospatial database design, although RE considers risk analysis as part of requirements analysis activities [26], risk analysis is not always undertaken iteratively in conformance to RE guidelines and ISO risk standards.

C. Paper Contribution

Though, the literature offers a variety of contributions in the areas of risk analysis and RE, a lot must still be done to fully integrate approaches of these areas in the geospatial database design process. Our aim in this paper is to propose a framework that integrates iteratively risk analysis into RE activities, particularly in a context of geospatial data reuse during designing the database. We strongly believe that a systematic risk analysis performed iteratively within the requirement analysis phase of geospatial database design can help discovering new requirements and improve the knowledge about the risks related to the misuse of the underlying geospatial data. Moreover, most of the work presented in literature consists mainly in approaches and frameworks with limited consideration to end-users implicit requirements during risk analysis activities. In this paper, we extend the scope of project risk management by including risk analysis of geospatial data as an iterative and collaborative activity in project risk management in conformance to RE guidelines and ISO standards; we propose an iterative mechanism to improve the collection of the intended usages of geospatial data by helping geo-IT experts to discover implicit requirements of end-users during the requirement analysis process.

III. CHALLENGES OF RISK ANALYSIS IN THE CONTEXT OF GEOSPATIAL DATABASE DESIGN

A. Deficiency of Unsystematic Risk Analysis

The output of a risk analysis process is tightly dependant on the initial context within which the analysis is performed. The relation between the risk analysis output and the context is rather fluid as the context evolves over time. However, the changes in the context of the risk analysis and its underlying consequences are often overlooked by project management: risk analysis is usually performed once, and even skipped sometimes or done with little rigor, mainly at the beginning of the project.
Besides, risk register, which is a commonly known tool typically dedicated to record project's risks, is rarely maintained after its constitution and its initial feeding: although a risk recorded in the register might have been addressed later in the project and become a non-issue, there is no guarantee to update the register because of the non-blocking nature of that activity; similarly, new risks might rise with the emergence of news needs or even with the introduction of ad-hoc changes in the final system. Consequently, the risks register may not fully reflect the up-to-date context of risks.

Additionally, the risk register structure is semantically poor: it does not allow the traceability of the intended usages, from their formulation until their translation into specifications, while keeping track of the related risks. It is usually a tabulated structure used in practice to document risks textually.

As stated in ISO-31000 standard, enhanced risk management includes continual communications with external and internal stakeholders, as part of good governance. During the risk analysis process, collaboration and communication lead to an improvement in the knowledge about the problem-space. It may also lead to the formulation of insights often critical to the success of the project. Without a proper mechanism to collect, organize and keep trace of this influx of information (i.e. end-users objectives and the underlying risks, knowledge of application domain experts) and its elaboration into insights and design entities (i.e. translation of end-users assertions into specifications by application domain experts), geospatial database design could miss some risky issues that need to be addressed for the delivery of a satisfactory solution.

Finally, the risks concerned by the present work are risks of geospatial data misuse. For example, there may be a risk related to spatial inaccuracy in the case of distance calculation between the street and the border of buildings in order to verify the compliance of those buildings to municipal regulation in terms of zoning. Another example is the calculation of the optimal path to be followed by an emergency vehicle based on sporadically updated data. In the first example, the risk is assessed in terms of financial loss but for the second example, the consequences may be life losses. A number of geospatial data misuse cases have been described in literature [33] in relation to their underlying risks, and the potential harm and severity of that risks [10].

B. Shortcomings of Traditional Geospatial Database Design

Geospatial data projects need to be able to rely on a sufficient level of end-user collaboration during the design of the GI database in order to identify early the risks that may arise while using the data. Such attitude is part of a good professional ethics and not doing so may involve professional liability [31],[32]. In fact, intended use is considered as “available information” that might be leveraged to reduce the risks of reasonably foreseeable misuse [10]. However, the focus of the design process is not primarily on risks, but on data.

Besides, the traditional geospatial data design usually involves experts having variable skills in some specific areas. Geospatial database designers may be experts in IT or GeoIT, as they can be experts in another specific application domain. Difference in expertise determines the perspective of the designer. Consequently, there may sometimes be evidence about end-user needs (e.g. application domain experts could assume some usage patterns of the data based on their own experience) or about some technical choices (e.g. geo-IT experts could assume the adequacy of the designed model to a set of analytical operations they intend to use).

Finally, organizational constraints may impact the design process. For example, budget limitation may lead the design team to stick to the “in scope” issues rather than exploring the existence of risky issues or the existence of non-functional requirements having potential risky consequences.

IV. COLLABORATIVE RISK ANALYSIS IN GEOSPATIAL DATABASE DESIGN

Geospatial database design is a process where at least some ad-hoc activities of risk analysis have to take place: risk analysis is important in the software design phase and constitutes a prerequisite for evaluating criticality of the system [4] and taking the necessary countermeasures. However, different expertises are required to perform the risk analysis within the design phase [1],[2]: application domain expertise (e.g. ecology, epidemiology, transportation, and security), information technologies expertise (e.g. system engineers and database designers) and Geospatial Information Technology - GeoIT expertise (e.g. geomatics engineers, GIS developers, and geographers). Considering the differences in perspectives, backgrounds and objectives between experts, there may exist a divergence in the way they analyze risks. Accordingly, there is a need to bridge the gap between those that are experts about the domain and its requirements (i.e. application domain experts), those that are experts in the design and the construction of the artifacts that together satisfy the domain requirements (i.e. GeoIT experts), and those that are experts in software and database design (i.e. IT experts) in order to get a common understanding of the implied risks related to the geospatial data to be used.

