# BIM AND KNOWLEDGE BASED RISK MANAGEMENT SYSTEM

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of
Doctor of Philosophy

by

**Yang Zou**

School of Engineering
University of Liverpool
September 2017

# DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award. This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD. This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own. I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed: 邹阳

Yang Zou (Candidate)

6 September 2017

# ABSTRACT


The use of Building Information Modelling (BIM) for construction project risk management has become a growing research trend. However, it was observed that BIM-based risk management has not been widely used in practice and two important gaps leading to this problem are: 1) very few theories exist that can explain how BIM can be aligned with traditional techniques and integrated into existing processes for project risk management; and 2) current BIM solutions have very limited support on risk communication and information management during the project development process.

To overcome these limitations, this research proposes a new approach that two traditional risk management techniques, Risk Breakdown Structure (RBS) and Case-based Reasoning (CBR), can be integrated into BIM-based platforms and an active linkage between the risk information and BIM can be established to support the project lifecycle. The core motivations behind the proposed solution are: 1) a tailored RBS could be used as a knowledge-based approach to classify, store and manage the information of a risk database in a proper structure and risk information in RBS could be linked to BIM for review, visualisation and communication; and 2) knowledge and experience stored in past risk reports could contribute to avoiding similar risks in new situations and the most relevant cases can be linked to BIM to support decision making during the project lifecycle. The scope of this research is limited to bridge projects; however, the basic methods and principles could be also applied to other types of projects.

This research is in three phases. In the first stage, this research analysed the conceptual separation of BIM and the linkage rules between different types of risk and BIM. Specifically, an integrated bridge information model was divided into four Level of Contents (LOCs) and six technical systems based on the analysis of the Industry Foundation Classes (IFC) specification, a critical review of previous studies and the author's project experience. Then a knowledge-based risk database was developed through an extensive collection of risk data, a process of data mining, and further assessment and translation of the data. Built on the risk database, a tailored RBS was developed to categorise and manage this risk information and a set of linkage rules

between the tailored RBS and the four LOCs and six technical systems of BIM was established. Secondly, to further implement the linkage rules, a novel method to link BIM, RBS, and Work Breakdown Structure (WBS) to be a risk management system was developed. A prototype system was created based on Navisworks and the Microsoft SQL Server to support the implementation of the proposed approach. The system allows not only the storage of risk information in a central database but also to link the related risk information in the BIM model for review, visualisation and simulation. Thirdly, to facilitate the use of previous knowledge and experience for BIM-based risk management, the research proposed an approach of combining the use of two Natural Language Processing (NLP) techniques, i.e. Vector Space Model (VSM) and semantic query expansion, and outlined a new framework for the risk case retrieval system. A prototype was developed using the Python programming language to support the implementation of the proposed method. Preliminary testing results show that the proposed system is capable of retrieving relevant cases automatically and to return, for example, the top 10 similar cases.

The main contribution of this research is the approach of integrating RBS and CBR into BIM through active linkages. The practical significance of this research is that the proposed approach enables the development of BIM-based risk management software to improve the risk identification, analysis, and information management during the project development process. This research provides evidence that traditional techniques can be aligned with BIM for risk management. One significant advantage of the proposed method is to combine the benefits of both traditional techniques and BIM for lifecycle project risk management and have the minimum disruption to the existing working processes.

# ACKNOWLEDGEMENTS

First and most importantly, I would express my deepest love and gratitude to my parents, Wenying Huan and Kaijun Zou. It is them who have brought me into life, accompanied me to grow, and taught me how to know this world. Their positive attitudes towards study, working as well as life have influenced me profoundly and they always encourage me to do what I really want to. Although they have not been involved in any work of my PhD research, this thesis cannot be completed without their continuous support and concern.

I am grateful as well to my wife, Shancha Xu, for her full support on my dream of pursing a PhD. Making the decision to live in the UK for more than 2 years was not easy because this has totally changed her career path as well as the future plan of our family.

The key people making my PhD to become a reality are my supervisors, Professor Arto Kiviniemi and Dr Steve Jones, who accepted me as one of their PhD students, guided me to obtain research scholarships and supported my study from the first idea to its completion with their best patience. They always encourage me to keep my eyes open and give confidence on study, life and the future. I am honoured to be lucky enough that I could have worked with my supervisors during the two and a half years. I would apply the life philosophy, knowledge and research methodologies learned from my supervisors into my future work.

In the PhD process, I have received very valuable comments and suggestions from a number of industry experts, Alastair Soane, Gordon Crick, Martin Simpson, David Philp, and Benedict Wallbank. The Navisworks model was provided by James Walsh, a MEng Civil Engineering undergraduate of the University of Liverpool. In addition, I have enjoyed working with a group of lovely colleagues, Elaheh Gholami, Sujesh Sujan, Gulnaz Aksenova, Mohammad Qabshoqa, Sa’id Kori, Zuhair Nasar, Pooya Sanjari, who organised activities regularly and made me feel that I am part of their community.

The PhD research has been funded by the University of Liverpool, China Scholarship Council (CSC), and University of Liverpool Graduate Association (Hong Kong).

# LIST OF PUBLICATIONS

## Peer-Refereed Journal Articles

1. **Zou, Y.**, Kiviniemi, A., and Jones, S.W. & Walsh, J., 2017. "A Framework and Prototype for Risk Information Management through Integrating Risk Breakdown Structure into 3D/4D BIM", Journal of Construction Engineering and Management, ASCE. (submitted) **(This paper is used in Chapter 6)**

2. **Zou, Y.**, Kiviniemi, A., and Jones, S.W., 2017. "Retrieving Similar Cases for Construction Project Risk Management Using Natural Language Processing Techniques", *Automation in Construction*, Elsevier, 80, pp. 66-76, http://dx.doi.org/10.1016/j.autcon.2017.04.003. **(This paper is used in Chapter 7)**

3. **Zou, Y.**, Kiviniemi, A., and Jones, S.W., 2017. "A Review of Risk Management through BIM and BIM-related Technologies", *Safety Science*, Elsevier, 97, pp. 88-98, http://dx.doi.org/10.1016/j.ssci.2015.12.027. **(This paper is used in Chapter 2)**

4. **Zou, Y.**, Kiviniemi, A., and Jones, S.W., 2016. "Developing a Tailored RBS Linking to BIM for Risk Management of Bridge Projects", *Engineering, Construction and Architectural Management (ECAM)*, Emerald, 23(6), pp. 727-750, http://dx.doi.org/10.1108/ECAM-01-2016-0009. **(This paper is used in Chapter 5)**

## Peer-Refereed International Conference Papers

1. **Zou, Y.**, Kiviniemi, A., and Jones, S.W., 2017. "A conceptual model of collaborative risk management system", in *Proceedings of 5$^{th}$ International Workshop of When Social Science Meets BIM and Lean in Aalborg, Denmark*, 25-27 January 2017.

2. **Zou, Y.**, Jones, S.W., and Kiviniemi, A., 2015. "BIM and Knowledge Based Risk Management System: A Conceptual Model", in *Proceedings of CITA BIM Gathering 2015 International Conference in Dublin, Ireland, 12-13 November 2015*, Construction IT Alliance, Dublin, pp. 63-68.

3. **Zou, Y.**, Kiviniemi, A., and Jones, S.W., 2015. "BIM-based Risk Management: Challenges and Opportunities", in *Proceedings of the 32nd CIB W78 International Conference in Eindhoven, The Netherlands, 27-29 October 2015*, pp. 847-855.

# TABLE OF CONTENT

# LIST OF TABLES

Page

# LIST OF FIGURES

Page

# LIST OF ABBREVIATIONS AND SYMBOLS

The following abbreviations and symbols have been used in this thesis:

| | |
|---|---|
| 2D | Two-dimensional |
| 3D | Three-dimensional |
| 4D | Four-dimensional |
| 4D CAD | Four-dimensional computer aided design |
| AEC | Architecture, Engineering and Construction |
| AI | Artificial Intelligence |
| API | Application Programming Interface |
| BIM | Building Information Modelling |
| BIMKSM | BIM-based Knowledge Sharing Management |
| BIM-VE | BIM based virtual environment |
| BKRMS | BIM and Knowledge Based Risk Management System |
| C# | C-sharp |
| CBR | Case-based Reasoning |
| CBRL | Case-based Reasoning Library |
| CDE | Common Data Environment |
| CDM | Construction Design and Management |
| CMMS | Computerised Maintenance Management System |
| CSHM | Construction Safety and Health Monitoring |
| DFSP | Design-for-safety-process |
| DSS | Decision Support System |
| EDM | Express Data Manager |
| EV | Earned Value |
| FM | Facility Management |
| FN | False Negative |
| FP | False Positive |
| FTA | Fault Tree Analysis |
| GIS | Geographic Information Systems |
| GPS | Global Positioning System |
| ICT | Information and Communication Technologies |
| IFC | Industry Foundation Classes |

| | |
|---|---|
| ILO | International Labour Organization |
| IR | Information retrieval |
| KM | Knowledge Management |
| LOC | Level of Content |
| NBIMS | National BIM Standard |
| NLP | Natural Language Processing |
| NLTK | Natural Language Toolkit |
| NN | Neural Networks |
| POP | Product, organisation and process |
| PPE | Personal Protective Equipment |
| RBM | Risk Breakdown Matrix |
| RBS | Risk Breakdown Structure |
| RCRS | Risk Case Retrieval System |
| RFID | Radio Frequency Identification |
| SoA | Sequence of Actions |
| SPMT | Self-propelled Modular Transporter |
| SQL | Structural Query Language |
| SWOT | Strength-weakness-opportunity-threats |
| TF-IDF | Term Frequency-Inverse Document Frequency |
| TN | True Negative |
| TP | True Positive |
| UCN | Universal Circulation Network |
| UI | User Interface |
| UWB | Ultra-Wideband |
| VP | Virtual Prototyping |
| VR | Virtual Reality |
| VSM | Vector Space Model |
| WBS | Work Breakdown Structure |

| | |
|---|---|
| $q$ | Query |
| $d_j$ | Document |
| $k_i$ | Index term |
| $f_{i,j}$ | Frequency of index term $k_i$ in a document |
| $(k_i, d_j)$ | Term-document pair |
| $N$ | Total number of documents in the document set |
| $w_{i,j}$ | TF-IDF weight |
| $n_i$ | Number of documents containing the term $k_i$ |
| $sim(d_j, q)$ | Degree of similarity between the document $d_j$ and the query $q$ |
| $\|\vec{q}\|$ | Norms of the query vectors |
| $\|\vec{d_j}\|$ | Norms of the document vectors |
| $\vec{d_j} \cdot \vec{q}$ | Inner product of the document and query vectors. |
| $q_o$ | Original query |
| $q_e$ | Expanded query |
| $\lambda$ | Coefficient for the effect of query expansion |
| Precision | Fraction of retrieved documents that is relevant and is used to measure the percentage of relevant documents in all retrieved documents |
| Recall | Fraction of relevant documents that has been retrieved and used for measuring the percentage of retrieved documents in all relevant documents |
| F | Harmonic mean of $precision$ and $recall$ |
| P@10 | Percentage of relevant risk cases in the top 10 similar documents |

# Chapter 1. INTRODUCTION

## 1.1 Problem description

### 1.1.1 Risks in the Architecture, Engineering and Construction Industry

The Architecture, Engineering and Construction (AEC) industry has witnessed a rapid development all around the world during the last few decades. Large-scale projects have become widespread and international, new project delivery methodologies are being adopted, design theory and tools are constantly improving, creative and new approaches, methods, and materials of construction are being introduced (Bryde *et al.*, 2013). AEC projects such as buildings and infrastructure systems are part of the fabric of urban spatial planning and design, and have an immediate impact on and a direct relation to the accommodation of land use for the future growth of cities (Colding, 2007). However, high accident rates and hazardous activities in the AEC industry not only lead to a poor reputation but pose a threat to its future innovation and evolution. The scope of a risk is very broad and consists of issues such as damage or failure of structures, injury or loss of life, budget overruns, and delays to the construction schedule, which are caused by various reasons such as design deficiency, material failure, inexperienced operatives, and weak management. For instance, in the United States, 503 bridge collapses were reported between 1989 and 2000 (Wardhana and Hadipriono, 2003), and according to official records over 26,000 workers lost their lives on construction sites from 1989 to 2013 (Zhang *et al.*, 2013). It was estimated that over 60,000 on-site fatal accidents happen every year globally (ILO, 2005). In China, although the number of construction supervision companies has increased from

52 in 1989 to 5123 in 2000 (Liu *et al.*, 2004), unwanted hazards related to safety, time, and cost were observed frequently due to poor risk management (Tam *et al.*, 2004).

An AEC project starts with planning and design followed by the construction stage lasting for months or years, and eventually the project will come into the operational period that may last for decades before demolition. Different risks may be present in each of the different stages of the project and product lifecycle. There are a wide range of risks that may lead to hazards. In recent years, with the rapid development of society, risks are gradually growing because of the increasing structural complexity and project size, and the adoption of new and complex construction methods (Shim *et al.*, 2012). To reduce the possibility of these hazards occurring and to achieve project goals successfully, there is a high demand for managing risks effectively throughout a project's lifecycle.

### 1.1.2 Challenges in traditional risk management

With their complex and dynamic nature, construction projects normally last for many years through design, construction, and operation until demolition. They require a large amount of multi-disciplinary knowledge input and both internal and external collaboration and communication to manage project risks well. In the project development process, it is the minimum and mandatory requirement to check against regulations or standards, e.g. in the UK there are the Construction Design and Management (CDM) regulations (HSE, 2015). Other methods that are commonly used in the current construction culture for reducing risks include: design review and approval, construction progress meetings and site inspections, and education and safety training for workers (Goldenhar *et al.*, 2001).

However, these are still static and traditional methods (Alaeddini and Dogan, 2011) which are heavily reliant on multi-disciplinary knowledge and experience (Shim *et al.*, 2012). As a result, many researchers (Zhang *et al.*, 2014, Hartmann *et al.*, 2012, Shim *et al.*, 2012) point out that traditional methods are facing with a number of challenges in the real world:

- Traditional risk management is a multi-disciplinary knowledge and experience based manual undertaking, which is time-consuming, error-prone and inefficient (Zhang *et al.*, 2013, Hartmann *et al.*, 2012, Shim *et al.*, 2012). For example, many designers and contractors are still working on 2D platforms and using 2D drawings to communicate project information. Detecting errors from 2D drawings is time-consuming and difficult. It is also a challenge to combine 2D drawings, site information, and text-based documents together when considering risks. As the methods are manual and mainly based on experience, observed results and decisions are often subjective and error-prone (Zhang *et al.*, 2013).

- Risk knowledge management is fragmented and insufficient, and risk knowledge transfer from project to project is difficult (Zhang *et al.*, 2013). Construction projects often last for several years and need teamwork and collaboration of a large number of participants from different disciplines. Team members gain valuable knowledge and experience from every project and most of them will leave the project after finishing their particular job. Significant project experience is held by individuals and it is difficult to capture and manage the large database of human knowledge effectively and extract valuable information for later use.

- Risk sharing and communication tends to be incomplete and inconsistent (Tah and Carr, 2001). Projects are completed by a team cooperatively, any common risks will be identified and treated individually, and the corresponding information will be documented and sometimes this work will be ignored or forgotten (Kazi, 2005). This may lead to the risk that information cannot be presented, shared, recorded, and updated effectively during the development process of a project.

### 1.1.3 Limitation in current BIM-based risk management

In response to these challenges in traditional risk management, there is currently a new research trend of utilising Building Information Modelling (BIM) and BIM-related tools to assist in early risk identification, accident prevention, risk communication, etc., which is defined as "BIM-based risk management" in this research.

BIM is a process involving the generation, exchange and management of accurate building information in digital formats, which allows better control and analysis than manual processes and could facilitate collaboration and communication (Eastman *et al.*, 2011). In this way, the use of BIM allows the visualisation of a building and simulation of its construction in a computer environment before real implementation and improves the collaboration and communication by data interoperability (Laakso and Kiviniemi, 2012), which provides a solution for the challenges of traditional risk management discussed in Section 1.1.2. With the rapid development and use of BIM, some researchers have tried to implement BIM for managing project risks in the last few years. For example, physical clashes can be detected automatically in BIM and engineers can quickly make changes or modifications to models in a parametric way (Hartmann *et al.*, 2008). Fast quantity calculation and accurate cost estimating leads to better planning and management of the project budget (Azhar, 2011). The use of open data standards could improve the data exchange between software and reduce data loss (Laakso and Kiviniemi, 2012). In addition, BIM can be used for managing some specific risks, e.g. evacuation simulation for fire accidents in buildings (Wang *et al.*, 2014), automatic checking of fall risks in BIM through model checkers (Zhang *et al.*, 2013). It can be observed from these efforts that the visualisation feature of BIM can effectively improve risk identification and analysis at an early stage.

However, although the existing efforts of using BIM for risk management have both theoretical and practical significance to the industry, most of them are still at a conceptual or prototyping stage and have not been broadly used and tested in practice (Forsythe, 2014). An extensive literature review of risk management using BIM and BIM-related digital technologies is outlined in Chapter 2 and two main gaps in current BIM-based risk management are expanded as follows:

#### ***1.1.3.1 Very few theories exist that can explain how to align traditional techniques with BIM for construction project risk management***

The current BIM-based design platforms do not have modules for risk identification, analysis and information documentation. Most of current research efforts in using BIM and BIM-related technologies for risk management focus on new developments from a technology-oriented perspective, while little is known about the possibility of integrating BIM into traditional risk management methods or existing organisational work processes. Design and construction organisations have their existing working structure and framework and it is naturally easier for them to accept new developments that have been proved to be reliable and effective.

To strengthen the practical applicability of BIM-based risk management, some studies (Hartmann *et al.*, 2012, Shim *et al.*, 2012) suggested aligning traditional techniques and processes with BIM as an integrated solution for managing project risks and have demonstrated its feasibility and benefits. For example, Shim *et al.* (2012) proposed a conceptual process model to visualise risk information in BIM for assisting design and construction management of a challenging cable stayed bridge project in Korea. Ding *et al.* (2016) presented a framework for risk knowledge management and developed a prototype with a user interface that enables it to link risk information to objects in BIM. Another study, from a technology pull view, integrated BIM into traditional construction risk management and tested its practical performance in a large infrastructure project (Hartmann *et al.*, 2012). A construction management tool called Vico Control (Vico Software, 2017) developed a risk analysis module which enables the extraction of a schedule and resource information from BIM for Monte Carlo simulation. However, very limited research or tools exists in this field. Both theoretical developments and practical studies need to be further explored.

#### *1.1.3.2 Current BIM solutions have very limited support on risk communication and information management during the project development process*

Today large construction projects are often complex in design, construction and management and their lifecycle may extend for decades, where different types of risk may exist and need to be treated in an appropriate manner. In the dynamic construction process, different disciplines need to collaborate and communicate mutually to construct a one-off product and project information is generated in and transferred between various platforms. However, a few existing studies (Kwak and Stoddard, 2004, Han *et al.*, 2008) pointed out that the importance of risk management is often overlooked and it is difficult to share and communicate risk information. Without effective information management, risk information may be lost or difficult to be transferred to the people in charge, which may lead to some major risks that are not identified and treated timely.

The current BIM-based design tools fail to support the generation and management of risk information during the project development process. Moreover, the current open BIM standards, e.g. Industry Foundation Classes (IFC), do not define schemas for risk management. A good feature of BIM is to provide a platform to support the dynamic process of construction projects through storing, using and managing digital information. For example, Goedert and Meadati (2008) demonstrated a method to integrate construction process documentation into BIM for management. In addition, Becerik-Gerber *et al.* (2011) proved that BIM can be used for facility data management. Therefore, the capability of BIM could be extended to support project risk information management and communication. The importance of effective risk information communication and management in complex construction projects as well as the potential of using BIM to facilitate communication and collaboration for project risk management have been discussed in a number of studies, e.g. (Han *et al.*, 2008, Frewer, 2004). In addition, some recent studies (Shim *et al.*, 2012) have discussed its potential and presented conceptual frameworks. Nevertheless, very limited research has been found in this area.

## 1.2 Research questions

In response to the two knowledge gaps discussed in Section 1.1.3, the following two main research questions were formulated after an extensive review of relevant literature in Chapter 2 on managing construction project risks through both traditional methods and processes, and BIM-related technologies:

- **Research Question 1 (RQ1)**: How can existing techniques align with BIM for construction project risk management?

- **Research Question 2 (RQ2)**: How can risk knowledge and information be managed and visualised in the BIM environment during the project development process?

The two research questions are connected to each other. RQ1 explores why, what and how some existing risk management techniques can be integrated into BIM. RQ2 investigates how the generated risk knowledge and information can be visualised and managed in the BIM environment.

## 1.3 Intuition

The intuition of this research is that some existing risk management techniques can be integrated into BIM-based platforms and an active linkage between the risk information and BIM can be established to support the project lifecycle. This proposed concept of a BIM and Knowledge based Risk Management System (BKRMS) is presented in Figure 1. The motivations of the proposed solution are described in Section 1.4 and the starting points for the proposed solution are discussed in detail in Chapter 4. The principle behind this solution is that traditional risk management techniques can be integrated into BIM-based platforms and an active linkage between the risk information and BIM can be established to support the project lifecycle.

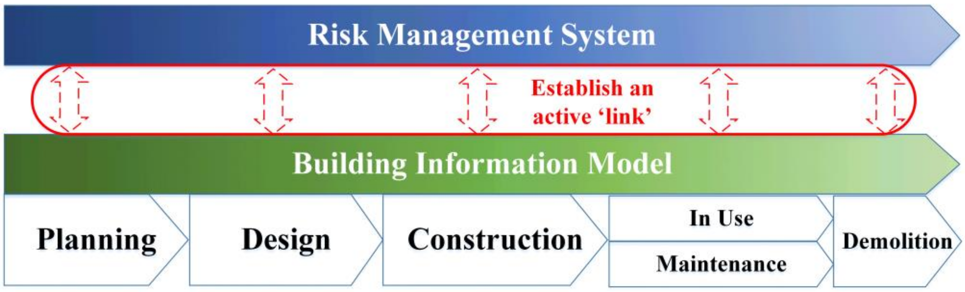



Figure 1. Establishing active linkages between BIM and Risk Management System

## 1.4 Overview of point of departure

The proposed framework of BKRMS is illustrated in Figure 2 and consists of two main modules: the BIM Module and the Risk Module where the two modules are linked to each other. In the BIM Module, design, environmental and other general project information can be captured firstly to generate a 3D BIM. The second step is to collect and analyse the construction information, schedule and work breakdown tasks, which can be connected with the 3D BIM to generate a 4D BIM. The Risk Module provides a user interface for managing any updates and changes of identified risks in BIM and is supported by two traditional risk management techniques, Risk Breakdown Structure (RBS) and Case-based Reasoning (CBR). On one hand, the RBS can be used as a knowledge-based approach to classify, store and manage the information of a risk database in a proper structure and risk information in RBS could be linked to 3D/4D BIM for review, visualisation and communication. On the other hand, the Case-based Reasoning Library (CBRL) is a collection of both successful and unsuccessful cases that can provide risk management knowledge about project cases for analysing the ongoing project and helping decision makers investigate possible solutions. In addition, the CBRL is able to retain and update new cases from the BIM Module.

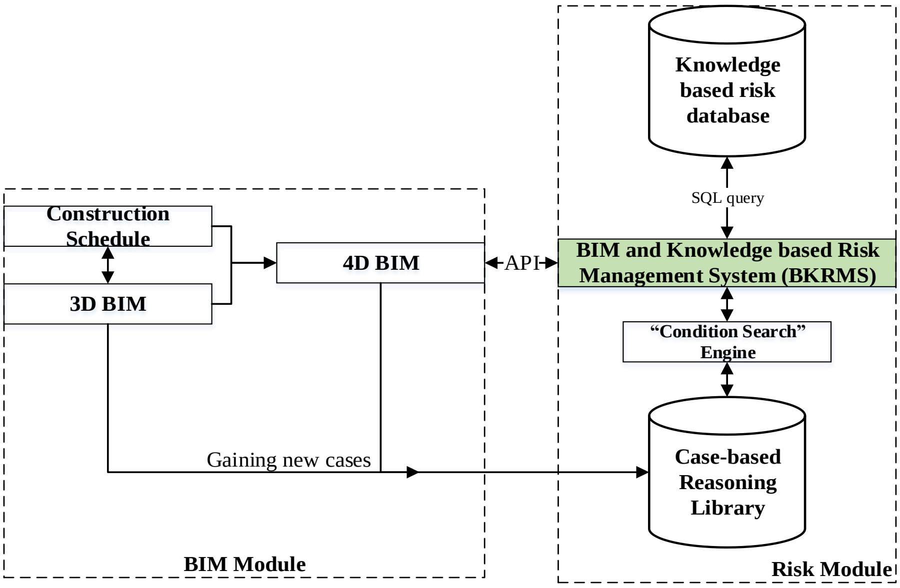



Figure 2. Framework of the BKRMS

The idea of this research was motivated by the following observations. First, Kiviniemi (2005) demonstrated a methodology to manage user requirements during the lifecycle of a project by establishing an active link between requirements models and building information models, and successfully illustrated that user requirement information can be divided into different levels and linked with BIM. For this research, potentially the linkage approach can be adopted to address the risk information management and support the lifecycle project risk management. Its feasibility and benefits have been discussed and proved by a number of existing studies (Shim *et al.*, 2012, Ding *et al.*, 2016).

Secondly, the study conducted by Shim *et al.* (2012) presented a conceptual diagram for visualising risk information in BIM and pointed out that RBS has the ability to facilitate the understanding and communication of risks in risk identification and analysis processes. RBS, in concept, is a hierarchical structure that allows all types of risk factors and events to be organised by groups and categories (Holzmann and Spiegler, 2011). It is an open, flexible and easily updatable tool and could offer a global view on risk exposure (Tah and Carr, 2001, Mehdizadeh *et al.*, 2013). The main advantages of RBS include: 1) to increase overall understanding of risks and facilitate risk communication; 2) to help locate identified risks into relevant risk categories and make special strategies to treat them easily; and 3) to provide an architecture for managing the risk database and developing risk management software.

Thirdly, in general, the process of solving new problems based on experience of similar past problems is known as Case-Based Reasoning (CBR) (Jonassen and Hernandez-Serrano, 2002), which examines what has taken place in the past and applies it to a new situation (Kolodner, 1993). It could be of particular help in identifying and mitigating project risks at early stages, e.g. design and construction planning. In order to facilitate CBR for practical use in the construction industry, some efforts have been observed in collecting risk cases and establishing a risk case database, e.g. (Structural-Safety, 2016, Zhang *et al.*, 2016). However, as a risk case database often contains a huge amount of data where reports are written in everyday human language, manually reviewing, analysing and understanding these reports is time-consuming, labour-intensive and inefficient work. Failure in extracting 'correct' cases and information within a limited time may mean that the importance of learning from past experience is missed. In recent years, with the development and growing use of Natural Language Processing (NLP) in the computer science discipline, some researchers have been trying to introduce NLP into the construction industry to address the analysis and management issues of textual documents, e.g. retrieval of CAD drawings (Hsu, 2013), automatic analysis of injury reports (Tixier *et al.*, 2016), and automatic clustering of construction project documents based on textual similarity (Al Qady and Kandil, 2014). It could be seen that NLP is a promising technique in assisting the knowledge and case retrieval of CBR.

## 1.5 Research aim and objectives

The overall aim of this research is to develop a methodology of integrating traditional techniques with BIM for project risk management and information sharing and improve the applicability of BIM-based risk management.

In order to accomplish the research aim, conceptual frameworks were designed in Sections 1.3 and 1.4 on top of three main starting points (Chapter 4), and the following three research objectives were formulated:

**Objective 1: To develop a RBS and a Case-based Reasoning Library (CBRL) for bridge projects.**

The first research objective involves two tasks:

- Collection and detailed analysis of project risks from a holistic view.
- Development of a tailored RBS for classifying and managing these project risks and a CBRL as a knowledge database.

**Objective 2: To develop a linkage between Risks and BIM for bridge projects.**

The second research objective consists of three tasks:

- Development of the linkage rules between the proposed RBS and BIM.
- Development of a tool prototype for implementation of the proposed linkage.
- Validation of the linkage approach and tool prototype.

**Objective 3: To develop a method to support the fast risk case retrieval from the CBRL.**

The third research objective consists of three tasks:

- Development of a NLP based approach for quickly and accurately retrieving valuable data from risk and accident report documents.
- Development of a tool prototype for implementation of the proposed approach.
- Validation of the proposed method and tool prototype.

## 1.6 Research method

The research method is briefly introduced in Sections 1.6.1-1.6.5 and further explained and discussed in Chapter 3.

### 1.6.1 Literature Survey and Solution Development

A comprehensive literature review was conducted in the first stage of this research to: 1) identify the knowledge gaps that currently exist in BIM-based risk management, 2) explore the potential solution to fill the identified gaps, and 3) limit the scope of this research. The detailed analysis of literature and existing gaps are discussed in Chapter 2 and the Point of Departure of the proposed solution is introduced in Chapter 4.

### 1.6.2 Development of the RBS and CBRL

Based on the proposed solution, the next stage is to further collect, analyse and classify the risks to develop a RBS and establish the CBRL for bridge projects. The development process of the RBS is introduced in Chapter 5. The CBRL here refers to a collection of risk case documents written in everyday language and can be used to support the risk identification and analysis, and decision making during the project process. The collection and development of the CBRL is illustrated in Section 7.2.

### 1.6.3 Development of linkage between BIM and RBS

With the observation that different types of risk can affect the project differently, an active linked relationship was established between the RBS and BIM. To support the linkage, a new framework of the BIM-based risk management system was proposed and prototyping was used for the implementation and validation of the proposed approach and tool prototype. This work is introduced in Chapters 5 and 6.

### 1.6.4 Information retrieval from the CBRL

Built on the established CBRL, this research proposed a method through combining the use of two NLP techniques, i.e. Vector Space Model (VSM) and semantic query expansion, for risk case retrieval and outlines a framework for this risk case retrieval system. A prototype system was developed using the Python programming language to support the implementation and validation of the proposed method. This work is presented in Chapter 7.

### 1.6.5 Validation of the research

The purposes of the validation are to check: 1) if the proposed solution is able to address the identified knowledge gaps in current BIM-based risk management, and 2) if the proposed BKRMS can be implemented for practical use and integrated into the risk management process.

The proposed methodology was primarily validated through prototype development and testing, where the prototype refers to the common practice in software engineering to develop an early release or incomplete versions of a software to simulate and test a concept or process (Smith, 1991). Specifically, one prototype, i.e. a Navisworks plugin, was developed to support the linkage between the RBS and BIM, and its implementation process is illustrated by an example case study (Sections 6.3 and 6.6); another prototype, i.e. a risk case retrieval system, was developed for fast information retrieval from the CBRL and the system validation is discussed in Section 7.6.

In addition, a group of industry experts were invited to participate in semi-structured interviews to evaluate the feasibility and usefulness of BKRMS from an industry perspective, where comments and suggestions for future research were also noted (Section 8.1).

## 1.7 Guide to the thesis

The thesis is organised into eight chapters. The outline of the chapters is as follows:

Chapter 1 Introduction – summarises the observed problem, research questions, intuition to address the knowledge gaps, overview of the starting points, research objectives and methods.

Chapter 2 Background and Literature Review – presents a summary of existing risk management methods and processes, and the state of the art of BIM-based risk management. The research problems and knowledge gaps are concluded from the core findings of this chapter.

Chapter 3 Research methodology – discusses the methodology used in this research to determine the research questions.

Chapter 4 Point of Departure – describes what existing knowledge and methods contribute to the starting points for the proposed solution presented in Chapters 5-7.

Chapter 5 Risk Analysis and RBS Development – discusses the conceptual separation of BIM and the process of developing a tailored Risk Breakdown Structure (RBS) for bridge projects and formalising an active link between the resulting RBS and BIM.

Chapter 6 Integrating RBM into 3D/4D BIM for Risk Identification and Communication – builds on the results of Chapter 4 and introduces a method of interconnecting BIM, RBS and Work Breakdown Structure (WBS) as a risk management system. It then develops prototype and tests the implementability of the proposed system through a case study.

Chapter 7 Retrieving Similar Cases from the CBRL using NLP Techniques – describes a method for retrieving similar cases from the CBRL through combining the use of two NLP techniques. It then introduces the development of a prototype and tests the functionality of the proposed method using a number of sample queries.

Chapter 8 Discussion and Conclusions – firstly describes the validation of the usefulness of the proposed methods presented in Chapters 5-7 through interviews with industry experts. This chapter then completes the thesis through providing a summary of the work involving theoretical and practical contributions, limitations and suggested future research, and conclusion of the research.

The overall thesis structure is presented in Figure 3.

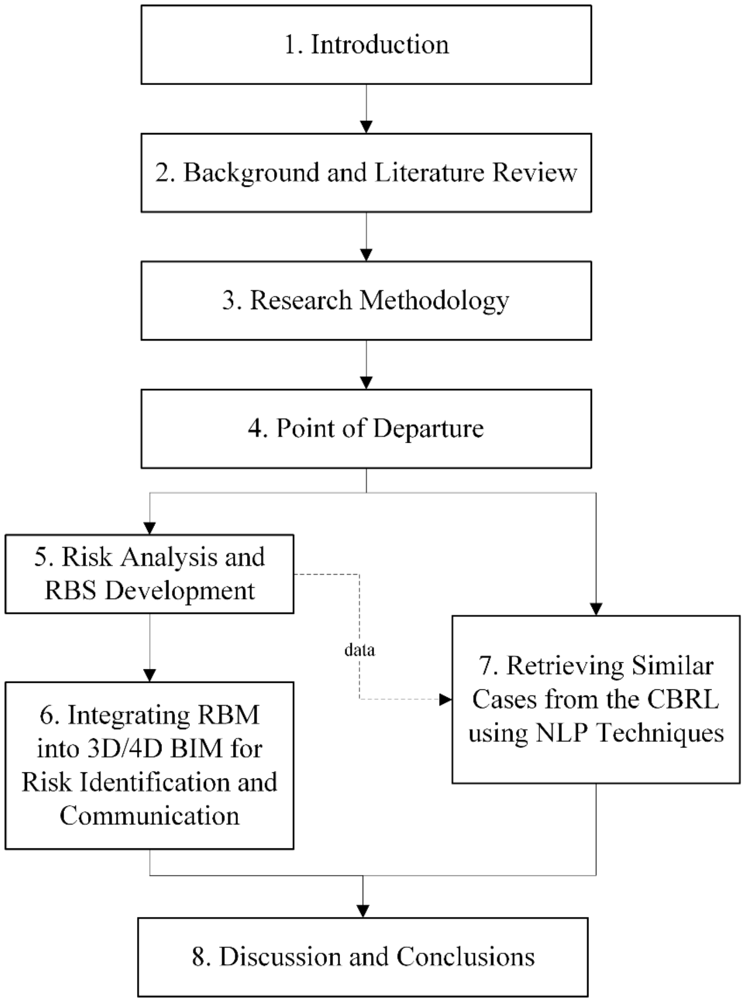



Figure 3. Structure of the thesis

## 1.8 Scope and limitations of the research

The purpose of this research is to link risk knowledge and information to BIM. The risks of BIM technologies and implementation were not within the scope. The focus of the research has been identifying typical risks that may affect a construction project (e.g. design error, structural safety and financial risk) and developing the knowledge-based risk database and CBRL.

Bridges were chosen as the type of projects to study the RBS and the linkage between the resulting RBS and BIM. Although the scope of this research is limited to bridges, the basic methods and principles presented in this research could be also applied to other types of AEC project.

Another aim of this research was to study the use of NLP into CBR to support lifecycle risk management and decision making. Generally the implementation cycle of CBR contains four main processes: RETRIEVE, REUSE, REVISE, and RETAIN (known as ‘the four REs’), where RETRIEVE is the first and the most important process in any CBR systems (Lu *et al.*, 2013). Only risk case retrieval (i.e. RETRIEVE) was within the scope of this research.

# Chapter 2. BACKGROUND AND LITERATURE REVIEW

## 2.1 The fundamentals of risk management

The term "risk" was known in the English language from the 17th century and was derived from an original meaning to run into danger or to go against a rock (McElwee, 2007). Today the concept of risk is adopted in many different fields and with a variety of different words, such as "hazard", "threat", "challenge", or "uncertainty". In the AEC industry, risks have a two-edged nature, e.g. "*the likelihood of unwanted hazards and the corresponding consequences*" (Zou *et al.*, 2007), "*the likelihood and consequence of risks*" (Williams, 1996), "*a combination of the likelihood and consequences of the hazard*" (Vrouwenvelder *et al.*, 2001).

Risk management is a system aiming to recognise, quantify, and manage all risks exposed in the business or project (Flanagan and Norman, 1993). PMBOK® (Project Management Body of Knowledge) describes it as a process in relation to planning, identifying, analysing, responding, and monitoring project risks and one of the ten knowledge areas in which a project manager must be competent (PMI, 2004). The International Organization for Standardization (ISO, 2009) defines the process of risk management involving applying a systemic and logical method for establishing the context, creating a communication and consultation mechanism, and constructing risk management identification, analysis, evaluation, treatment, monitoring, and recording in a project. In accordance with these definitions, risk management in the AEC context is a logical, systematic, and comprehensive approach to identifying and analysing risks, and treating them with the help of communication and consultation to successfully achieve project goals. The systematic process includes risk identification, analysis, evaluation, treatment, monitoring and review (Banaitiene and Banaitis, 2012, ISO, 2009, Zou *et al.*, 2007), where risk identification aims to find out the range of potential risks and risk analysis plays a core role in the whole process. When risks cannot be eliminated, early and effective identification and assessment of risks become necessary for effective risk management in a successful project (Zou *et al.*, 2007). All activities of a project involve risks (ISO, 2009) and there is an immediate and direct relationship of objectives between the whole project and risk management.

A set of techniques has been developed to identify, analyse and evaluate risks. The techniques, according to ISO (2009), can be divided into qualitative and quantitative analysis. The former includes Delphi, checklists, strength-weakness-opportunity-threats (SWOT) analysis, risk rating scales, etc., while the latter includes environmental risk assessment, neural networks (NN), row tie analysis, reliability centred maintenance, risk indices, and others. However, although the above methods are important techniques for risk management, they are confined to static control management and play only a limited role in practice (Zhang *et al.*, 2014). The implementation of traditional risk management is still a manual undertaking, the assessment is heavily reliant on experience and mathematical analysis, and the decision making is frequently based on knowledge and experience based intuition, which always leads to a decreased efficiency in the real environment (Shim *et al.*, 2012).

## 2.2 The general process of risk management

Based on a review of the literature, expert interviews, and the author's project experience, the current general risk management framework used in the UK AEC industry is summarised in Figure 4. The framework prescribes a long-term risk management strategy and a process that allows participants to work collaboratively to manage risks in a systematic way. The core philosophy of this method, defined in the Risk Mitigation Model, is that the main scope for identifying and mitigating risks should be as early as possible, especially in the design or planning phases, which is regulated in the UK's Construction Design and Management (CDM) Regulations (HSE, 2015). Ideally, most of the foreseeable risks should be "designed out" during the planning or design stages, and the residual risks should be managed during the construction and subsequent phases.

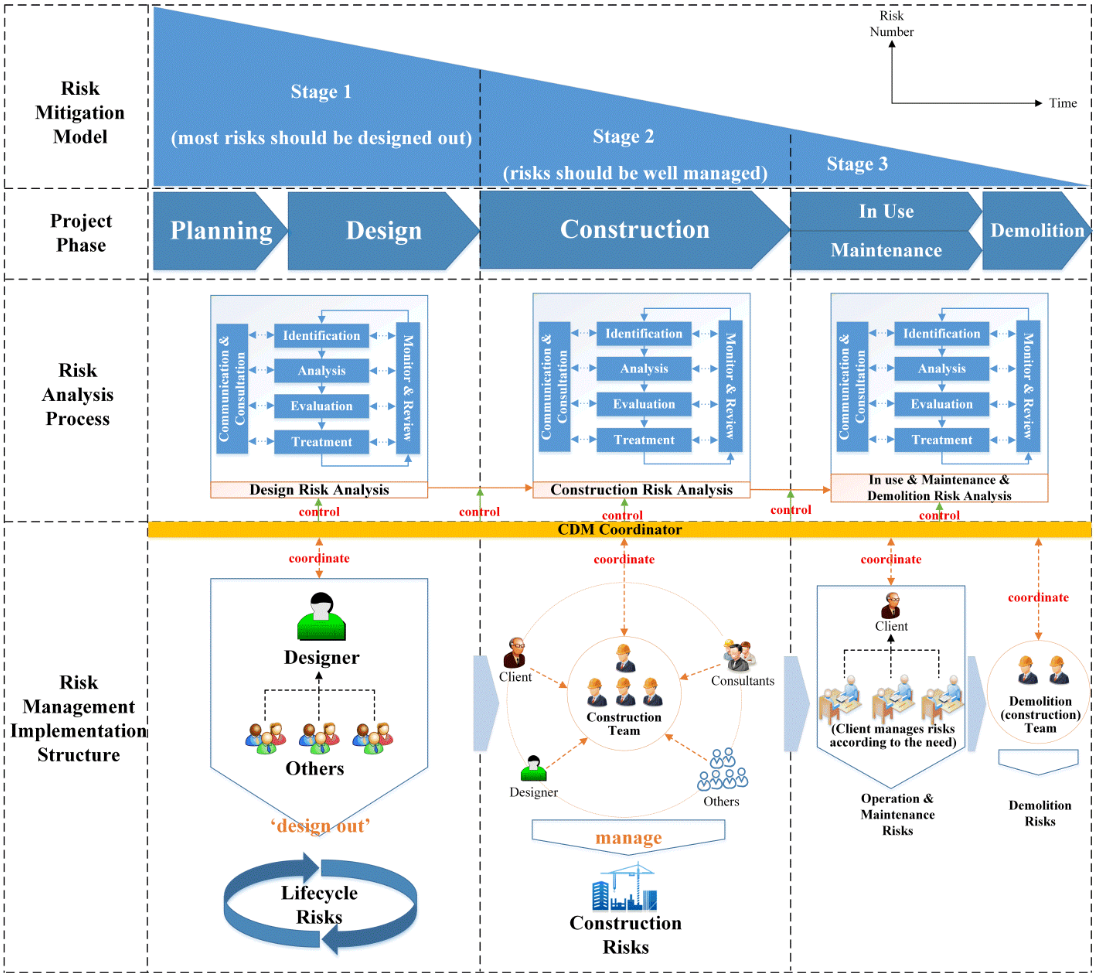



Figure 4. General Risk Management Framework

However, some challenges in the above process are: (1) in-time knowledge capture and analysis, (2) the management of multi-disciplinary knowledge and experience, and (3) effective communication environment. Valuable knowledge and experience are gained from previous projects and this can be used to contribute to future work. In this case, the effective management of this large database of human knowledge and experience, as well as flexible and accurate data extraction, become a precondition for the success of risk management. As the project is handed over from designer to contractor, and then from contractor to the client, people will normally leave the project after completing their tasks and large amounts of risk information may be lost if it is not properly recorded and communicated to other project participants (Kazi, 2005).