In the following sections, we first depict the collaborative Delphi-based mechanism intended to help achieving the common understanding of the risks related to the manipulated geospatial data. We then propose to record the collaboratively identified risks within a formal structure, i.e. a risk taxonomy, based on a risk classification for geospatial risks. We finally present the architecture of the implemented prototype and detail a usage scenario of that implementation.

A. Collaboration Using the Delphi Method

Since its origins, the Delphi method, originally developed at the RAND Corporation as a tool to assist forecasting, has broadly been used in the domain of decision making as a collaborative technique for generating scenarios about critical
events that may happen in the future. Considering the uncertainty about the occurrence or non-occurrence of the events, building a Delphi scenario may become a challenging issue for the involved experts. It is even more challenging when those experts have different expertises as underlined above.

The Delphi method output makes explicit a set of requirements and produces a collective estimation of cross impacts of risky issues (see Fig. 1).

In the context of the present work, the Delphi method is not used as a standalone approach. Many researches contend that using the Delphi method as only part of a wider process may well prove a means to enhance its utility [3]. Following these researches, we used the Delphi method to involve experts in providing their judgment about the risks but we also involved end-users in a second collaborative step using a web 2.0 collaborative platform in order to collect non-functional requirements (i.e. goals, objectives and constraints) and assess their cross impacts on risk analysis (Fig. 2).

The use of Delphi method with other techniques (e.g. collaborative workshops, Nominal Group Technique, focus groups and face-to-face meetings) may produce more satisfactory results [5],[6] and make them more coherent [7]. The Section IV.B about a proposed tool to help detecting risks and the Section IV.C about the Prototype Architecture and Implementation describe in more details how the Delphi method is put in practice in the context of the present work.

B. Geospatial Risk Repository: a Semantically Enriched Risk Register

Risk analysis activities usually include conversations, brainstorming and face-to-face meetings. With these techniques, insights into the problem space, clarification of assumptions, and deeper understandings are discussed.
Because of their ad-hoc nature, these techniques do not guarantee a systematic and methodological risk analysis. It is the case for the great majority of projects: risk analysis is usually considered as a disorganized effort that occurs sporadically or, sometimes does not occur at all. Tools should provide a mechanism to help detecting unexpected risks systematically and transcribe them to a more structured register, e.g. a geospatial risk repository (GRR), which is semantically richer than the traditional risk register used in GORE. The potential of such a repository is broader than a simple record tool: the risk information may be exposed to GI systems and CASE (Computer Aided Software Engineering) tools using service-oriented interfaces (Fig. 3).

The proposed tool helps in detecting systematically the risks but also store them with the aim to make them available for the decision making process. The collected risk information (i.e. the link between the geospatial data and its underlying risks) is available as "an exposed service to be consumed", and no more as "a data to be reused": this allows GI systems to have full control on the way they manage the received risk information without altering the original data.

Here two examples of the use of risk information:
- in GIS tools, depending on the level of danger of the risk, the ISO 3864-2 [30] standard suggests the use of graphical warnings icons, safety signs and safety labels [9]. Those graphical symbols can be displayed to a GIS end-user in order to inform him about the risks related to data he is intending to use.
- in CASE tools, geo-IT experts in database design can rely on the information of the GRR in the same way they rely on metadata (i.e. risk information is also "data about data"). The CASE tool consumes a service exposed by the GRR to provide risk information about an attribute, or a specific value of an attribute. According the returned result provided by the GRR and displayed in a pop-up window within the CASE tool, the database designer could conclude that his analysis may be faulty if he does not consider the displayed risk information in his model.

The role of experts is important because their judgments can provide valuable information, particularly in view of the limited availability of data regarding uncertainties in a given project. The use of experts’ judgment is motivated by the lack of historical data and the desire to obtain as much information as possible about risks related to the use of geospatial datasets. Combining experts’ opinions to implicit requirements expressed by end-users in terms of non-functional requirements (i.e. objectives, constraints, and goals) summarizes the accumulated information for risk analysis. The GRR is also enriched by the information about the context of usage of the geospatial data, with includes the links between non-functional requirements and risks.

C. Prototype of the Collaborative Platform and the Geospatial Risk Repository: Architecture and Implementation

The proposed approach is implemented using web-based technologies. The implementation is performed with respect to a service-oriented architecture that divides the prototype into two web modules: (1) a collaborative platform and (2) a geospatial risk repository. The collaborative platform is technically separated into two components: a back-end and a front-end. It functionally covers two types of process: a Delphi-based process in a first step where only experts are involved and a Volunteered process in a second step where both experts and end-users are involved.
The geospatial risk repository is composed only from a back-end component. The data layer of the two modules, i.e. their database, is situated at the back-end level.