## 2.3 Information and Communication Technologies (ICT) for risk management

To overcome these obstacles, ICT, e.g. BIM, 4D CAD, and Virtual Reality (VR), has been applied in the AEC industry to manage risks. For instance, construction safety risk planning and identification is an issue addressed by 3D/4D visualisation (Hartmann *et al.*, 2008). BIM could help automatically detect physical spatial clashes (Chiu *et al.*, 2011) and specific requirements of building codes could be interpreted to machine-read rules and checked automatically in Industry Foundation Classes (IFC) information models (Eastman *et al.*, 2009). Heng Li *et al.* (2013) presented a proactive monitoring system using Global Positioning System (GPS) in combination with Radio Frequency Identification (RFID) to improve the safety of blind lifting of mobile/tower cranes. The next section will review and discuss these developments critically in detail.

Two reasons could explain the increasing interest and adoption of ICT for risk management. The first reason is that as the industry has benefited from salient technical advantages of BIM and other digital technologies, a natural consequence is to investigate their possibilities in risk management. These new techniques could not only provide new design tools and management methods (Eastman *et al.*, 2011) but significantly facilitate the collaboration, communication, and cooperation for both within and between organisations (Dossick and Neff, 2011), which are essential requirements for managing risks successfully. The second reason comes from a strong thrust from the government policy makers who have realised the importance of integrating ICT with risk management.  Evidence of this is the new version of CDM regulations that will cover ICT such as BIM after 2015 (Joyce and Houghton, 2014) replacing the older version that was introduced in the UK initially in 1996 for improving safety and risk management.

## 2.4 Survey of BIM and BIM-related technologies for managing risks

The state-of-the-art of the use of BIM and BIM-related technologies for risk management is summarised in this section. The technologies referred here include BIM, automatic rule checking, knowledge based systems, reactive and proactive safety systems based on information technology. There is a distinct difference between reactive and proactive safety systems for risk management. Forsythe (2014) and Teizer *et al.* (2010) pointed out reactive systems using information technologies such as VR, 4D CAD, and GIS seldom use real-time data and need a post data collection processing effort for analysis, while in contrast proactive technologies can collect and analyse real-time data, and provide real-time warning and immediate feedback to construction site about dangers in time. It has been found that BIM, on one hand, can be used as a systematic risk management tool in the development process and, on the other, can perform as a core data generator and platform to allow other BIM-related tools for further risk analysis, where most of these technologies can be used interactively in related investigations.

### 2.4.1 Managing risks through BIM

Over the last few years, with the rapid development of theory and computer applications, BIM has achieved a remarkable awareness in the AEC industry and there is a significant increase of the adoption of BIM to support the planning, design, construction, operation and maintenance phases (Volk *et al.*, 2014). Instead of being just considered as a technology, BIM is becoming a systematic method and process that is changing the project delivery (Porwal and Hewage, 2013), designing (Liu *et al.*, 2014), and the communication and organisational management of construction (Hardin, 2011). Although most papers utilising BIM as an advanced tool to manage project risks such as design errors, quality, and budget do not often refer to risk management intentionally, the process of applying BIM can be seen, to some extent, as a systematic way for managing risks. Examples are presented in Table 1.

Table 1. Examples for applying or developing BIM for risk management

| Functionality | Benefits for risk management | Research | Practice |
|---|---|---|---|
| 3D visualisation | Facilitating early risk identification and risk communication | (Hartmann *et al.*, 2008) | (Liu *et al.*, 2014, Shim *et al.*, 2012) |
| Clash detection | Automation of detecting physical conflicts in model | (Hartmann *et al.*, 2008, Tang *et al.*, 2011) | (Liu *et al.*, 2014, Chiu *et al.*, 2011) |
| 4D construction scheduling/planning | Facilitating early risk identification and risk communication; improving construction management level | (Hartmann *et al.*, 2008, Hardin, 2011, Whyte, 2002) | (Liu *et al.*, 2014, Chiu *et al.*, 2011) |
| 5D cost estimation or cash flow modelling | Planning, controlling and managing budget and cost reasonably | (Hartmann *et al.*, 2008, Hardin, 2011, Whyte, 2002, Marzouk and Hisham, 2014) | (Motawa and Almarshad, 2013) |
| Construction progress tracking | Improving management level for quality, safety, time, and budget | (Eastman *et al.*, 2011, Bhatla *et al.*, 2012) | - |
| Safety management | Reducing personnel safety hazards | (Whyte, 2002, Teizer, 2008) | - |
| Space management | Improving the consideration of space distribution and management in design | (Hartmann *et al.*, 2008, Kim *et al.*, 2012) | - |
| Quality control | Improving construction quality | (Chen and Luo, 2014) | - |
| Structural analysis | Improving structural safety | (Sacks and Barak, 2008, Seung-Il Lee *et al.*, 2012, Shim *et al.*, 2012) | (Liu *et al.*, 2014) |
| Risk scenario planning | Reducing personnel safety hazards | (Hardin, 2011, Azhar, 2011) | (Hartmann *et al.*, 2012) |
| Operation and maintenance (Q&M), facility management (FM) | Improving management level and reducing risks | (Volk *et al.*, 2014, Becerik-Gerber *et al.*, 2011) | - |
| Interoperability | Reducing information loss of data exchange | (Laakso and Kiviniemi, 2012, Ji *et al.*, 2013) | - |
| Collaboration and communication facilitation | Facilitating early risk identification and risk communication | (Dossick and Neff, 2011, Grilo and Jardim-Goncalves, 2010, Porwal and Hewage, 2013) | - |
| Urban planning and design | Integrating planning and design of urban space and AEC projects; facilitating land-use planning, design and management | (Ghang Lee *et al.*, 2012, Kim *et al.*, 2011, Rajabifard *et al.*, 2012) | (Ghang Lee *et al.*, 2012) |

In the planning and design stages, one of the main risks is how the design aligns with the determined project feasibility, secured budget, and established governance regime (Miller *et al.*, 2001). This is an area where BIM has the potential to manage the risks. For example, the visualisation of preliminary design by 3D/4D models could help engineers build and modify the model quickly in a parametric way to meet the stakeholders' requirements (Hartmann *et al.*, 2008). The short videos or virtual walkthroughs which simulate the view of a person walking through the building can rapidly improve stakeholders' understanding of the project (Whyte, 2002). Meanwhile, neutral data formats such as the IFC that store standard and customised data for all project elements could provide an interoperable digital representation of all project elements enabling interoperability between BIM software applications (Laakso and Kiviniemi, 2012), which could increase the repeated use of data and reduce the possibility of errors.

At the construction stage, there is often a huge pressure for the construction team to complete the project safely within budget and schedule, and various risks and uncertainties exist in this period. To identify construction risks at an early stage and optimise the construction sequences, Chiu *et al.* (2011) conducted a clash detection and a 4D simulation of the construction of a steel bridge. Chen and Luo (2014) extended the 4D model to cover quality management based on construction codes and established a quality control model in a product, organisation and process (POP) data definition structure, which was used and validated in the construction of the Wuhan International EXPO centre. In addition, Marzouk and Hisham (2014) used BIM's ability of cost estimation to develop an application that integrates BIM with Earned Value (EV) for cost and schedule control, and determines the project status at specific reporting dates for infrastructure bridges.

It has also been found in this review that although the majority of efforts still focus on applying BIM to the design and construction phase, BIM can also be used in other processes and phases, e.g. facility management (Becerik-Gerber *et al.*, 2011), maintenance management (Volk *et al.*, 2014), and demolition (Cheng and Ma, 2013). In addition, a BIM-based collaboration and communication environment could naturally facilitate the early risk identification and mitigation (Dossick and Neff, 2011, Grilo and Jardim-Goncalves, 2010).

### 2.4.2 Knowledge based systems

In the AEC industry, every project produces valuable knowledge and experience which can contribute significantly to managing risks in future projects. It is essential to manage this information properly and communicate it effectively in all stages of the whole project lifecycle (Tah and Carr, 2001). This idea has been recognised and adopted for a long time by researchers to manage project risks. For example, Total-Safety (Carter and Smith, 2006) is a method statement development module within an ICT tool that could assist engineers to formulate method statements with a high level of risk identification by extracting safety information from a knowledge based database. When a construction method is chosen, the tool can return all known risks associated with different tasks as the knowledge basis for further risk assessment. Similarly, Cooke *et al.* (2008) proposed a web-based decision support program named ToolSHeD to integrate assessment of safety risk into design process. The principle of ToolSHeD is to structure the knowledge obtained from industry standards, national guidelines and codes of Australia, and other information sources, and employ this knowledge for assessing risks in complicated situations of buildings.

The integration of BIM and knowledge based systems has been seen as a new trend. Deshpande *et al.* (2014) proposed a new method to capture, extract, and store information and knowledge from BIMs, and presented a framework for classification and dissemination of the knowledge. To strengthen its practical application, Ho *et al.* (2013) developed a BIM-based Knowledge Sharing Management (BIMKSM) system that could enable managers and engineers to share knowledge and experience in the

BIM environment. Aiming at managing safety risks in design, Qi *et al.* (2011) developed a dictionary of construction worker suggestions and a constraint model to store the formalised suggestions. Then in the BIM environment, designers could utilise rule checking software for identifying safety risks during the planning and design phases, and mitigating risks and optimising their designs. The system consists of three parts: BIM as the main information input, a knowledge based system, and a risk identification module. Motamedi *et al.* (2014) integrated the use of knowledge management (KM) and BIM to investigate an approach for detecting failure root-cause which could help facility management (FM) technicians identify and solve problems from their cognitive and perceptual reasoning. Integrated with BIM, a Computerised Maintenance Management System (CMMS) was developed to store inspection and maintenance data. In addition, a knowledge based BIM system was presented by Motawa and Almarshad (2013) to capture and store various types of information and knowledge created by different participants in the construction project in order to support decision making for building maintenance.

### 2.4.3 Automatic rule checking

In definition, the term *Automatic Rule Checking* is the use of a computer program to assess a design based on objects' configuration (Eastman *et al.*, 2009) and its purpose is to encode rules and criteria by interpretation and thus building models could be checked against these machine-read rules automatically with results, for example, "pass", "fail", "warning", or "unknown" (Borrmann *et al.*, 2009).

Regulations and rules written by experts have traditionally been comprehended, interpreted and used in a manual way. Thus, these rules are sometimes conflictive and incomplete, and the corresponding implementation is often limited by people's understanding, interpretation, and reasoning capability. To computerise this process and improve the effectiveness, the research of automatic code checking or rule compliance started in the 1960s. Soon afterwards, a lot of effort was put into interpreting particular requirements to computerised codes, logically structuring and managing rules, and developing rule-based systems (Fenves, 1966, Rasdorf and Lakmazaheri, 1990, Fenves *et al.*, 1995, Garrett Jr and Fenves, 1987). In the late 1990s, due to the fast growth rule-based systems for building models, the development of IFCs brought on the initial exploration of building model schema for checking building codes. This review has observed three development directions in the area of automatic rule checking during the last two decades – (1) building design codes compliance, (2) construction safety checking, and (3) special requirements checking, which will be discussed further in detail below. A comprehensive review, which introduced the main steps and software platforms of automatic rule checking, was reported by Eastman *et al.* (2009).

The most common application of rule checking is to ensure the design work is compliant with numerous building codes, which are normally known as the minimum standards for construction objects such as buildings and infrastructure projects. To computerise this work, two major activities are needed to achieve this goal: 1) to formalise the building code and BIM into building rule models and building design representation models respectively; and 2) to implement both models in computer programs and execute rule objects over design objects in compliance checking automatically (Yang and Xu, 2004). Substantial efforts in this area have been made in recent years. For example, Delis and Delis (1995) proposed a method which could encode fire code requirements in a knowledge based system for analysing the performance of fire safety in the completed building design. Balachandran *et al.* (1991)

developed an approach to processing non-measurable code provisions for verifying building designs automatically. Solihin (2004) developed the e-PlanCheck system by using the IFC model and Express Data Manager (EDM) for assessing the code compliance in Singapore. One of the latest efforts in this area is an on-going project in the US funded by Fiatech to develop AutoCodes expecting to improve automatic code checking capability for BIM standards and guidelines, and US building model codes (Fiatech, 2013).

The second development direction is to check construction safety rules. To prevent any human safety accidents on site, it is essential to identify and mitigate these risks in design, and inspect, monitor and manage safety in construction. Hence the design stage is the best opportunity to mitigate most of these risks if potential hazards could be well identified and planned, and corresponding measures to control these risks can be chosen correctly (Bansal, 2011). Yi and Langford (2006) collected and analysed historical safety records and proposed a theory that could estimate a project's risk distribution. Sulankivi *et al.* (2013) presented a theory to identify safety risks, which are unknowingly built into the construction activities at the design stage and developed a BIM-based automatic safety rule-checking prototype. The approach works by simulating the construction sequences and tasks with embedded safety rules. Aiming at fall protection, Zhang *et al.* (2013) formalised the fall protection rules of the Occupational Safety and Health Administration (OSHA) and other best practices into a table-based safety rule translation algorithm, and implemented a rule-based checking system in BIM to plan and simulate safety issues at an early stage. The feasibility has been shown by implementing this approach in Tekla Structures.

The last application direction of development is for checking specific requirements of buildings, such as the circulation problems, space requirements, and special site considerations. For instance, Han *et al.* (2002) presented a hybrid method that used encoding prescriptive-based provisions and supplemented them with a performance-based approach to facilitate conformance and applicability analysis for accessibility. Lee (2010) developed a new approach to checking occupant circulation rules automatically in the US Courts Design Guide, which could assist circulation rule checking in the development processes of a courthouse's design. Lee *et al.* (2010) proposed a computational approach called the Universal Circulation Network (UCN) for checking walking distances between buildings by implementing a length-weighted graph structure for building models, and developed a plug-in on top of the Solibri Model Checker.

### 2.4.4 Safety risk management through reactive IT-based safety systems

The AEC industry is still faced with a particular challenge of high accident rates – over 6 percent in Hong Kong for instance (OSHC, 2008). To detect health and safety (OHS) risks in time and mitigate them before any hazards occur, reactive IT-based safety systems have been used in conjunction with BIM to achieve this goal. Forsythe (2014) and Zhou *et al.* (2012) summarised these technologies including, for example, database technology, Virtual Reality (VR), 4D CAD, Geographic Information Systems (GIS), which are discussed in this sub-section.

#### *2.4.4.1 Database technology*

Experience and knowledge learned from past accidents provide a better perception to prevent hazards in future work (Gambatese *et al.*, 2005). An obvious step from this is database technology that could be used to store valuable knowledge, capture accurate information and then intelligently extract them based on specific selection criteria (Forsythe, 2014). For example, Imhof (2004) collected 360 cases of bridge failures and established an online database to help learn from past accidents, analyse the risk distribution and summarise the main risk factors that led to bridge collapse, which allows a better understanding of the mechanism of an accident and a better insight of how to prevent hazards in the future. Yu (2009) developed a knowledge based decision support model on the basis of knowledge representation and reasoning features to assist clients to evaluate competence of potential designers, principal contractors, and CDM coordinators. Furthermore, to improve the performance and capability, an enhanced online database called Construction Safety and Health Monitoring (CSHM) system was developed to enable remote access, speedy data collection and retrieval, and expert communication (Cheung *et al.*, 2004).

#### *2.4.4.2 Virtual Reality*

Virtual Reality (VR) is an important area in current BIM research and vice versa (Gu and London, 2010). Conceptually, VR is a virtual system that consists of a computer capable of real-time animation, controlled through a group of equipment for simulating physical presence in places in the real world (Steuer, 1992). VR has been used to provide a 3D, virtual and interactive computer environment for training site workers to become aware of identified on-site safety risks (e.g. (Guo *et al.*, 2012)) and formalising strategies and measures of potential hazards by simulating the dangerous scenarios (e.g. (Wang *et al.*, 2014)). Specifically, Guo *et al.* (2012) presented a game based interactive multi-client platform for safety training to improve construction site operation safety. Embedded with identified hazards, the platform provides a virtual environment where trainees can learn and practice operating methods and construction sequences, which closely resemble the real working on-site environment. The presented platform also encourages trainees to work collaboratively with others in operating the construction site. Although technological development looks extremely important in VR for managing safety risks, how these developed technologies could be adopted and implemented in practice becomes another concern. Therefore, after summarising the main factors that may cause construction accidents, Guo *et al.* (2013) proposed a conceptual framework to adopt Virtual Prototyping (VP), consisting of three core components: (1) modelling and simulation, (2) identification of unsafe factors, and (3) safety training, to support construction health and safety risk management for both technicians and workers. For improving the building emergency management, Wang *et al.* (2014) developed a BIM based virtual environment (BIM-VE) to address two key issues: "*(1) timely two-way information flow and its applications during the emergency and (2) convenient and simple way to increase evacuation awareness*". In addition, VR can also be incorporated with database technology for managing construction safety risks. For example, Hadikusumo and Rowlinson (2002), Hadikusumo and Rowlinson (2004) created a design-for-safety-process (DFSP) tool to aid safety risk identification when producing the construction plans and schedules in the design stage. This tool comprises three components: (1) the DFSP database, (2) the virtual reality construction components and processes, and (3) virtual reality functions. The DFSP database stores a full list of common dangerous conditions and actions, local accident reports and rules. The integration of the VR components and DFSP database allows users to walk through in a virtual project environment from a first-person view and to identify safety risks within construction components and related processes, and to choose preventative measures for those identified risks.

#### *2.4.4.3 4D CAD*

Early research of applying four-dimensional computer aided design (4D CAD) for construction planning to identify potential problems, mitigate risks, and optimise construction schedule and processes started in the early 1990s (Heesom and Mahdjoubi, 2004). The core concept of 4D CAD is to add 4D construction schedule information into a 3D model to establish a collaboration and communication media and clear visual insights of the construction sequences for the construction team (Koo and Fischer, 2000). It is observed that the most common application of 4D CAD for safety risk management is to establish an extensive 4D CAD model by gathering all design data about building objects and construction processes, activities and sequences, and conduct further risk analysis on the basis of the model. For instance, Benjaoran and Bhokha (2010) presented a 4D CAD model to integrate safety risk and construction management. Rule-based algorithms for working-at-height risks were formalised, interpreted, and visualised into the model. A rule-based system was then used to extract information from the 4D CAD model to detect working-at-height risks automatically and forecast necessary measures including safety activities and requirements. In structural analysis, Hu and Zhang proposed a new method in their two papers (Hu and Zhang, 2011, Zhang and Hu, 2011) to analyse safety and conflict by incorporating BIM, 4D CAD, time-dependent structural analysis, and clash detection, and then implemented this theoretical solution by developing an integrated archetypal system named 4D-GCPSU 2009. A group of researchers from Finland's VTT Technical Research Centre demonstrated a BIM-based safety management and communication system that develops construction procedures and BIM for 4D safety planning, management, and communication, where BIM and 4D CAD are utilised as the central technologies (Kiviniemi *et al.*, 2011).

#### *2.4.4.4 Geographic Information Systems*

While BIM is defined to develop objects' geometric data into the maximum level of detail, a Geographic Information System (GIS) is a collection of environmental information from the macro perspective (Irizarry and Karan, 2012, Zhou *et al.*, 2012). GIS can be integrated into a Decision Support System (DSS) to monitor and control safety risks (Cheng *et al.*, 2002). Along a similar line, Bansal (2011) successfully applied GIS to predict places and activities where there was an increased likelihood of hazards in a building project in India because BIM and 4D modelling could not provide the capability for features like 3D components editing, topography modelling, geospatial analysis, and generation and updating of schedules. Bansal and Pal (2007) also proved GIS has the potential to help cost estimation and visualisation. Recently, several studies have been conducted to explore how to integrate BIM and GIS to improve construction site safety risk management and optimisation. For example, Irizarry and Karan (2012) integrated the use of BIM and GIS and proposed a GIS-BIM model to assist identification and optimisation of the feasibility for the location of tower cranes. In this work, BIM software was first used to generate geometry information of the construction site, and the GIS model then extracted data from the BIM to determine the proper combination of tower cranes for location optimisation. The analysis output linking to the BIM platform can suggest one or more possible areas including all supply points and demand.

### 2.4.5 Proactive IT-based safety systems

As described in the previous sections, reactive IT-based safety systems are able to provide 4D simulation and virtual prototyping to assist safety risk identification and construction safety management planning. However, as planning is by nature a predictive process established on previous knowledge and experience, the construction projects have a habit of changing during the dynamic processes of a project lifecycle (Forsythe, 2014). To manage those unplanned changes and unexpected safety risks, it is important to track the hazard areas, collect real-time data from the sites for further analysis, and give immediate warning or feedback to the active construction workspace before the actual occurrence of hazards, which is what proactive IT-based safety systems could help (Teizer *et al.*, 2007). To achieve this objective, proactive IT-based safety systems can be created by combining one or more information technologies, BIM, and possibly other techniques. Teizer *et al.* (2007) and Forsythe (2014) summarised the related technologies, approaches, their features, and current situation and development. The core philosophy behind proactive IT-based safety systems is to create a virtual environment where accurate positions of both static and moving objects can be tracked, the corresponding data from the real world can then be collected in real time and analysed by formalised safety algorithms, and, most importantly, information of hazards could be delivered in real-time and effective mitigation measures can be taken in time.

Currently, most efforts of proactive IT-based safety systems focus on tracking the static and moving objects in particular construction activities such as excavator and crane usage. For example, Kim *et al.* (2004) presented a theoretical model of a human-assisted obstacle-avoidance system with a 3D workspace model, and a sparse point cloud approach was described for modelling static objects or zones which may lead to hazards or have been identified to have risks. The framework includes algorithms for an obstacle avoidance system as well as for 3D workspace modelling. To apply this theory, McLaughlin *et al.* (2004) developed an obstacle detection system to allow machines to navigate around equipment safely. Radio frequency wave spectrum technology was applied by Allread (2009) to warn workers in real time where blind spots occur for machine operators and when they are in danger. To improve the safety of blind lifting of mobile/tower cranes, Heng Li *et al.* (2013) presented a real-time monitoring system which integrates the use of Radio Frequency Identification (RFID) and Global Positioning System (GPS). The system can detect the interactive proximity between unauthorised work or the entrance of personnel and the crane. When workers were present within a risk zone, a warning was sent to the safety management team. Other proactive technologies have been used in this area including, laser scanning (Cheng and Teizer, 2014), remote sensing and actuating technology (Teizer *et al.*, 2010), and wireless communication (Wu *et al.*, 2013).

In order to improve the tracking accuracy and reliability, Teizer *et al.* (2013) used Ultra-Wideband (UWB) to deal with the indoor and outdoor settings and to provide the 3D and 4D location values accurately in real time. To enhance the risk management in large transit projects, Ding and Zhou (2013) developed a web-based system for safety early warning in urban metro construction. From this review, it has also been observed that sensors receiving passive warning signals are commonly embedded into Personal Protective Equipment (PPE), such as safety helmets, hats, and shoes, for enhancing the portability of these warning devises, e.g. (Teizer *et al.*, 2010, Abderrahim *et al.*, 2005).

## 2.5 Implications of BIM-based risk management

The purpose of this section is twofold: (1) to provide an overview discussion of BIM-based risk management, and (2) to summarise the shortcomings of related technologies.

The literature shows that BIM and numerous BIM-related digital technologies have been developed to assist risk management during a project's lifecycle. These technologies, discussed in the previous sub-sections, include BIM, automatic rule checking, knowledge based systems, reactive and proactive safety systems.

Applications managing some particular risks can be developed based on either a single technology or a combination of several technologies as illustrated, for instance, in the 4D-GCPSU 2009 system. What can be seen from all of the above efforts is that there has been an emphasis on identifying and mitigating risks as early as possible, and managing real-time risks before any occurrences of hazards. Meanwhile, the findings show that despite considerable developmental work, most of their focus has been on exploiting new technologies to mitigate single risks in particular scenarios for design and construction stages, such as the prevention of falling accidents through automatic rule checking. The management of construction personnel safety risk is a main interest so far, e.g. in Sections 2.4.4 and 2.4.5.

However, there is a need to point out that most existing studies are at a conceptual or prototyping stage because of existing limitations. For example, an important challenge for knowledge based systems is how to ensure the knowledge and experience shared by a limited number of professionals are complete and "correct" information of the potential risks. Although in current AEC projects, successful project risk management is still heavily reliant on all participants' experience and knowledge, as discussed in Section 2.2, different people have different educational backgrounds, knowledge bases, and project experience, and the process of risk management through knowledge sharing is naturally complicated. Eastman *et al.* (2009) highlighted three main problems in current automatic rule checking systems: (1) most common rule checking systems rely on IFC as input and currently are limited in what they support; (2) rule checking at the scale of all sections of a project's codes is a massive undertaking. A critical problem is how to identify and verify the potential errors in the rule checking algorithms and building models; (3) current efforts enable checking the final state of a design but fail to support its development process. Although several reactive IT-based safety systems have been applied for safety risks planning before actual operation, as described in Section 2.4.4, a significant shortcoming exists. The planning process is by nature established on knowledge and experience-based human assumptions. As construction is a dynamic process which may last for many years and involves frequently unexpected changes and unplanned risks, operational risk management cannot normally fully comply with the original planning. Regarding this issue, an additional method is to work on a collaborative 4D construction planning platform by collecting as much reliable multi-discipline knowledge and experience as possible (Zhou *et al.*, 2009). Another alternative approach is to use proactive technologies for real-time data collection and treatment, as described in Section 2.4.5. However, much of the cited work on proactive systems is still very young. Some particular hazardous scenarios in, for example, excavation and lifting have been considered. Meantime, so far most of these efforts only focus on technical development, and these technologies have not reached the stage of "human factor" testing (Forsythe, 2014). Therefore there is still a long way to go before the wide use of these new technologies for risk management will be common in the workplace.

## 2.6 Discussions

An important aspect of this research is to find out challenges and research gaps in current BIM-based risk management through a systematic and critical review, which is discussed as follows:

### 2.6.1 A Multi-disciplinary system-thinking

This review indicates that developing new technologies to assist with the management of construction safety risks is currently a popular research topic. However, any AEC project starts with planning and design followed by the construction stage lasting for months or years, and eventually the project will come into the operation period that may last for decades before demolition. Various types of risks (e.g. structural safety risk, financial risk, environmental risk, supply risk) may be present in the different stages of the project and product lifecycle. People with different knowledge backgrounds and from different domains may be involved in the dynamic process of risk management. ISO (2009) stated that "*risk management is a logic and systematic method*". Hence, it is clear that the concept of multi-disciplinary system-thinking should be embedded in the research of BIM-based risk management.

### 2.6.2 Implementation method and process

The findings show that despite considerable development work, much of the focus has been on exploiting and developing new technologies to treat specific risks in a particular scenario, which were also mentioned by Zhou *et al.* (2012) and Forsythe (2014). Since AEC projects are one-off endeavours with numerous special features and risks existing during the whole dynamic process, any new methods for risk management are valuable when core project participants start to use these enhanced technologies as part of their daily work. The complete implementation framework or method of BIM-based risk management consisting of fragmented activities and processes are equally important as technical developments. Finally, the people, who work collaboratively in a project team using these technologies for managing risks, make the projects successful, and profitable. Based on these observations, an important research topic is to investigate how BIM and BIM-related technologies can be implemented in real projects to achieve their best value.

### 2.6.3 Integration of BIM-based and traditional methods for risk management

Another knowledge gap observed in this review is that there are nearly no studies focusing on integrating BIM and BIM-related digital technologies with the traditional methods, processes, and techniques for risk management. Numerous investigations (Shim *et al.*, 2012, Hartmann *et al.*, 2012, Zhang *et al.*, 2014) have pointed out that the traditional method is heavily reliant on experience and multi-disciplinary knowledge, and common risk assessment techniques include Fault Tree Analysis (FTA) (Suresh *et al.*, 1996), decision trees (Dey, 2002), and neural networks (NN) (Khoshgoftaar and Lanning, 1995), etc. These general methods have been commonly applied by the AEC industry and play a significant role in real projects. Clearly, there is a need to combine BIM-based and traditional risk management to improve practical applicability. The potential and benefits have been proved by several instances. For example, Shim *et al.* (2012) converted the traditional risk management method into visual information in a visualisation environment to improve the efficiency for practitioners in dynamic risk management in terms of schedule, cost and safety to assist the design and construction and management of a challenging cable stayed bridge project. Another study, from a “technology pull” perspective, aligned BIM with risk management into a large infrastructure project to test its practical performance (Hartmann *et al.*, 2012).

### 2.6.4 BIM-based risk management as part of the development process

Undoubtedly risks may be present in the different stages of the project and product lifecycle and the performance of risk management has a direct influence on whether the project can be fulfilled successfully on-time and within budget. In the UK, the CDM rules are a compulsory legislation requirement that indicates all risk analysis for a project starts with the designer. It is the designer who has to assess the risks that may occur during the construction, use of the project, maintenance (including equipment replacement), and demolition. It is the responsibility of the designer to "design out" and eliminate the risks wherever possible. If this is not possible it is the responsibility of the designer to minimise the risks. When a contractor is appointed, the analysis of risks continues but now with the assistance of specialists in construction. A construction project is normally divided into a number of sub-projects for managing risks at a sub-project level by considering different activities and processes individually. Each sub-project may have separate designers and contractors with their own risks to identify and manage. A group of risk specialists (experts from multi-disciplines) hired by the project team then need to collaborate with project members to identify and investigate the potential risks by interviews and discussions. A group of paper-based risk documents (e.g. risk start-up report, risk inventory) are then compiled in this process. To implement risk management, specialists who play facilitating roles during the risk management process need to attend the project control meetings and keep tracking progress, and give advice on specific construction activities. However, the project team, especially the managers, is required to be responsible for the application of the risk management cycle. It is extremely important to point out that many people will be involved in the risk management during the lifecycle, so that any updated risk information, decisions and changes should be recorded and communicated effectively. Therefore, BIM-based risk management is expected to facilitate efficient risk communication and support the dynamic development process of a project.

# Chapter 3. RESEARCH METHODOLOGY

## 3.1 Introduction

Connecting to the state-of-the-art review of literature in Chapter 2, this chapter will introduce and discuss the methodology and choice of methods used in this research. The chapter starts with an overview of the commonly used research approaches in the field of Construction ICT. It then justifies the mixed use of research methods for achieving the stated research objectives. The Chapter 4 will introduce the basic theories and approaches as starting points for the development of the thesis.

## 3.2 Types of research

Research is defined as a process of enquiry and examination undertaken on a systematic basis using disciplined methods to discover unknown relationships, create further knowledge and use it for devising new applications (Easterby-Smith *et al.*, 1991, OECD, 2002). Research approaches may vary within and between science, technology and humanity in different ways depending on epistemologies. A research approach defines the means or modes of data collection, analysis, and how specific results are to be calculated and concluded (Howell, 2012).

Research in the field of digital technologies in the construction domain is not a purely technical question, as with the study of engineering or ICT: it often involves many aspects, e.g. engineering, technology, management and social science (Whyte, 2000). Research in this area is expected to explore the fundamental theories that can explain obstacles and gaps, and use the approaches of ICT as a starting point to develop solutions to improve digitalisation, automation, collaboration and productivity in the context of digital technologies for construction projects. Proper research design is a precondition for seeking answers to research questions and validating the results, where an important step is to justify the rationale of the selection of research methods.

Although there are different ways to categorise research methods and many approaches have evolved from a historical view, especially considering that traditionally research followed the natural science approaches (Kimmance, 2002), today some scientists tend to classify these research methods into three basic types: 1) qualitative, 2) quantitative, and 3) mixed methods (Creswell, 2013). A major difference between qualitative and quantitative research shows that the former is often considered as non-numerical examination using open ended questions to discover underlying meanings and patterns of relationships, while the latter refers to "numerical representation and manipulation of observations for describing and explaining the phenomena that those observations reflect" (Babbie, 2015). However, the qualitative and quantitative research should not be considered as polar opposites and it is more accurate to say a research tends to be more qualitative than quantitative or vice versa (Creswell, 2013, Newman and Benz, 1998). Mixed methods research occupies in the middle of qualitative and quantitative research because it combines the use of elements from both qualitative and quantitative approaches (Creswell, 2013).

### 3.2.1 Qualitative research approach

Qualitative research was originally developed in the social science field to explore and understand the social phenomena or human problems (Coombes, 2001). It is conceived as a wide methodological method that consists of many research methods (Alasuutari, 2010). The process of this research involves emerging questions and procedures, collecting data in a relatively subjective way, analysing data inductively, and making interpretations of the meaning of the data to get the general rules and themes (Creswell, 2013). In other words, qualitative research is often employed to carefully collect and examine a rich collection of data under a well-designed guideline to better understand and explain a phenomenon.

Actually, qualitative research refers to a relatively wide methodology encompassing many research methods, and has now been used and developed in both social sciences and the fields of natural science. Contemporary qualitative research has been characterised by a distinct turn toward more interpretive, postmodern, and critical practices, and it was identified that five main types of paradigms are used: positivism, post-positivism, critical theories, constructivism, and participatory/cooperative paradigms (Lincoln *et al.*, 2011). There are various qualitative research approaches associated with the qualitative paradigm including, for example, grounded theory, phenomenology, epistemology, critical theory, case study, action research, participant observation, visual analysis, discourse analysis, etc.

Many studies that provide in-depth discussions on the qualitative philosophy perspectives and methods are available, e.g. (Lincoln *et al.*, 2011, Creswell, 2013, Myers and Avison, 2002, Baskerville and Wood-Harper, 1996).

### 3.2.2 Quantitative research approach

Quantitative research is often contrasted with qualitative research and, as its name suggests, refers to the process of systematically investigating observable phenomena through collecting, processing and analysing "numerable" data using appropriate statistical, mathematical or computational methods (Given, 2008). The purpose of quantitative research is to test objective theories or discover underlying relationships through examination and analysis of variables (Creswell, 2013). In a way, quantitative research requires that these variables can be counted or measured so that the numerical data can be analysed and interpreted using statistical procedures. Quantitative research was initially developed in the natural science field to study natural phenomena and is believed to be the oldest type of research (Kimmance, 2002). The core belief within this type of research coheres with immanent principles of causation and stresses objectivity, measurability and repeatability of variables, where complex problems can be addressed by reductionism (Locke *et al.*, 2009). Therefore, quantitative research expects researchers to keep away from the research process and use objective data and unbiased result to describe the generality of the existing reality.

Locke *et al.* (2009) summarised that the main branches of quantitative research include: 1) descriptive, 2) correlational/predictive, 3) quasi-experimental, 4) single-subjects, and 5) meta-analysis. According to Kimmance (2002), the quantitative approaches can be further broken down into lower levels including, for example, survey, laboratory experiments, structured observations, statistics, and mathematical modelling. More comprehensive discussions on quantitative research can be found from, e.g. (Locke *et al.*, 2009, Newman and Benz, 1998, Creswell, 2013).

### 3.2.3 Mixed methods research

Mixed methods (also known as integrating, synthesis, hybrid or multi-method) research is an approach to investigation that involves collection of both qualitative and quantitative data, integrating the data from a variety of sources, and following well-designed theoretical frameworks to examine complex phenomena (Creswell, 2013). The term 'mixed' implies the use of both qualitative and quantitative data and methods, and thus mixed methods research is distinct from simply a combination of multi-qualitative or quantitative methods. The core assumption or purpose of this type of research is that the combined use of both qualitative and quantitative elements can produce a more accurate and complete understanding towards a complex phenomenon or problem. An advantage of mixed methods research is the counteraction of the inherent threats to validity, generality and reliability and overcome the intrinsic biases or weakness of the observed problem (Kimmance, 2002, Gable, 1994). According to Creswell (2013), mixed methods research design has three basic forms (i.e. convergent parallel mixed methods, explanatory sequential mixed methods, and exploratory sequential mixed methods) and three advanced types (i.e. embedded mixed methods, transformative mixed methods, and multiphase mixed methods).

## 3.3 Methodology adopted

Research of Construction ICT lies in an inter-disciplinary area that covers many aspects, e.g. engineering, technology, management, and social science. One important reason that can explain this phenomenon is that, in a sense, every construction project is unique and construction is by nature a complex manual process that involves human efforts, engineering knowledge, experience-based decision-making, use and operation of instruments, etc. Research in this area is not the same as pure engineering or management studies that deal with a single-aspect problem: data may be collected in both qualitative and quantitative ways from a variety of sources. It is observed that many researchers tend to use a mixed methods methodology to guide their research to gain a better or complete understanding, especially in the research area of applying ICT for construction management. For example, Kimmance (2002) employed a mixed methods strategy to develop a customised research framework called HIPPY model to guide the doctoral research of developing an integrated product and process information modelling system for onsite construction.

The main idea and overall methodology to address the aims and objectives of this research was summarised in Chapter 1. In order to overcome the knowledge gap of managing risk knowledge and information within the BIM environment, it has been discussed that existing risk management techniques can be integrated into BIM to establish a BIM-based risk management system. The research is closely related to two main aspects: managing risk data form knowledge-based perspectives and integrating risk data into BIM through using ICT as an enabling technology. A mixed methods approach that can synthesise both aspects and combines both qualitative and quantitative approaches was used to investigate solutions for the observed issues.

The "CIFE Horseshoe Method" (Kunz and Fischer, 2007) was used to guide the overall direction of this research (Figure 5). Initially developed by Center for Integrated Facility Engineering (CIFE) of Stanford University, the "Horseshoe" method defines a structured framework to plan and manage theoretical research in the construction industry. In this framework, intuition refers to the big idea that may explain the nature of the problems being investigated and Point of Departure (POD) linked to intuition describes what is already known about the problem and what basic theories or approaches can be used as a starting point to support the research development. The validation makes research results reliable and can prove the research findings answer the research questions. Established on the validation, contributions towards the theories or approaches discussed in the POD can be claimed.

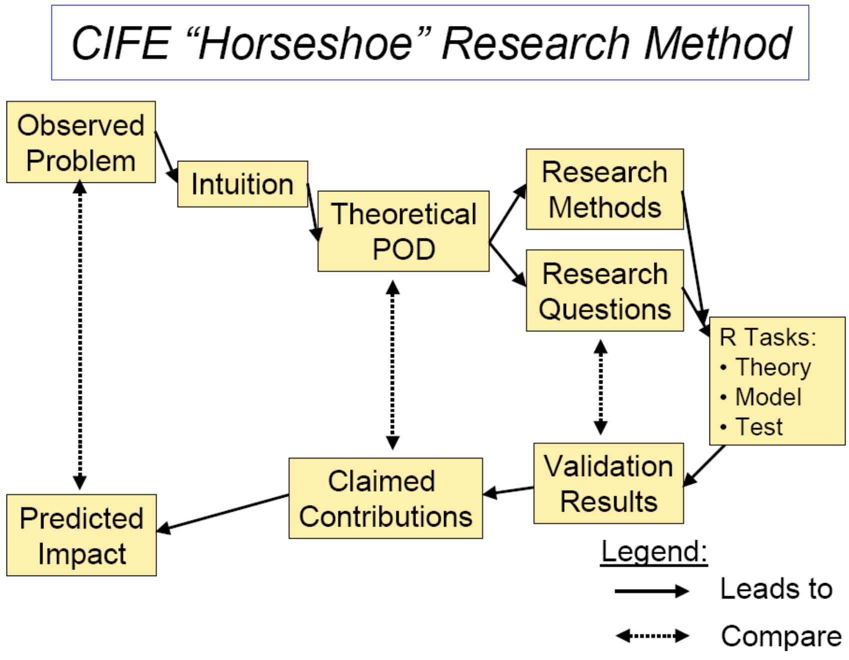



Figure 5. CIFE Horseshoe Method (Kunz and Fischer, 2007)

In Chapter 1, it is observed that currently very few theories exist that can explain how to align traditional risk management techniques with BIM and the current BIM solutions have very limited support on risk communication and information management. The intuition to the observed problem is to integrate risk management techniques into BIM-based platforms and establish an active linkage between the risk information and BIM. It is discussed in Section 1.4 that some previous research provided the evidence that RBS and CBR can be used for the development of this research and an overall concept framework is proposed. On one hand, the RBS is a hierarchical representation of different types of risk and can be used for managing the risk information in a pre-defined database. On the other hand, as some risk knowledge is stored in reports that are written in human language, CBR can facilitate the use of previous risk knowledge in analysis of new situations. The theoretical POD of the proposed solution in this research is addressed in Chapter 4. For the use of CBR, the focus of this research lies on the method of quickly obtaining the most relevant information from the risk report database. One important reason is that current BIM tools (e.g. Autodesk Navisworks 2017) already support linking a document (e.g. a risk report) to BIM objects while the key difficulty in the construction workplace is how to find the most valuable information within a limited time.

The research methods used in this research include: literature review, interview with industry experts, concept modelling, prototype development and evaluation. There is a need to clarify that the selection of research methods is to achieve the following three objectives outlined in Chapter 1:

- To develop a RBS and a CBRL for bridge projects.
- To develop a linkage between Risks and BIM for bridge projects.
- To develop a method to support the fast risk case retrieval from the CBRL during the project development process.

The research tasks, methods, and validation are summarised in a diagram shown in Figure 6.