As illustrated in Fig. 3, end-users and experts interact with the front-end components of the collaborative platform. Collaboration features are available throughout web interfaces. For example, Delphi iterations illustrated in Fig. 3 are facilitated using those interfaces; it allows experts to interact about the problem formulation and to design the questionnaire. Next, end-users contribute with their knowledge to enrich the Delphi method results: they may express their intentions of use which are translated into implicit and non-functional requirements.

The risk repository provides for the structured storage of risk issues. The storage of those risks is performed in respect with a pre-defined risk classification [9].

The non-functional requirements collected during the Volunteered Stage feeds the back-end of the collaborative platform with formal requirements. In risk analysis, non-functional requirements (e.g. intentions of use, goals, usage context, and constraints) include the understanding of the context of the risks [16]. Without establishing the context that defines the parameters to be taken into account when managing risks, the understanding of the risks may be undermined. Users who interact with the collaborative platform contribute to establish the context and help defining the scope of the database design.

Finally, both the back-end of the collaborative platform and the GRR store the relationships resulting from Delphi iterations, i.e. the links between the geospatial usage context, the underlying requirements, and the resulting risks, which keeps track of the traceability links and provides usable information to support decision making.

V. DISCUSSION

The usefulness of the collaborative framework developed in this work and its implementation as a web platform may be discussed by comparing it to previous approaches that share similar objectives. Accordingly, the following advantages and disadvantages can be attributed to the collaborative risk-based approach presented in this paper.

First, in a context of geospatial database design, basing risk analysis on a collaborative platform can be an effective way to bring more active users into the design process. According to RE principles, this helps discovering new requirements that may be derived from user-contributed content about risky data usages. Compared to other approaches that have limited consideration to end-users implicit requirements during risk analysis activities (e.g. [9]), this approach helps identifying foreseeable risks about the potential uses of the data. As such, if RE is considered as the process of determining requirements [2], performing RE activities during geospatial risk analysis should in turn fine-tune design requirements leading to improving the knowledge of experts (i.e. geo-IT experts and application domain experts) about the underlying risks.

Second, compared to previous risk-based approaches [9], [34], considering intentions of usages in the risk analysis process during the design stage extends the scope of risk analysis. This scope is no more limited to a set of risks identified at the beginning of the design stage [9] or on risks identified after data production [34]. Accordingly, our approach includes foreseeable risk aspects in the scope of risk analysis which prevents experts from overlooking relevant requirements from the output specifications.

The approach developed still presents a number of limitations that would deserve attention in further studies. The implementation used a basic structure for risk classification [35]. A richer structure, such as a risk ontology for example (e.g. [36],[37]), could be considered for future work in order to allow semantic reasoning on risks and knowledge extraction. Also, when using this prototype, information about risks and about the collaborative process is displayed on a dashboard: it is not integrated into a CASE tool. However, the prototype is extensible: as such, it would be useful to adapt CASE tools in a way to integrate the prototype output into their graphical interface. This would help database designers visualizing collaborative and risks indicators within the same system.

In order to confirm the usefulness of the approach, application domain experts have been asked to complete a test that asked them to go through a usage scenario. They first interact, using Delphi iterations, to formulate a set of questions about a database design issue. Next to the user-contributed step, where users express their usage intentions of the underlying data, domain application experts revisit the exposed design issue. As a first result, experts noticed the discovery of new requirements derived from the intentions of use expressed by data end-users. As a second result, all the domain application experts agree about the improvement of their knowledge about the contributed risky issues: this newly acquired knowledge helps them producing a more "safe-by-design" geospatial data model.

VI. CONCLUSION

This paper has described a Collaborative Framework built upon best practices in risk analysis, requirements engineering and ISO standards. This risk-based Framework integrates the Delphi method to collect experts’ opinions in regards to risky issues in a first step, and helps identifying non-functional requirements expressed by end-users using a Volunteered process in a second step. The Framework provides a GRR (geospatial risk repository) to record risks as well as to support traceability of intentions of use, requirements, risks and their underlying context (i.e. usage context and risk analysis context). The GRR delivers risk information that may be used, graphically for example, in CASE tools, in GI systems to automatically produce contextual warning or to generate geospatial data quality reports, etc. Traceability, i.e. keeping track of the different links between the identified elements (i.e. intentions of use, requirements, risks and their
underlying context), is a prerequisite for reusing risk information by other applications.

Our long-term research objective is to improve, through a participative process, the knowledge about risks related to inappropriate usage of geospatial data. The Collaborative Framework presented in this paper helps identifying risky issues and provides a platform that involves experts and end-users in the geospatial database design. It also helps detecting new requirements and improving existing ones. Future research would explore more in detail the appropriate representation of risks within the repository and would depict practical usage scenarios of reusing the output of the Framework.

REFERENCES


[33] Public Protection and Ethical Dissemination of Geospatial Data - Social and Legal Aspects. Examples of geospatial data misuse: http://dataquality.seg.ulaval.ca/