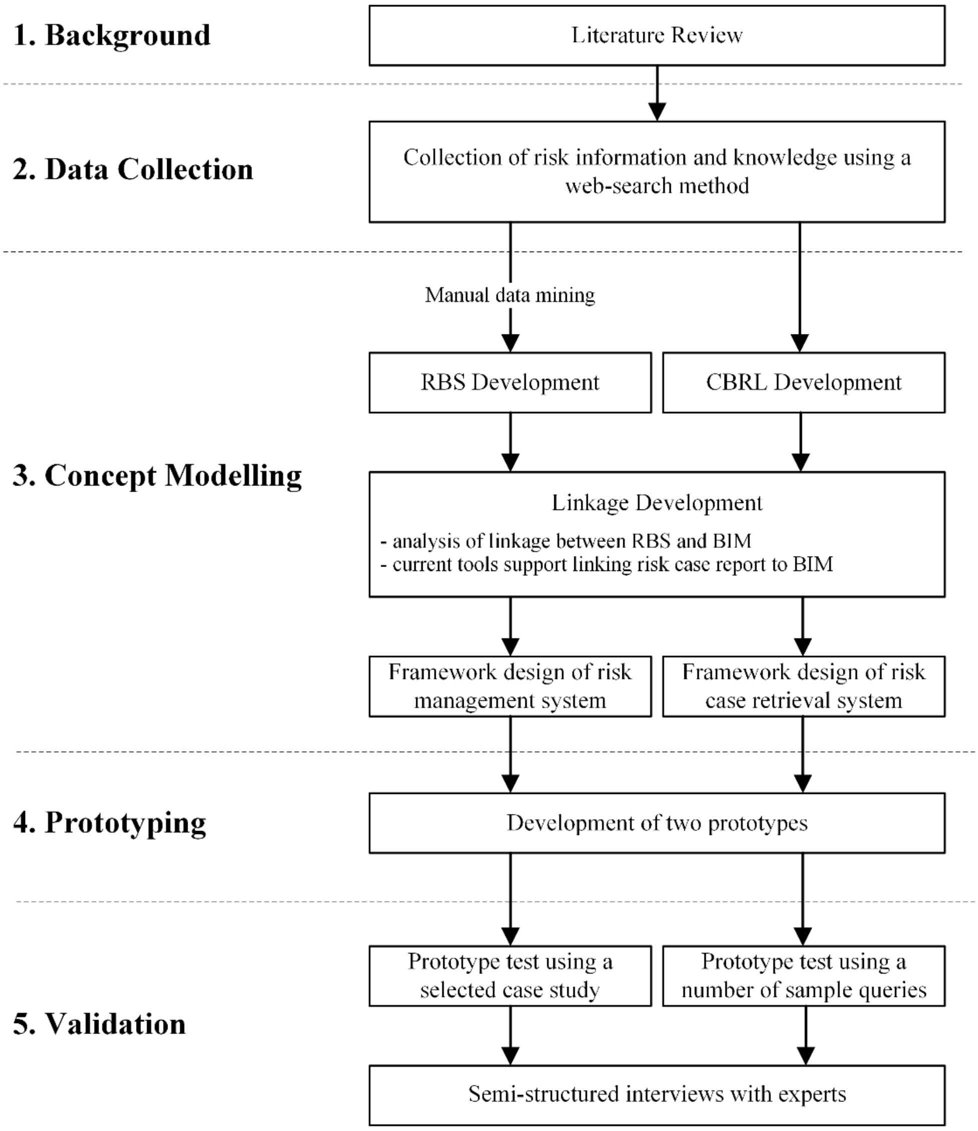



Figure 6. Diagram to show the research tasks, methods and validation

### 3.3.1 Literature review

In this research, an extensive review of literature was conducted in two stages.

The first stage of the literature review documented in Chapter 2 was to obtain an understanding of the overall picture of the research area, and identify the knowledge gaps that currently exist in BIM-based risk management. The topic of "risks of implementing BIM" and papers that are not published in English were not within the scope of this review. Specifically, a three-step approach was used. In the first step, the fundamentals, general process, and main challenges of traditional risk management were summarised through an extensive literature review and several expert interviews for comprehensive understanding of the relation between the traditional methods and BIM-based risk management. The process identified a set of keywords for data collection as the basis for the next step. The main keywords were “BIM”, “building information model”, “risk”, “risk assessment”, “risk analysis”, ”risk management”, “knowledge management”, “safety”, “quality”, “time”, “cost”, and “budget”. In the second step, these keywords were applied to a web search in online academic publication databases, i.e. “Web of Science”, “Engineering Village”, “Scopus”, and “Google Scholar”, for collecting academic and applied publications related to this topic. Then the state-of-the-art of these technologies were classified and surveyed as follows: (1) BIM, (2) automatic rule checking, (3) knowledge based systems, (4) reactive IT-based safety systems (i.e. database technology, VR, 4D CAD, GIS), and (5) proactive IT-based safety systems (e.g. GPS, RFID, laser scanning). The scope of the survey included articles in leading journals of this area (e.g. *Safety Science*, *Automation in Construction*, *International Journal of Project Management*, *Journal of Computing in Civil Engineering*, *Information Technology in Construction*, *Reliability Engineering & System Safety*), publications from conference proceedings and other sources of professional associations, standard committees (e.g. HSE, ISO) and authorities. In the third step, all publications were analysed critically and compared with the traditional risk management methods to identify current obstacles and future work to close these gaps.

From the initial literature review, it was identified that future research of BIM-based risk management should (1) have a multidisciplinary system-thinking, (2) investigate implementation methods and processes, (3) integrate traditional risk management with new technologies, and (4) support the project development process. As discussed in Section 1.1.3, two main knowledge gaps to be addressed in this research were then concluded after an extensive survey of relevant literature as: (1) very few theories exist that can explain how to align traditional techniques with BIM for construction project risk management; (2) current BIM solutions have very limited support on risk communication and information management during the project development process.

Built on the well-defined problem and clearly articulated intuition, the second stage of the literature review discussed in Chapter 4 was for searching and discussing the theoretical starting points for the research.

### 3.3.2 Development of RBS through a manual data mining method

In order to achieve the second objective of developing a linkage between the tailored RBS and BIM for bridge projects, the first step was to understand and identify the basic hierarchical structure of content of an integrated bridge information model. As only limited studies were found, a comprehensive analysis based on the existing primary element hierarchy of IFC models for buildings, existing studies, e.g. (Kiviniemi, 2005, Shim *et al.*, 2012) and the author's project experience on bridge design and construction was conducted to divide an integrated bridge information model conceptually into different LOCs and technical systems. This division was the theoretical basis in further steps for linking different groups of risks to the particular levels of a bridge information model.

The second step employed a manual data mining approach (Jun Lee and Siau, 2001, Gargano and Raggad, 1999) to collect, identify and categorise risk information. It started with an extensive web-search to collect academic publications, bridge project risk assessment reports, and standards and guidelines that documented risk information in the past or potential risks that may affect bridge projects. As construction projects share a large number of common risks and there are only a limited number of documents focusing on bridge related risks, the scope of collecting academic publications and related standards or guidelines was extended to all construction projects. A manual text mining process was then conducted through careful study of each document and interpreting and understanding the text in its relevant context to identify the valuable risk information (e.g. risk category, risk factor, risk description, and possible mitigation measures or strategies) in 80 collected documents. As there is currently no consensus on how to develop the RBS (Mehdizadeh *et al.*, 2013), a list of key words (e.g. project risk, external risk, global risk, design risk) were identified from previous studies (Tah and Carr, 2001, Choi and Mahadevan, 2008, Mehdizadeh *et al.*, 2013) as an initial hierarchy for allocating and managing the collated risk information according to the source of risk. All identified risk information was stored in an initial database, which was defined as the ‘risk pool’ in this research. After this, similar risks were translated to one format and all risk information was well structured to develop a knowledge-based risk database.

### 3.3.3 Concept development of linkage

Built on the results obtained in the second step, the next step further categorised risks to generate a tailored RBS. The location of different types of risk in the RBS were classified according to their relationships with the 4 LOCs, e.g. structure-related risks are related to bridge-level while the financial risks are related to the project-level. To further improve the practical applicability of implementing the linked relationship between RBS and BIM, a critical analysis was then conducted to determine on which level the different risks should be allocated to bridge projects and 13 sub-models of linkage were developed. Finally, risks at the lowest level of the generated RBS were classified into four groups (i.e. project, surrounding environment, site, and bridge) and a conceptual model was established to link four LOCs and six technical systems of BIM to the tailored RBS.

### 3.3.4 Development of CBRL using a web search method

To collect risk reports for establishing the CBRL as part of first research objective, a web search method was used. In total 590 risk cases were collected from the following major organisational and governmental construction accident databases: (1) Structural Safety (Structural-Safety, 2016), (2) the National Institute for Occupational Safety and Health (NIOSH, 2016), (3) WorkSafeBC (WorkSafeBC, 2016), (4) Occupational Safety and Health Administration (OSHA, 2016), and (5) others (e.g. some published papers that document construction accidents). The source distribution of collected risk cases is shown in Figure 7 and the category distribution is presented in Figure 8. Although collecting as many risk cases as possible from every category of project risks could improve the reliability of the proposed approach, this study stopped collecting more cases due to the following reasons: (1) the focus of this study was on developing a NLP based general approach for risk case retrieval instead of establishing a complete risk case database; (2) it was observed that some risks (e.g. collapse of structure, loss of life) that may lead to severe consequences attract more attention while there are very few detailed reports available on those risks that are not so dangerous, e.g. financial loss, time overrun.

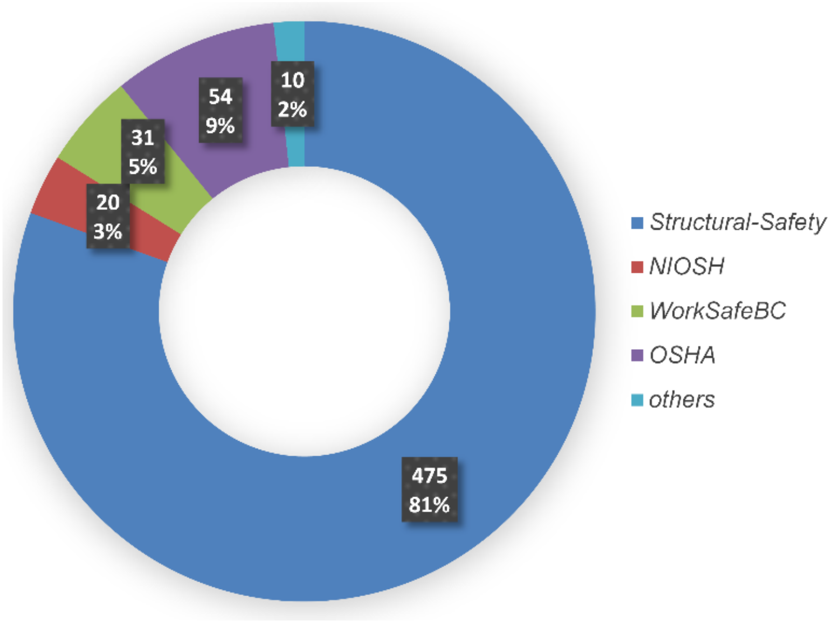



Figure 7. Source distribution of collected risk cases

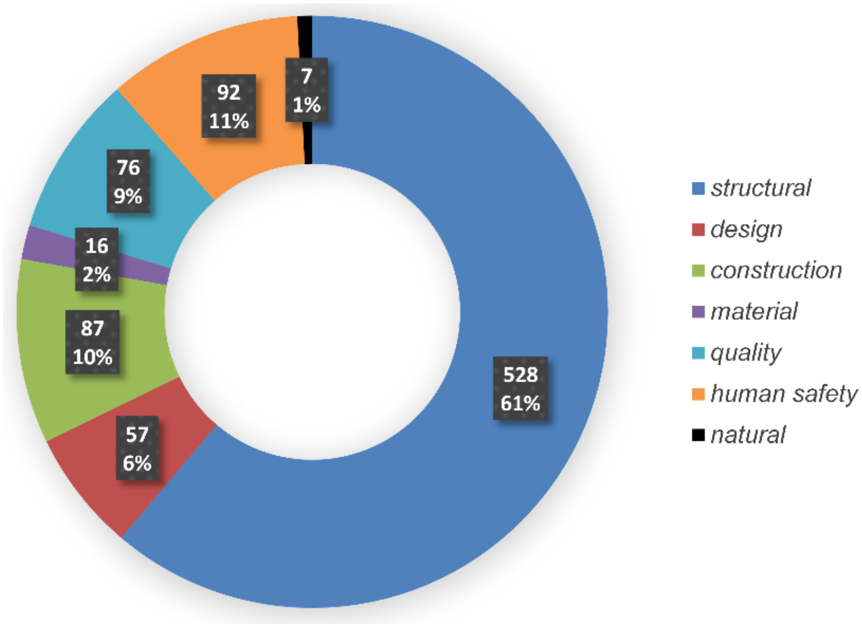



Figure 8. Category distribution of collected risk cases

### 3.3.5 Framework development for risk case retrieval

In order to improve the efficiency and performance of risk case retrieval from the CBRL established in Section 3.3.4, this research developed an approach of combining the use of VSM and semantic query expansion, and outlined a framework for this Risk Case Retrieval System. It was an important step to achieve the third research objective and the detailed development process is described in Chapter 7.

### 3.3.6 Prototype development and evaluation

As part of the second and third research objectives, the method of prototype software development was used for testing the implementability of the proposed solution, which is a widely adopted method for testing and validating concept in the area of Construction ICT and provides the feasibility of further developing the concept for commercial use (Smith, 1991).

Specifically, two prototypes were developed. The first one was a plugin into the Navisworks environment to support the linkage between RBS and BIM. The strategy for demonstration and evaluation of the Navisworks plugin was through a selected case study of a standard steel footbridge. The second prototype was a Python program that is capable of retrieving similar risk cases from the CBRL according to users' queries. The evaluation strategy was to test 10 selected queries and observe that their retrieval results were at an acceptable level.

Details of the development and evaluation process of the two prototypes are discussed in Chapters 6 and 7.

### 3.3.7 Expert interviews

A series of important discussions with industry experts were involved during the whole process to guide the development of the research. The valuable suggestions and comments from industry experts played a complementary role to shape the research to contribute to both the existing body of academic knowledge and addressing practical challenges.

A number of semi-structured interviews with leading experts, from design consultancies, contractors and software companies, were conducted to validate the usefulness of the proposed solution of this research. The interviews were designed as semi-structured to allow the interviewees to explore and gain an overall understanding through presentations and free discussions, and then develop answers to the well-designed questions. The background of the invited industry experts is summarised in Table 2. The detailed interview process is described in Section 8.1 and the experts' reports are listed in Appendix C.

Table 2. Background of interviewed experts

| No. | Name | Expertise | Experience | Method | Duration |
|---|---|---|---|---|---|
| 1 | Alastair Soane | Structural Safety | 50 years | Face-to-face | 1.5 hours |
| 2 | Gordon Crick | Construction Safety | 25 years | Face-to-face | 1.5 hours |
| 3 | Benedict Wallbank | BIM and Architecture | 35 years | Face-to-face | 1 hour |
| 4 | David Philp | BIM | 23 years | Video conference | 1 hour |
| 5 | Martin Simpson | BIM and Stadium Design | 20 years | Face-to-face | 1.5 hours |

# Chapter 4. POINT OF DEPARTURE

As stated in Sections 1.1.3 and 1.2, very few existing theories can explain how traditional risk management techniques can be integrated into BIM and how risk information can be communicated and managed in the BIM environment during the project development process. To overcome these knowledge gaps, a solution presented in Sections 1.3 and 1.4 can build on the following three main starting points:

- BIM as a systematic solution
- Hierarchy of project risks
- Management and use of previous knowledge and experience

Firstly, BIM as a systematic solution illustrates: 1) the potential of using the visualisation feature of BIM (i.e. 3D/4D BIM) to assist early risk identification and analysis, 2) why and how, in concept, the BIM model could be divided into different Level of Contents (LOCs) and technical systems for risk management purpose, and 3) the feasibility and reasons of establishing a link between risks and BIM for lifecycle risk information sharing and management. Secondly, a hierarchy of project risks describes what traditional techniques are available for classifying and managing different types of risk, and why the RBS is chosen as a core traditional technique that can be linked to BIM. Thirdly, the management of previous knowledge and experience demonstrates: 1) why CBR can facilitate the use of previous knowledge and experience for solving new problems, 2) what barriers exist in implementing CBR, and 3) why and how NLP can be integrated with CBR to support risk management and decision making during the project development process.

## 4.1 BIM as a systematic solution

### 4.1.1 The application of 3D/4D BIM

Three-dimensional/four-dimensional (3D/4D) models are two important concepts of BIM and have been increasingly used in construction projects (Hartmann *et al.*, 2008), where the fourth dimension refers to the time- and schedule-related information. 3D BIM can be created by a number of existing tools, e.g. Graphisoft ArchiCAD (Graphisoft, 2016), Autodesk Revit (Autodesk, 2016b), Tekla Structures (Tekla, 2016). Other tools allow the time schedule to be linked to the 3D model to establish a 4D BIM, e.g. Autodesk Navisworks (Autodesk, 2016a), Synchro PRO (Synchro, 2016).

One important motivation for choosing BIM as a core technology of the proposed BKRMS is that 3D/4D BIM could effectively facilitate the concept of "early risk identification and prevention" through the visualisation of a project and the simulation of construction activities during the entire construction process in a computer based virtual environment before the onsite implementation of construction. The 3D/4D BIM is also a way to help recall previous knowledge and experience to solve potential problems. For example, BIM itself has been proven as a systematic way to assist early identification and assessment of risks for design and construction through 3D visualisation (Grilo and Jardim-Goncalves, 2010), 4D scheduling (Zhang and Hu, 2011), and 5D cost estimating (Mitchell, 2012). The spatial visualisation and dynamic modelling of a project in a computer system could effectively facilitate early risk identification and communication (Liu *et al.*, 2014), and assist strategy and decision making to improve safety, time and cost management in construction (Hardin, 2011). Meanwhile, open data standards such as IFC that store standard and customised data for all project elements provide an interoperable digital representation of all project elements enabling interoperability between BIM software and applications (Laakso and Kiviniemi, 2012), which can increase the repeated use of data and reduce the possibility of errors. With the growing development of BIM in the AEC industry, some efforts that could further integrate BIM with risk management have been observed, e.g. automatic rule checking (Eastman *et al.*, 2009, Zhang *et al.*, 2013, Sulankivi *et al.*, 2013), proactive IT-based safety systems (Forsythe, 2014), and safety training in a virtual gaming environment (Guo *et al.*, 2012).

### 4.1.2 Separation of BIM into LOCs and Systems

In concept, data from different contents and disciplines is gradually defined and added to build the integrated bridge information model and BIM can be considered as a shared knowledge and information repository to support the whole project lifecycle. The data and their structure in a complete BIM are extremely complex, therefore some researchers, e.g. Fischer and Kam (2002) and Haymaker *et al.* (2003), realised that there is a need for division of the integrated information model to meet particular needs. In 2005, Kiviniemi (2005) proposed a formal solution for dividing a project’s data set into several sub-models and linking user requirements with these sub-models. Similarly, as current open BIM standards such as IFC are only supported in a limited way on infrastructure structures such as bridges, Shim *et al.* (2012) divided the integrated bridge information model into five levels for different purposes of use, e.g. structural analysis, structural detailing, and construction simulation.

LOC is defined in this research as the primary hierarchical structure of content of BIM. LOC could be used for decomposing an integrated 3D information model into separated sub-models according to different information content and different disciplines, which could enable the ‘correct’ information to be extracted, used and communicated in an efficient way to meet particular requirements, e.g. structural analysis, construction scheduling simulation, and risk management. Currently there is no explicit method for dividing the model into LOCs but the division should meet particular needs and requirements. Kiviniemi (2005) defined the technical system as an aggregation of objects that have a common purpose or function or to provide a service, which originates from the definition of ifcSystem by buildingSMART (2016a). Although some researchers (Tah *et al.*, 1999, Shim *et al.*, 2012) tried to summarise the component objects for a bridge information model, no study has been found to classify and group bridge component objects that have a common purpose or function or to provide a service to be a part of a technical system.

In this research, the integrated bridge information model is, in concept, divided into 4 LOCs and 6 technical systems based on analysis of the Industry Foundation Classes (IFC) specification, a critical review of previous studies and the author’s project experience. One motivation for the division in this research is that risks can affect and have impacts on BIM differently. For example, some risks (e.g. financial and legal) may affect the overall project while other risks (e.g. structural, site safety) may only influence the structure or construction site. Therefore, a potential linkage between risks and the particular LOCs or systems of BIM could help facilitate the understanding of how the particular risk may affect BIM as well as improve the risk identification efficiency.

In addition, a model can be generated through one or more of the various BIM-based software, e.g. Graphisoft ArchiCAD (Graphisoft, 2016), Autodesk Revit (Autodesk, 2016b), Tekla Structures (Tekla, 2016). Although the models created by the software can serve as a basis for the linkage between risks and BIM, these software have been developed by using different methods and following different standards and this means that the hierarchies of these models are different. Moreover, the internal structures of models created by these software might not be publicly available. As a result, the linkage rules between risks and the conceptual LOCs and systems of an integrated bridge information model can have more generality and applicability in understanding the relationships between risks and BIM.

### 4.1.3 Lifecycle risk information management

As discussed in Section 1.1.2, three challenges still exist in traditional risk management and can be expanded as follows.

- Traditional risk management is still a knowledge and experience based manual undertaking, and numerous investigations (Shim *et al.*, 2012, Hartmann *et al.*, 2012, Zhang *et al.*, 2014) have concluded it is time-consuming, error-prone and highly inefficient. In real projects many practitioners still work on two dimensional (2D) platforms and use 2D drawings and paper-based documents to convey the product information. In this process, although some simple techniques, such as checklists, could assist risk identification and analysis (HSE, 2015), it is a significant challenge to combine and link 2D drawings, on-site observations and paper-based documents together for identification and consideration of risks. Decisions are to a large extent made through a "brainstorming" exercise based on existing knowledge and previous experience.

- Risk knowledge management is fragmented and insufficient, and risk knowledge transfer from project to project is difficult. Multi-disciplinary knowledge and experience play a key role in traditional risk management and the corresponding decision making. Project participants, e.g. clients, architects and engineers, gain valuable knowledge and experience from every project and can use them to contribute to future work. In this case, the effective management of this large database of human knowledge and experience as well as flexible and accurate data extraction become a precondition for the success of risk management. However, unlike some manufactured products that can be made automatically, every AEC project has its unique characteristics that are distinct from others (Clough *et al.*, 2000). In addition, the process of any AEC project is dynamic and new experience and new lessons come to light nearly every day. Consequently, another significant challenge is how to effectively manage the "database" of human knowledge and experience as well as extract the correct data flexibly and accurately.

- Communication and collaboration needs to be improved in traditional risk management. Since projects are completed by a team cooperatively, any common risks will be identified and treated individually, and the corresponding information will be documented and sometimes this work will be ignored or forgotten (Kazi, 2005). This may lead to the risk that information cannot be presented, shared, recorded, and updated effectively during the development process of a project. As the project is handed over from designer to contractor, and then from contractor to the client, people will normally leave the project after completing their tasks. Thus, large amounts of risk information may be lost if it is not recorded properly and communicated to other project participants.

It can be summarised that in current risk management it is still a huge problem to manage, share and communicate the risk information to support the project development process, where some existing studies show BIM has the potential to overcome this challenge. For example, the National BIM Standard (NBIMS) of the United States (NIBS, 2007) defines BIM as: "*BIM is a digital representation of physical and functional characteristics of a facility. A BIM is a shared knowledge resource for information about a facility forming a reliable basis for decisions during its lifecycle; defined as existing from earliest conception to demolition*". It emphasises that BIM can be considered as a data enriched model of building projects and its information is generated gradually and managed effectively to support the project development process. For example, design models created by BIM-based design tools can support the construction planning for contractors as well as decision making for the clients. Open data standards such as the IFC that store standard and customised data for all project elements could provide an interoperable digital representation of all project elements enabling interoperability between BIM-based applications (Laakso and Kiviniemi, 2012). In addition, Deshpande *et al.* (2014) proposed a method to capture, extract, and store information and knowledge from BIMs, and presented a framework for classification and dissemination of the knowledge. Therefore, BIM is an enabling technology for lifecycle information management and has the potential to facilitate risk communication and information management.

A construction management tool called Vico Control (Vico Software, 2017) developed a risk analysis module which enables to extract schedule and resource information from BIM for Monte Carlo simulation. However, it focuses on a narrow view of risk management on producing schedules and does not cover the relations between different types of risks and the LOCs and systems of BIM, and thus does not enable wider management of risk information with BIM.

To facilitate the risk information management, a core approach used in this research is a linkage between risks and BIM, which will be described in detail in Chapter 5. The following two reasons could explain why this study decided to establish a linkage instead of embedding risk information into BIM. First, the scope of risk is very broad including, for example, finance, environment, health and safety. Unlike information relating to geometry or materials that can be described easily and stored in a standard format, most risk information is written and stored in everyday language and people from different disciplines consider and record risks in different ways, which leads to the difficulty in embedding risk data into BIM. Secondly, nowadays different disciplines are still using different platforms or tools to contribute to a construction project. Storing the risk database in the risk management system and linking the related risk data to BIM could reduce the risk of information loss arising from updates or changes to the BIM and information transfer between different platforms.

## 4.2 Hierarchy of project risks

According to ISO 31010:2009, risk management is a logic and systematic method that involves a set of activities and processes for establishing the context, facilitating risk communication, identifying, analysing, evaluating, treating risks, and recording and reporting the corresponding results properly in a timely manner (ISO, 2009). AEC projects start with planning and design, followed by a construction stage that may last for many months, and eventually will come into the operation stage that may last for many decades before demolition. Different risks are present in the different stages of the project and product lifecycle. This means that regardless of the activity, there is always a possibility that hazards will occur and the whole project may be affected depending on the type of risk and how severe the consequences are. The scope of a risk consists of many issues: damage or failure of structures, injury or loss of life, budget overruns, and delays to the construction schedule, etc. Consequently, all project participants need to improve their ability, knowledge and experience to manage risks during the project lifecycle to ensure a safe, successful, and sustainable project. Generally a complete risk management process contains a sequence of sub-processes including risk identification, risk analysis, risk evaluation, risk treatment and risk review (PMI, 2004, ISO, 2009), where the first and the most important step is to identify potential risks associated with project tasks (Zou *et al.*, 2007). Failure in identifying risks at an early stage may lead to unawareness and no treatment of serious risks.

To learn from and use past project knowledge and experience for managing risks, an effective way is to work out a comprehensive risk database containing all possible risks that may affect the project. The database could facilitate a systematic understanding of all project risks, and help the project team link risk information to real projects and make decisions quickly, e.g. (Kartam and Kartam, 2001, Wang and Chou, 2003). As construction is by nature a dynamic process with unexpected changes and risks and new information is added into the project every day, it is crucial to use a logical and rapid approach for classifying and structuring the large amount of information. Currently a variety of tools have been developed for risk classification, e.g. risk list (PMI, 2004), risk matrix (Markowski and Mannan, 2008), risk maps (Dey, 2010) and RBS (Holzmann and Spiegler, 2011).

In concept, RBS is a hierarchical structure that allows all types of risk factors and events to be well organised by groups and categories (Holzmann and Spiegler, 2011). It is an open, flexible and easily updatable tool and could offer a global view on risk exposure (Tah and Carr, 2001, Mehdizadeh *et al.*, 2013). The main advantages of RBS include: 1) to increase overall understanding of risks and facilitate risk communication; 2) to help locate identified risks into relevant risk categories and make special strategies to treat them easily; and 3) to provide an architecture for managing risk database and developing risk management software. An example of a RBS is presented in Figure 9.

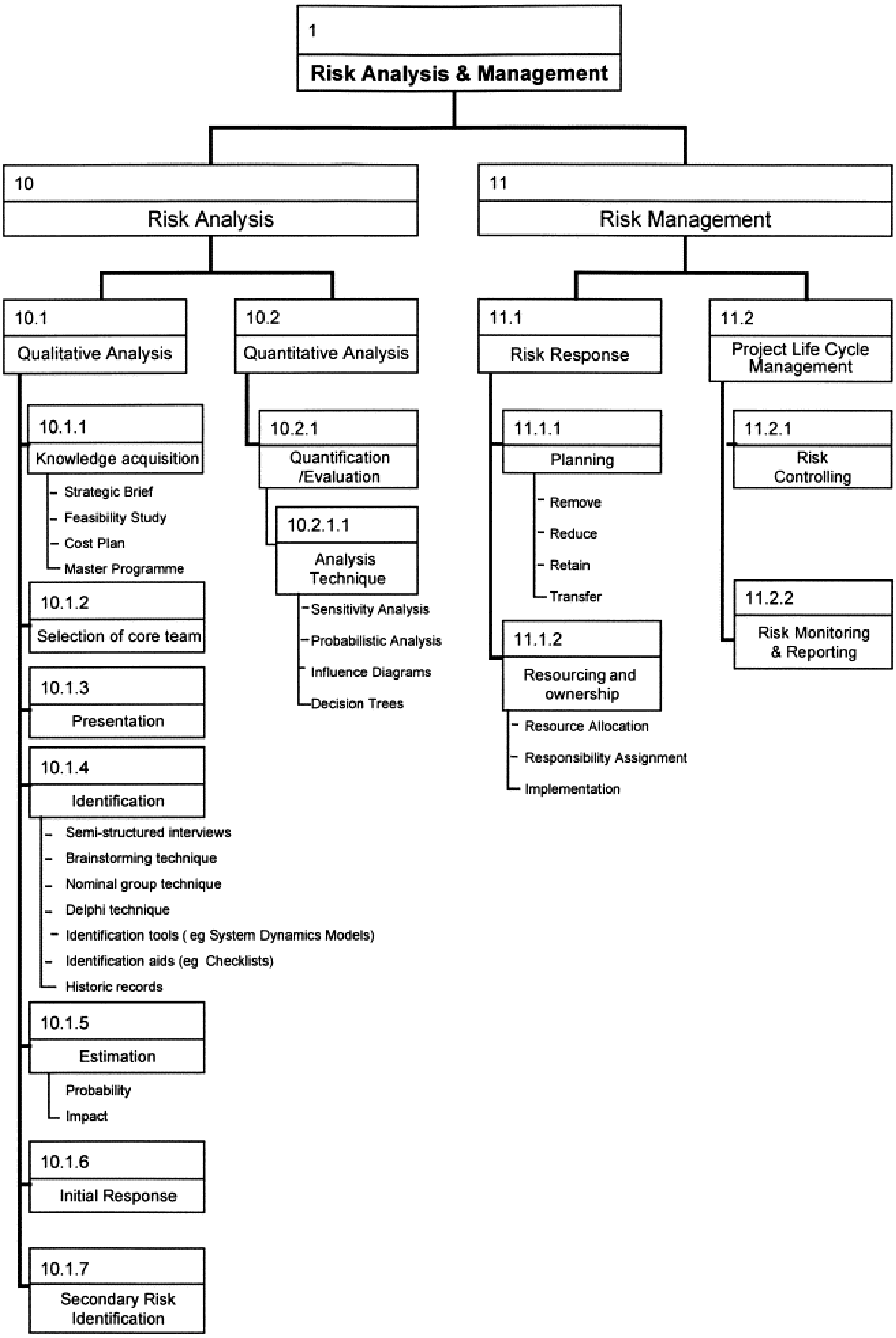



Figure 9. Example of a RBS (Chapman, 2001)

The objective of this study is to develop a methodology to merge the traditional technique and BIM as an integrated solution for risk management. RBS was chosen as a core traditional technique for the following observations.

- A recent study conducted by Shim *et al.* (2012) presented a conceptual diagram for visualising risk information in BIM and pointed out that RBS has the ability to facilitate the understanding and communication of risks in risk identification and analysis processes. The integrated use of RBS and BIM can take advantage of both the traditional method and BIM for managing risks. On the one hand, RBS enables risk information to be stored in a formal structure, used and communicated effectively. On the other hand, some features of BIM such as 3D visualisation and 4D construction scheduling can facilitate the risk identification, analysis and communication at an early project stage.

- The RBS of a project can be further associated with project tasks and developed to be a Risk Breakdown Matrix (RBM) for risk management. The RBM is in definition a cross matrix that interconnects the Work Breakdown Structure (WBS) of a project with its Risk Breakdown Structure (RBS), which is a simple but useful technique enabling the consideration of risks with their associated project activities (Aleshin, 2001, Hillson *et al.*, 2006). An example of a simple RBM is shown in Figure 10.

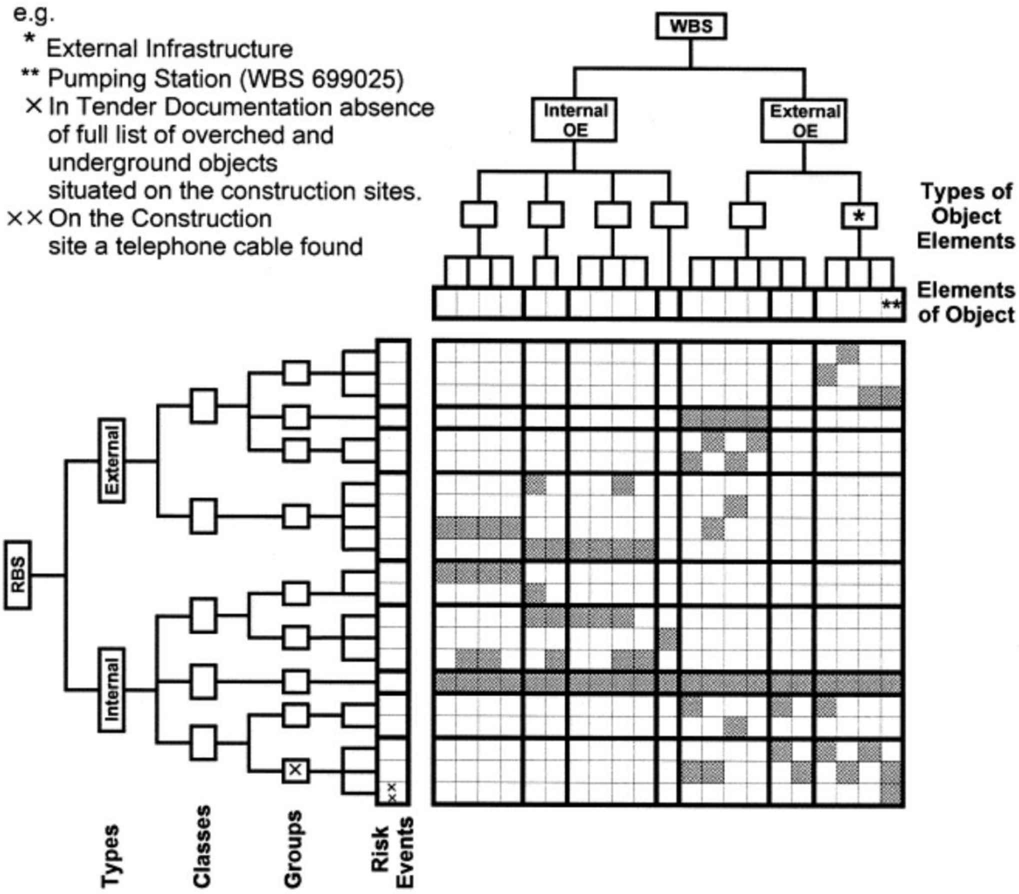



Figure 10. Example of RBM (Aleshin, 2001)

- As each risk is in association with one or more tasks in the WBS, the RBM enables the identification of how long the risk may exist and how many tasks the risk may affect. Meanwhile, it is observed that the WBS and 4D BIM share the same schedule or project activity information. Therefore combining the use of RBM and BIM provides the potential for considering risks associated with the project schedule and BIM.

## 4.3 Management and use of previous knowledge and experience

Construction is among the most hazardous and dangerous industries in the world (Sacks *et al.*, 2009). In the U.S., it is reported that over 157 bridges collapsed between 1989 and 2000 (Wardhana and Hadipriono, 2003), and more than 26,000 workers lost their lives on construction sites during the past two decades (Zhang *et al.*, 2013). Globally, the International Labour Organization (ILO) estimates that approximately 60,000 fatal accidents happen every year (ILO, 2005). Such serious accidents may not only lead to a bad reputation for the construction industry but also trigger further risks such as project failure, financial difficulty and time overruns. To avoid such serious accidents and improve the performance of risk management in future projects, a few studies (Dikmen *et al.*, 2008) suggested project practitioners should learn the valuable lessons from previous accidents and embed the consideration of risk management into the development process of a project. Learning from the past is a fundamental process in project risk management that helps individuals and organisations understand when, what and why incidents happened, and how to avoid repeating past mistakes (Goh and Chua, 2009b).

In general, the process of solving new problems based on experience of similar past problems is known as Case-Based Reasoning (CBR) (Jonassen and Hernandez-Serrano, 2002), which examines what has taken place in the past and applies it to a new situation (Kolodner, 1993), and could be of particular help in identifying and mitigating project risks at early stages, e.g. design and construction planning. In order to facilitate CBR for practical use in the construction industry, some efforts have been observed in collecting risk cases and establishing a risk case database. For example, Zhang *et al.* (2016) developed a database containing 249 incident cases to support risk management for metro operations in Shanghai. And there are more than 600 verified reports about structural risks on the Structural Safety website (Structural-Safety, 2016) and similarly the National Institute for Occupational Safety and Health (NIOSH, 2016) has established a database of over 249 reports on construction accidents. In addition, for identifying the reasons that contribute to collision injuries, Esmaeili and Hallowell (2012) reviewed and analysed over 300 accident reports. However, as a risk case database often contains a huge amount of data where reports are written in everyday language, manually reviewing, analysing and understanding these reports is time-consuming, labour-intensive and inefficient work. Failure in extracting 'correct' cases and information within a limited time often may mean that the importance of learning from past experience is missed. Hence, some researchers (Goh and Chua, 2009b, Tixier *et al.*, 2016, De Mantaras *et al.*, 2005) pointed out that a key challenge in current CBR research for project risk management is how to quickly and accurately retrieve relevant risk case data from the database so that knowledge and experience could be incorporated into new risk identification and assessment in a timely manner.

In recent years, with the development and growing use of Natural Language Processing (NLP) in the computer science discipline, some researchers have been trying to introduce NLP into the construction industry to address the analysis and management issues of textual documents, e.g. retrieval of CAD drawings (Hsu, 2013), automatic analysis of injury reports (Tixier *et al.*, 2016), and automatic clustering of construction project documents based on textual similarity (Al Qady and Kandil, 2014). It could be seen that NLP is a promising technique in assisting the knowledge and case retrieval of CBR. However, very few studies have been found in this field. In addition, Goh and Chua (2009b) stated that very few NLP tools nowadays appear to be suitable for the construction industry.

### 4.3.1 CBR in risk management

CBR is a branch of Artificial Intelligence (AI) and its origin can be traced back to the work of Roger Schank and his students in the early 1980s (Schank, 1983, De Mantaras *et al.*, 2005, Schank *et al.*, 2014). The core philosophy behind CBR is that previous knowledge and experience can be recalled and used as a starting point to solve new problems in many fields. In the project management domain, CBR has been recognised as an important technique for risk identification and analysis (Forbes *et al.*, 2008) and a number of applications have been developed, e.g. construction hazard identification (Goh and Chua, 2009b, Goh and Chua, 2009a), safety risk analysis in subway operations (Lu *et al.*, 2013), and construction supply chain risk management (Kumar and Viswanadham, 2007). Figure 11 shows the classical model of a CBR system adapted from a previous research by Aamodt and Plaza (1994). Basically, the implementation cycle of CBR contains four main processes: RETRIEVE, REUSE, REVISE, and RETAIN (known as 'the four REs'), where RETRIEVE is the first and the most important process in any CBR systems (Lu *et al.*, 2013).

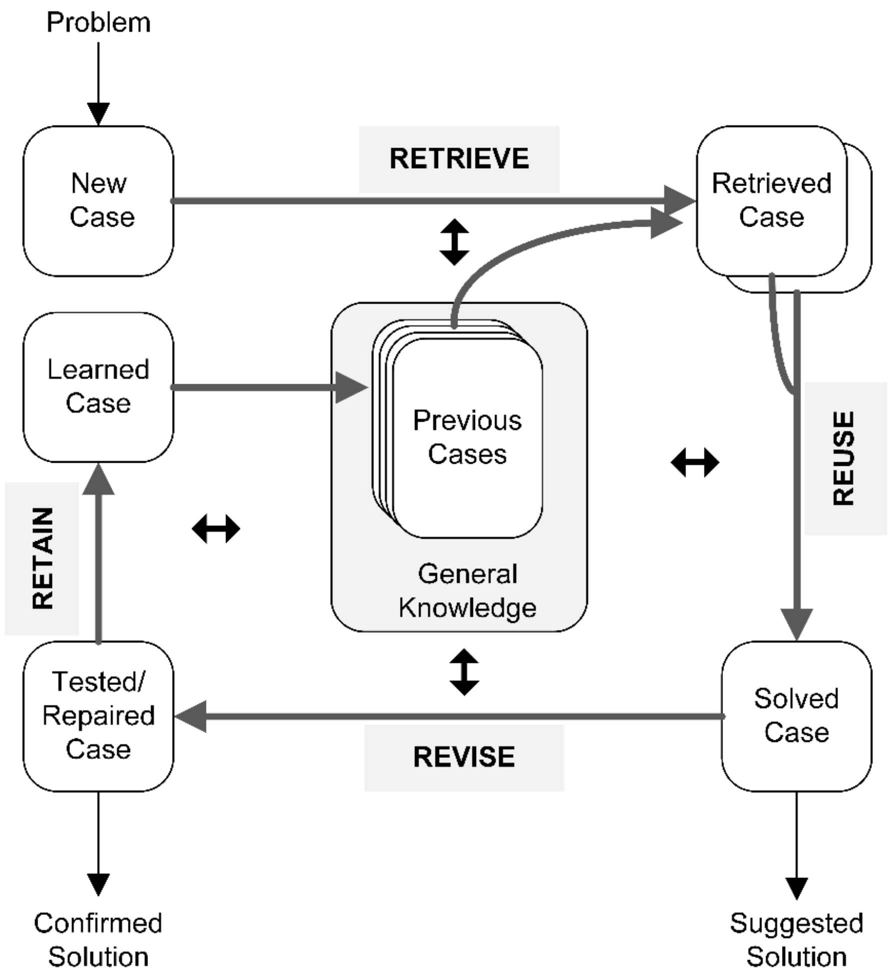



Figure 11. Classical model of a CBR system (Aamodt and Plaza, 1994)

RETRIEVE is a process of searching and determining the most similar and relevant case or cases (De Mantaras *et al.*, 2005, Aamodt and Plaza, 1994), and its importance can be viewed from the following three main aspects: (1) it acts as the only medium for helping individuals extract information from a risk case database; (2) as a risk case database often contains a large number of ‘human language’ based documents, the performance of case retrieval will have direct influence on the quality and accuracy of retrieved cases; and (3) the inefficiency of case retrieval seriously affects the user experience, which may lead to the importance of previous knowledge and experience being overlooked.

Currently scoring the similarity through allocating weights to factors is the most common method in case retrieval. For example, Lu *et al.* (2013) employed a semantic network approach to calculate the similarity value between two accident precursors. Karim and Adeli (2003) collected risk data into Excel tables and developed an attribute based schema for calculating the similarity between two cases. Goh and Chua (2009b) proposed a sub-concept approach based on a semantic network. Other efforts include, for example, evaluation of attributes (Kolodner, 1993), taxonomy tree approach (Cunningham, 2009), and the ontology-based method (Zhao *et al.*, 2009).

However, challenges and limitations also exist in current efforts, which are summarised as follows:

(1) Existing studies are very limited in scope. For example, the CBR system developed by Lu *et al.* (2013) predefined the potential accidents in subway operations and the similarity calculation is based on attributes that are to some extent subjective. Similarly, the prototype proposed by Karim and Adeli (2003) calculated the similarity index based on different weights of attributes and is only designed for highway work zone traffic management.

(2) A large amount of pre-processing or preparation work is needed. For instance, the sub-concept approach (Goh and Chua, 2009b) needs to establish a semantic network map of variables and each semantic network is constructed based on the analysis of cases and allocation of weights. Goh and Chua (2009b) acknowledged that organisations implementing the system need to consider the cost for establishing and maintaining the semantic networks and risk cases.

(3) Very few studies have been found that address the challenge of semantic similarity in case retrieval. Semantic similarity is defined as "*a metric defined over a set of terms*

*or documents, where the idea of distance between them is based on the likeness of their meaning or semantic content as opposed to similarity which can be estimated regarding their syntactical representation*" (Harispe *et al.*, 2015). Semantic similarity problems can be observed in, for example, synonyms (e.g. 'building' and 'house'), hyponyms (e.g. 'structure' and 'bridge'), and even related words (e.g. 'car' and 'bus'). Because risk case reports are all written in everyday human language and in different ways of expressing meaning by different individuals or organisations, the outcomes of case retrieval will be incomplete if a CBR system fails to consider semantic similarity. Therefore, De Mantaras *et al.* (2005) pointed out that improving the performance through more effective approaches to similarity assessment has been an important research focus in CBR.

### 4.3.2 Natural Language Processing

Natural language processing (NLP) is an interdisciplinary topic overlapping in computational linguistics, AI, and computer science that deals with the interactions between computer and human languages (Chowdhury, 2003). NLP started its early work in the 1950s in exploring the fully automatic translation between different languages (Bar-Hillel, 1960), and in recent years has seen a rapid increase in use and development in computer science. The application areas of NLP are very wide including, for example, machine translation, question answering, speech recognition and information retrieval (Jurafsky and Martin, 2009).

Information retrieval (IR) refers to the process and activity of extracting useful information from a collection of information resources (Baeza-Yates and Ribeiro-Neto, 2011). Due to the needs of managing and using the fast-growing volume of information (Bai, 2011), many IR systems have been developed and the best examples include web search engines (e.g. Google and Yahoo), and library resource retrieval systems (Murty and Jain, 1995).

In the construction industry, even a small project generates a large amount of digital information such as specifications, computer-aided drawings, and structural analysis reports (Soibelman *et al.*, 2008, Tixier *et al.*, 2016). In addition, in order to learn from past experience and avoid similar accidents on new projects, lots of investigations and analysis on previous accidents have been conducted and the resulting reports and feedbacks are important to improving the existing knowledge and standards (Kaminetzky, 2001). Currently major companies and organisations are using databases for managing those accident reports (Tixier *et al.*, 2016). However, new documents continually need to be added into databases and therefore the size of databases is increasing. Moreover, these reports are written in human language and in different ways of expression by different individuals or organisations, and a challenge is how to retrieve valuable and 'correct' information from the database quickly and efficiently.

To improve the use and management of 'human language' based engineering documents, a recent research trend is to take advantage of NLP. For example, Hsu (2013) made the use of the classical VSM and developed a Content-based CAD document Retrieval System (CCRS) for assisting the management of CAD drawings and quick retrieval of documents according to given queries. By taking the advantage of keywords extraction of NLP, Tixier *et al.* (2016) developed a prototype supported by the R programming language for automatically extracting precursors and outcomes from unstructured injury reports. Al Qady and Kandil (2014) proposed a method for automatic clustering of construction project documents based on textual similarity. Caldas and Soibelman (2003) developed a prototype system to automatically classify a large number of electronic text documents in a hierarchical order in the information management system. Another study took advantage of text mining and proposed an ontology-based text classification method for job hazard analysis (Chi *et al.*, 2014). In addition, Pereira *et al.* (2013) presented a solution to extract valuable information from incident reports in real time to assist incident duration prediction. However, very few studies exist in this field and new investigations are still needed.

It is observed that there are two main features in applying NLP into textual document management in the construction industry:

Firstly, most state-of-the-art studies of NLP still lie in the computer science discipline and most modern applications are often used to treat extremely large volumes of data e.g. extracting online information (Khribi *et al.*, 2008) and library management (Baeza-Yates and Ribeiro-Neto, 2011). In contrast, the sizes of electronic data in any construction project and risk cases in any database are relatively small. Hence, there is a need to select the appropriate methods and techniques for specific purposes. For example, Tixier *et al.* (2016) pointed out one difficulty in implementing machine learning for automatic safety keywords extraction is that the small number of injury reports cannot form a satisfactory training database and therefore they developed a NLP system based on hand-coded rules.

Secondly, unlike online websites containing often several aspects of information, construction project data and risk cases are relatively restricted to certain topics and thus there is a need to establish the context or rules in processing them. For instance, when applying ontology and text mining into job hazard analysis, the author predefined the list of potential safety hazards and emphasised the importance of defining the knowledge and resource scope into the construction safety domain (Hsu, 2013).

Some existing efforts (Al Qady and Kandil, 2014, Hsu, 2013, Tixier *et al.*, 2016) have shown that the application of NLP techniques in managing textual data is a new research trend in the construction industry and NLP has the potential to address the current challenges of case retrieval of CBR. However, very limited numbers of studies have been found in this area. In order to further improve the efficiency and performance of risk case retrieval, this study proposed an approach of combining the use of the two NLP techniques, i.e. VSM and semantic query expansion, and outlines a framework for the risk case retrieval system, as described in Chapter 7. The idea was motivated by the following observations:

- VSM is known as one of the most important IR models (Baeza-Yates and Ribeiro-Neto, 2011) and it can be used for information extraction, indexing and relevancy ranking, etc. For example, Caldas and Soibelman (2003) used VSM for characteristic information extraction and automatic classification of project documents. Similarly, Hsu (2013) embedded VSM as a core technique in their retrieval system of CAD drawings. Hence, VSM is potentially helpful in evaluating the relevance between user need and risk cases in a CBR system.

- Understanding the relations between words (e.g. hyponymy, synonymy) is an important step in fully using the concept of semantic similarity (Jurafsky and Martin, 2009). Thus, some individuals and organisations have started to establish lexical 'dictionaries' that pre-defined the semantic relationships between words, where the most commonly used resource for English sense relations is the WordNet lexical database (Fellbaum, 1998, Jurafsky and Martin, 2009). So far a number of studies (Gong *et al.*, 2005, Snasel *et al.*, 2005) have used WordNet for improving web retrieval through expanding the query terms using related words in WordNet and have proved this approach could partially address the semantic similarity issues and improve the performance and completeness of information retrieval. Therefore, the basic principle of semantic query expansion is also applicable for improving the completeness and quality of case retrieval.

This page is intentionally left blank

# Chapter 5. RISK ANALYSIS AND RBS DEVELOPMENT

In order to develop a tailored RBS for bridge projects and formalise an active 'link' between RBS and BIM, a three-step approach was used, which is described in Sections 3.3.3 and 3.3.4. The scope of this research is limited to bridge projects; however, the basic methods and principles could be also applied to other AEC projects.

## 5.1 Separation of the integrated bridge information model

In concept, data from different contents and disciplines is gradually defined and added to build the integrated bridge information model and BIM can be considered as a shared knowledge and information repository to support the whole project lifecycle. The data and their structure in a complete BIM are extremely complex, therefore some researchers, e.g. Fischer and Kam (2002) and Haymaker *et al.* (2003), realised that there is a need for separation of the integrated information model to meet particular needs. In 2005, Kiviniemi (2005) proposed a formal solution for dividing a project's data set into several sub-models and linking user requirements with these sub-models. Similarly, as current neutral information exchange formats such as IFC have limited supports on infrastructure structures such as bridges, Shim *et al.* (2012) divided the integrated bridge information model into five levels for different purposes of use, e.g. structural analysis, structural detailing, and construction simulation.

Level of Content (LOC) is defined in this study as the primary hierarchical structure of content of BIM. LOC could be used for decomposing an integrated 3D information model into separated sub-models according to different information content and different disciplines, which could enable the 'correct' information to be extracted, used and communicated in an efficient way to meet particular requirements, e.g. structural analysis, construction scheduling simulation, and risk management. Currently there is no explicit method for separating LOC but the separation should meet particular needs and requirements. Kiviniemi (2005) defined the technical system as an aggregation of objects that have a common purpose or function or to provide a service, which originates from the definition of ifcSystem by buildingSMART. Although some researchers (Tah *et al.*, 1999, Shim *et al.*, 2012) tried to summarise the component objects for bridge information model, no study has been found to classify and group bridge component objects that have a common purpose or function or to provide a service to be a part of a technical system.

After a critical analysis, this study divided the integrated bridge information model in concept into four LOCs (i.e. Project, Surrounding Environment, Site, and Bridge) and six technical systems (i.e. structural system, expansion joints system, decking system, drainage system, lighting system, and parapet system) for risk management, as shown in Figure 12.

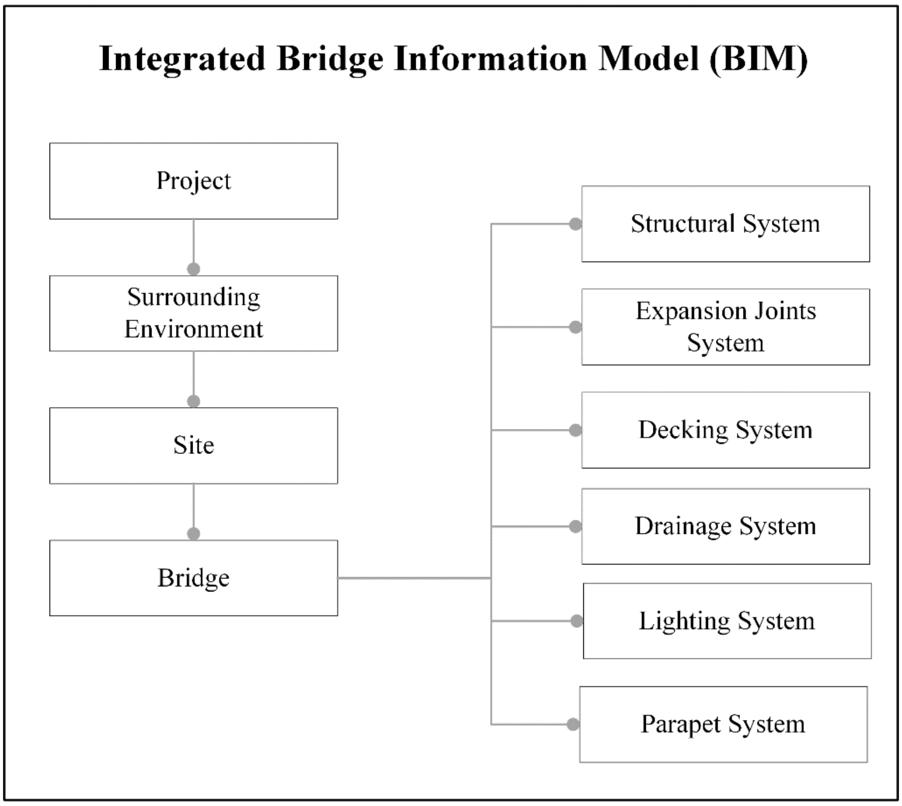



Figure 12. Conceptual Division of an integrated bridge information model

The separation of BIM into four LOCs and six technical systems for risk management in this study is based on the following four observations:

1) IFC is a neutral and open data model specification describing AEC project lifecycle data, which is developed and managed by the buildingSMART International (formerly known as International Alliance for Interoperability or IAI). IFC can be used for data exchange by a number of software, e.g. ArchiCAD® and Revit®. An IFC model is composed of IFC Entities built up in a hierarchical order and the primary IFC element hierarchy is: Project, Sites, Buildings, Storeys, Spaces, Elements (Eastman, 1999). Although currently IFC still has some limitations in supporting bridge and other infrastructure projects and one of the latest ongoing projects is to develop a standard format of IFC-Bridge (buildingSMART, 2016b), the principle of basic hierarchical

structure of IFC is also applicable for separation of LOCs and technical systems of a bridge information model.

2) For establishing a linkage between user requirements and BIM, Kiviniemi (2005) suggested that a model for client requirements could be divided into five basic levels (i.e. project, site, building, building storey, space) and 12 technical systems (e.g. building envelope, structural system, HVAC system) according to the IFC specification. This principle of separation could be an important guidance for this study for considering risks. The scope of risk is very wide – some, such as financial and political risks, may cause effects to the whole project and some others, such as the structural safety risks, may have a direct relation to part of the temporary structure or even a small component. Therefore, the core idea of this study is that risks from different sources could be divided into groups and linked to the four LOCs and six systems of an integrated information model, visualised in BIM and managed intelligently in a database during the development process.

3) Bridges share many common features with buildings. For example, both bridges and buildings are construction projects which will go through project phases such as briefing, design, construction, and maintenance. Although having different functionalities, both bridges and buildings are structures and need project participants (e.g. client, designer and contractor) to work collaboratively to complete the one-off endeavours. As a result, existing hierarchical structures of BIM could be important references for developing the LOCs and technical systems of the bridge information model. For instance, a number of studies (Ji *et al.*, 2013, Yabuki and Li, 2006) investigated developing a neutral data model IFC-bridge by extending the current standard IFC to cover bridge components. In addition, Tah *et al.* (1999) used the levels of hierarchy such as project, product, in-situ, and bridge when describing the object

classes of an integrated bridge information model. Therefore, in concept the bridge information models could also share some same LOCs with buildings, such as project, site and bridge.

However, for risk management, bridges are to some extent different from normal buildings. Liu *et al.* (2014) summarised that the characteristics of bridges include, for example, complex structural design, a large number of heavy components, and complicated site conditions. In addition, most bridges as part of a transport system need more information of a relatively large area such as the local economy, the potential number of citizens to be benefited, surrounding topography and geology, and existing roads and tunnels. For example, the Mersey Gateway Bridge as part of a big highway project is a cable-stayed structure with three towers being constructed in the UK. The design and access report (HBC, 2008) indicated that the design and construction of the bridge needs to combine information of the surrounding environment all together for considerations, e.g. the project influence on the local community, possible restrictions from the local airport and power station. For instance, the height of bridge towers should be restricted to meet the height zoning map by Liverpool Airport for safety purposes. Therefore, considering not only risks but other project decisions for bridges, a surrounding environment level between project and site is needed for the LOCs of infrastructure projects such as bridges. The surrounding environment defined by this research means in concept a relatively wide geographic area and opposites to the relatively narrow sense of site. In addition, as most bridges do not have complex requirements on space and storeys (Ryall *et al.*, 2000), LOCs such as space and storeys of buildings are not necessary for bridges.

4) Whatever the type of bridge is, according to Ryall *et al.* (2000), Zhao and Tonias (2012) and Fan (2012) bridges contain basically five major components (i.e.

superstructure, bearings, pier and pier caps, foundations, piles) and five minor components (i.e. deck pavement, drainage system, parapets, expansion joints, and lighting). The main purpose of most bridges is to span physical obstacles such as rivers and valleys, which is heavily reliant on structural components. Liu *et al.* (2014) also highlighted that one significant feature of bridges is the complex structural design. Therefore, this research grouped all structure-related components into the structural system and defined in total six technical systems as shown in Figure 12. The structural system includes bridge components such as girders, cross-beams, cables, towers (pylons), anchor blocks, bearings, abutments, piers.

## 5.2 Developing a knowledge-based risk database

This section demonstrates a knowledge-based approach of mining risk data to develop a knowledge-based risk database. As stated in Sections 4.2, a comprehensive risk database could be an important tool for helping the project decision makers develop an overall understanding of, quickly identify and effectively analyse and mitigate risks. Meanwhile, the knowledge-based risk database also provides information and theoretical basis for developing a RBS.

A number of existing studies have been conducted to obtain the 'complete' risk database, e.g. (Kartam and Kartam, 2001, Wang and Chou, 2003). Because the scope of risks is very broad, it is somehow difficult to obtain a complete risk database which can be applied to all industries and projects. As different risks are highly linked to certain type of projects and particular construction markets and conditions, most existing studies tend to apply some conditions to narrow the scope of the risk database and target particular types of project. For example, El-Sayegh (2008) summarised a list of 33 main risk factors for highway construction projects in the UAE through a critical literature review. Zayed *et al.* (2008) sent a questionnaire to 17 highway construction experts in China to collect a list of risks and then classified them into company level and project level for further analysis.

This study used a knowledge-based approach consisting of three basic steps (i.e. risk data collection, risk data mining, and risk data assessment and translation) to identify possible risks for bridge projects and develop the risk database, as shown in Figure 13.

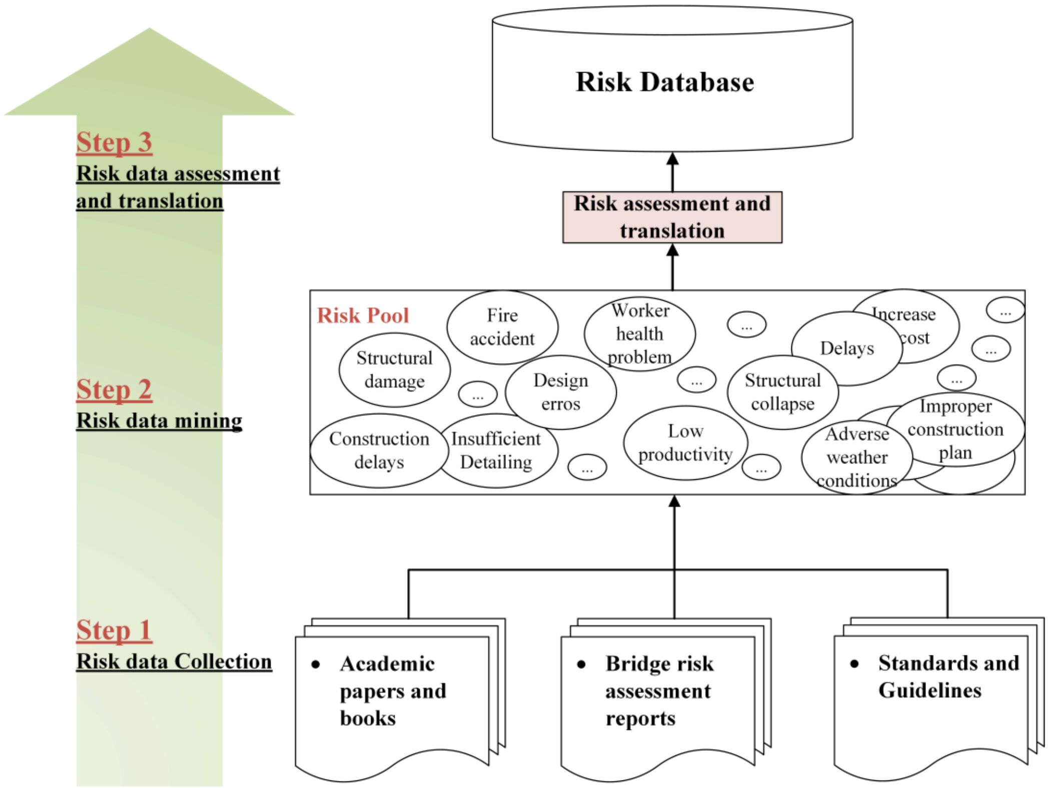



Figure 13. Process of developing a knowledge-based risk database

### 5.2.1 Step 1: Risk data collection

The first step in developing a knowledge-based risk database is to collect and prepare risk data as the basis for the next step. In a construction project where people are from different disciplines and have various educational backgrounds and work experience, it is relatively difficult to identify a complete list of risks through a limited number of interviews or surveys of literature and there is a need to investigate a wide range of sources to develop the risk database.

This study obtains risk data from 80 documents based on a web-search approach mainly from the following three sources: 1) academic publications, 2) bridge risk assessment reports, e.g. (Atkins, 2006, Structural-Safety, 1997), and 3) standards and guidelines on risk management, e.g. (HSE, 2015, Molenaar *et al.*, 2006, PMI, 2004). Reasons for choosing the three sources are: 1) the large number of published academic papers and books are easily accessed and contain research on the identification of risk factors for different countries and regions for all kinds of construction projects, which have summarised a relatively complete list of risks for construction projects; 2) risk assessment reports of bridges and related standards have recorded a number of identified risks in real environments and some are highly relevant to bridge projects, which are an important supplement for academic publications. However, there is a need to recognise that as construction projects, including not only bridges but also buildings, roads and industrial plants, share many common risks and there is only a limited number of publications focusing on bridge related risks, the scope of collecting academic papers and standards is extended to all construction projects.

### 5.2.2 Step 2: Risk data mining

The second step is to search for valuable risk information from the data collected in Step 1 by adopting a manual text mining approach. Specifically, a manual analysis through careful reading of each document and interpreting and understanding the text in its relevant context was conducted to identify and record the risk information. As the collected documents use different methods and standards to describe risks, e.g.

‘cost increase’ and ‘budget overrun’, it was then important to classify similar risks and put them into different risk groups individually according to the source of risk. Currently there is no consensus on how to develop the RBS (Mehdizadeh *et al.*, 2013), thus a list of key words (e.g. project risk, external risk, global risk, design risk) were identified from previous studies (Tah and Carr, 2001, Choi and Mahadevan, 2008, Mehdizadeh *et al.*, 2013) to be an initial hierarchy for allocating and managing the collated risk information. After this, all identified risk factors and corresponding information were organised into groups and stored in an initial database which is defined as a ‘risk pool’ in this research.

### 5.2.3 Step 3: Risk assessment and translation

In the third step, the identified risk factors and information were further assessed group by group, where the same or similar risks described in different ways were translated to the same format to avoid duplicated data. A concise knowledge-based risk database was then structured and developed. An example of the knowledge-based risk database is shown in Table 3 and its complete version with references is presented in Appendix A.

Table 3. Example of knowledge-based risk database

| Risk Category | Risk Factor | Risk Description | Possible Mitigation Strategy |
|---|---|---|---|
| Economic | Inflation | Price inflation of construction materials; Monetary inflation; Unanticipated local inflation and interest rates due to immature local economic and banking systems; Increase of wages and welfare | 1) Escalation Clause;<br>2) Price Contingency in the Bid;<br>3) Project Financing by a Reputable Owner;<br>4) Owner Purchase of Equipment & Material;<br>5) Providing Performance Bond and Prequalification of Suppliers;<br>6) Forward Contracts for Hedging Exchange Rate Changes |
| | Currency | Rate fluctuation; devaluation; difficulty in converting foreign currency | |
| | National and international impacts | Changes by International Associations such as OPEC | |
| | Inadequate market demand | Inadequate forecast of market demand; Owners' unreasonable upfront capital demand | |
| Design | Unqualified or defective design | Insufficient planning; Incomplete design scope; Difficult and complex construction; Improper site estimation; Improper material use; Lack of experience and knowledge in design; Inadequate specifications | 1) Changed Condition Clause (Delay);<br>2) Contractor Participates in Design;<br>3) Adoptable Design/ Construction Methods;<br>4) Changes to the Original Design |
| | Errors and mistakes | Carelessness; Lack of experience and knowledge in design; Inadequate specifications; Incorrect quantity calculation; Competence | |
| | Delays of design works | Low productivity; Work order change; Delays in design and regulatory approval | |
| Construction | Deviation between design and construction | Defective design and errors | 1) Contingency in the Bid;<br>2) Insurance for Liability from accidents;<br>3) Contract Clause for Time Extension Due to Delays;<br>4) Safety and Training Programmes from Employees;<br>5) Planning Procurement Activities in Advance |
| | Inadequate construction planning | Inadequate consideration on the actual condition of the construction site; Unfamiliarity with the design drawings and design intention; Insufficient site information and unforeseeable circumstances underground; Unreasonable personnel organisation and arrangement; Unreasonable materials and unreasonable equipment allocation; Lack of knowledge and experience | |
| | Improper construction methods | Unfeasible construction methods; Lack of knowledge and experience | |
| | Construction changes and delay | Third party delays; Delay of drawing supply; changes in work; Owner changes; Construction delay; Delayed site access; Late drawings and instructions; Delays in material supply; Improper intervention | |
| | Poor construction quality | Unqualified workmanship and skills; Improper material use; Violating construction standards; Cutting corners | |
| | Increase of cost | Cost of tests and samples | |
| | Low construction productivity | Obsolete technology and practices by local partners; poor skills or inadequate supervision; Shortage of skilled and unskilled workers; Foreign firms face difficulties in hiring and keeping suitable and valuable employees; Insufficient labour; Productivity of equipment | |
| | Improper project management | Improper project budgeting; Inadequate project organisation structure; Incompetence of local project team; Incompetence of subcontractor | |
| | Failure to identify defects | Insufficient inspections | |

## 5.3 Developing a tailored RBS linking to BIM

Built on the obtained knowledge-based risk database, this section further clusters data to develop a tailored RBS and proposes a conceptual model to link the resulting RBS to the four LOCs and six technical systems of an integrated bridge information model.

In this study, there are two major motivations for developing a tailored RBS. Firstly, although a number of RBS have existed, e.g. (Holzmann and Spiegler, 2011, Sigmund and Radujković, 2014, Tah and Carr, 2001), the current RBS vary in both form and content. Meanwhile, Mehdizadeh *et al.* (2013) stated that currently there is no consensus on the standards or general methods of developing a RBS and the RBS to be developed should satisfy the particular purposes and requirements. Furthermore, no existing studies have been found to develop a tailored RBS for linking it to BIM for risk management. Secondly, it was observed that a crucial role of RBS is to classify risks in a proper structure and the development of RBS is reliant on the collected risk data. However, only a limited number of studies have been found to focus on risks for bridge projects and most of them only partly summarise some of the major construction risks for their own country or local area (Q. F. Li *et al.*, 2013).

Through a critical analysis, the resulting RBS and its basic relationship with BIM are proposed as shown in Figure 14. Specifically, project risks in this research are basically divided into two main groups – external risks and internal risks. The idea has also been adopted by other researchers, e.g. (Fang *et al.*, 2004, El-Sayegh, 2008, Tah and Carr, 2001).

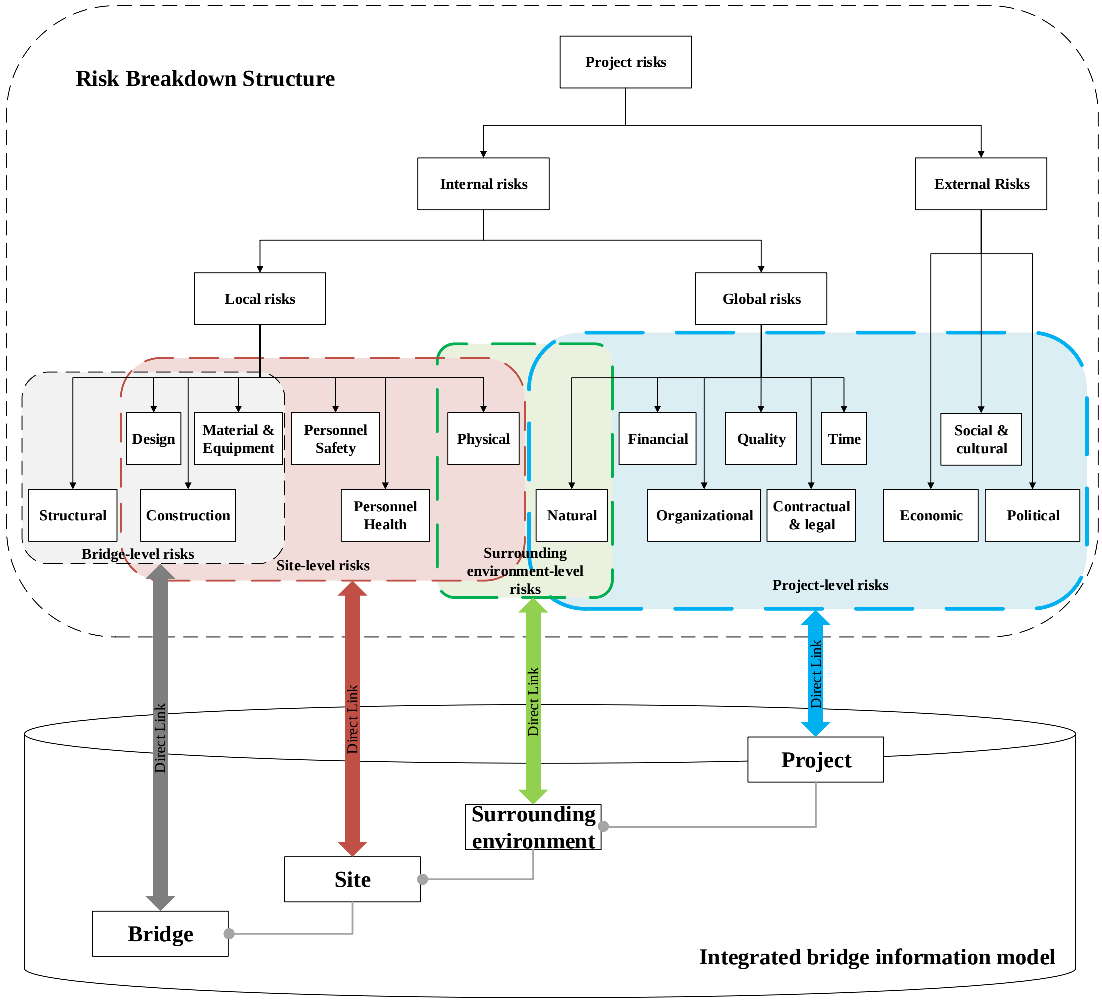



Figure 14. Basic linkage between RBS and BIM

External risks (Fang *et al.*, 2004, El-Sayegh, 2008, Tah and Carr, 2001) mean those risks that are outside the project and beyond the control of the project team. The external risks include political, economic, social and cultural risks. For example, the political risk may refer to the changes or variation of local laws and the economic risk could be the fluctuation of local currency. As external risks are at a macro level such as company or country levels and are not under the control of the project team, there is a need for a continuous scanning and forecasting through all phases of the project and drawing up company strategies to manage their effects (Tah and Carr, 2001).

Internal risks (Fang *et al.*, 2004, El-Sayegh, 2008, Tah and Carr, 2001) refer to those that are within the project and are more controllable by the project team. The scope of internal risks is much broader than external risks and there is a greater opportunity for the project team to manage them. The number of internal risks in the knowledge-based risk database is much larger than the number of external risks and the relation between different internal risks are inter-related and much more complex. Therefore, the internal risks were further divided into two groups – local and global – because some internal risks are related to the whole project whereas the others may cause effects local to the bridge or individual work packages (Tah and Carr, 2001).

To easily establish a conceptual relationship between the four LOCs of BIM and RBS, risks at the lowest level in the RBS were further classified into four groups - project-level, surrounding environment-level, site-level and bridge-level. The purpose of this classification is that: 1) risks from different sources are grouped together to help the project team have a better understanding for risk classification and communication; 2) as BIM can be divided into different LOCs in concept, different groups of risk in the resulting tailored RBS have a direct relationship to the different LOCs of an integrated bridge information model. Meanwhile, this classification does not mean that the risks information in practice will be strictly put in a particular group. In fact, risks are highly inter-related and it has been found that some types of risk overlap in two different levels and could cause effects on both levels. For example, material and equipment risks could refer to either risks in the material used in the bridge components or risks in transport and storage of material and equipment on site. Therefore, there is a need for detailed analysis of internal relationships between risks in the real environment.

The basic relationship of linkage shown in Figure 14 is that the four groups of risk in the RBS (project-level, surrounding environment-level, site-level and bridge-level) can be linked directly to the four LOCs in the integrated bridge information model (project, surrounding environment, site, bridge). This linkage presents a general framework integrating BIM and RBS for risk management of bridge projects. There is a need to point out that the concept of 'surrounding environment' refers to a relatively wide geographic area, which is within the project environment as opposed to the relatively narrow sense of the concept of 'site'. For instance, bridge projects, especially those as a part of the major highway or railway project, have to deal with the potential risks in a relatively big surrounding environment area instead of only on site, e.g. potential conflicts between the bridge and existing road network, financial and legal risks in removal and demolition of existing facilities, or natural risks (e.g. debris flow) nearby.

To further improve the practical applicability of implementing the linked relationship between RBS and BIM, 13 sub-models of linkage were developed and two examples are shown in Figure 15 and Figure 16. In total, risks are classified into 16 main categories (e.g. structural, design, financial) and a number of sub-categories (i.e. risk factors in Table 3. These risks could have both direct links and indirect links to the four LOCs and six technical systems.

The structural risks are used as an example to illustrate the detailed sub-model of linkage (see Figure 15). Structural risks as a part of bridge-level risks are directly linked to the bridge level and structural system. For example, potential damage or collapse of both temporary and permanent structures have immediate influence to the bridge structure and should be directly linked to the whole bridge or structural system. In addition, structural risks also have indirect links to expansion joints system, decking system, and drainage system. A practical example is that in the in-use phase rain may flow into the surface cracks of bridge deck slabs and corrode the steel reinforcement, which may influence the durability and safety of the whole bridge indirectly. As another example shown in Figure 16, health and safety risks (e.g. falls from height, site traffic accidents) in most cases take place within the site area and have a direct relationship to the features of the construction site, and therefore health and safety risks are linked directly to the site level of BIM. In addition, some indirect reasons may also trigger the safety issues and there is a need for the indirect linkage. For example, structural collapse of bridges may also lead to injuries or loss of live. And electrocution or fire accidents may be caused by an unsafe lighting system or parapet system.

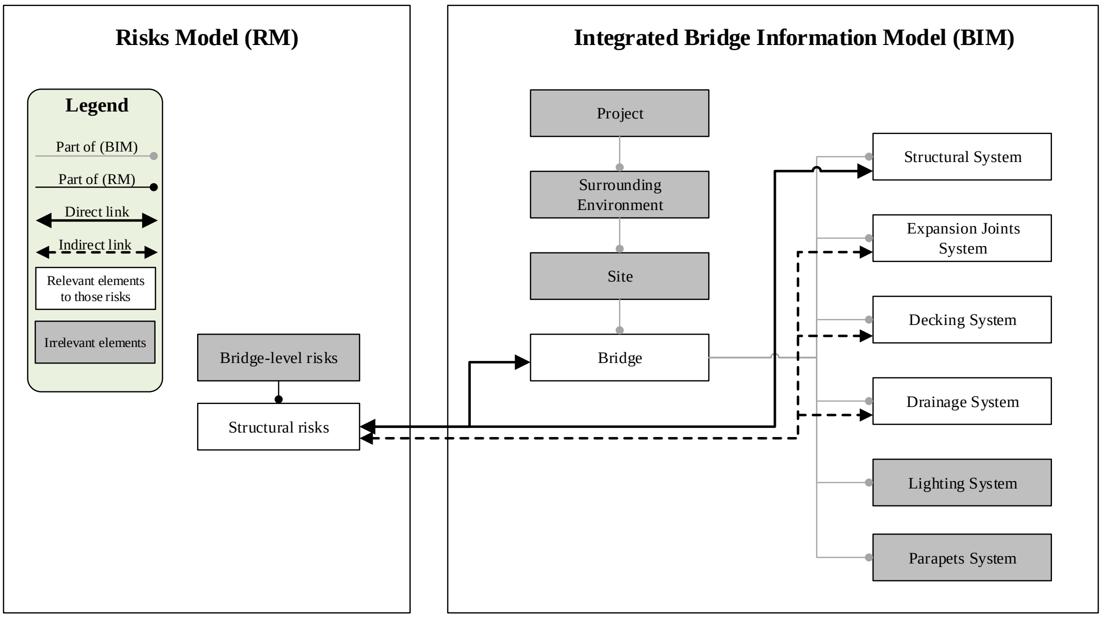



Figure 15. Linkage between structural risks and bridge information model

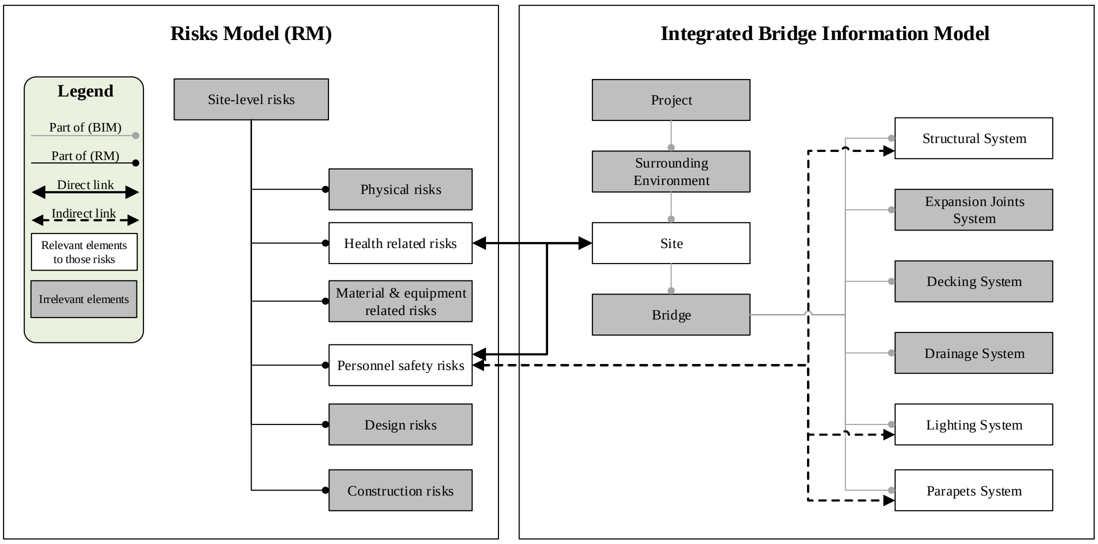



Figure 16. Linkage between health and safety risks and BIM

The rest of the developed sub-models will be described in Section 6.2.

This section developed a tailored RBS and a conceptual model for linking RBS and BIM. The tailored RBS as a hierarchical structure can be used for categorising and managing data in the knowledge-based risk database and could provide a global view on project risks. In addition, through linking risk information to the BIM, risks can be visualised and managed in the BIM throughout a project lifecycle. This proposed method merges the RBS with BIM as an integrated approach and take advantage of both methods and could effectively facilitate identification, analysis, communication, and decision making of risks.

# Chapter 6. INTEGRATING RBM INTO 3D/4D BIM FOR RISK IDENTIFICATION AND COMMUNICATION

Built on the conceptual division of BIM, tailored RBS and linkage rules, this chapter describes a novel method to link 4D BIM, RBS, and WBS to be a risk management system. Specifically, the overall framework of the 4D BIM and RBM based risk management system is presented in Section 6.1 and the full list of linkage rules are shown in Section 6.2. The overall structure, and development and implementation processes of this proposed system is illustrated in Section 6.3. Finally, an example case study for testing the feasibility of the proposed approach and tool is described in Section 6.6 and the system's benefits are summarised in Section 6.7. Although this study takes bridge engineering as an example, the methodology and principles can be applied to building engineering or other types of projects. In addition, the WBS varies from project by project and therefore the development of WBS is not included in the scope of this research.

## 6.1 Conceptual model of linking RBM to BIM

The proposed conceptual model for linking the RMB to BIM is shown in Figure 17 and it consists of three basic components: BIM, RBS and WBS. The motivations for developing this conceptual model are explained as follows.

- Firstly, RBS is a hierarchical representation of project risks arranged by category and WBS is used for breaking down a project into easily manageable tasks or components (PMI, 2004). The RBS and WBS of a project can be interconnected as a RBM, which is an easy-to-use tool for risk identification and data management and has a number of advantages: 1) the RBS organises different types of project risks into a hierarchically logical structure and can facilitate the understanding and communication of project risks from a systematic view; 2) with the observation that risks are in association with the processes of a project and may affect a certain period in the project lifecycle, combining the use of RBS and WBS enables the identification and understanding of how long the risk may exist and how many tasks the risk may affect. In addition, a number of previous studies have discussed the potential of using RBS and RBM for developing risk management software (Shim *et al.*, 2012, Aleshin, 2001, Hillson *et al.,* 2006).

- Secondly, different types of risk can affect a construction project in different ways and there are some "rules" between risks and BIM. For example, financial risks may affect the whole project while some structural risks may only have impacts on several structural elements or the structure. Therefore, there is a need to identify the relations between different types of risk and BIM, and establish the "rules" for linking RBS to BIM. The linkage rules could not only contribute to the use and management of risk data in the BIM environment but also facilitate the understanding of how a particular risk may affect BIM.

- Thirdly, WBS decomposes a construction project into tasks in a timeline-based order. As the concept of 4D BIM refers to construction schedule simulation by intelligently linking 3D BIM objects with time and schedule related information (Hartmann *et al.*, 2008), it is observed that WBS and 4D BIM share the same project schedule and activity information, and 3D BIM can be used in association with WBS to develop the 4D BIM. 4D BIM enables construction projects to be built in the virtual computer environment before real construction commences and recall related knowledge and experience for identifying any potential risks at an early stage.

- Fourthly, BIM can be seen as an information repository of a construction project and a separate central database can be used for risk data management. The related risk data can be linked to and visualised and reviewed in BIM. Then risk information can be generated from various BIM platforms and managed in a unified manner in the central database which can reduce the possibility of risk information loss caused by updates or changes of BIM or information transfer between different tools.

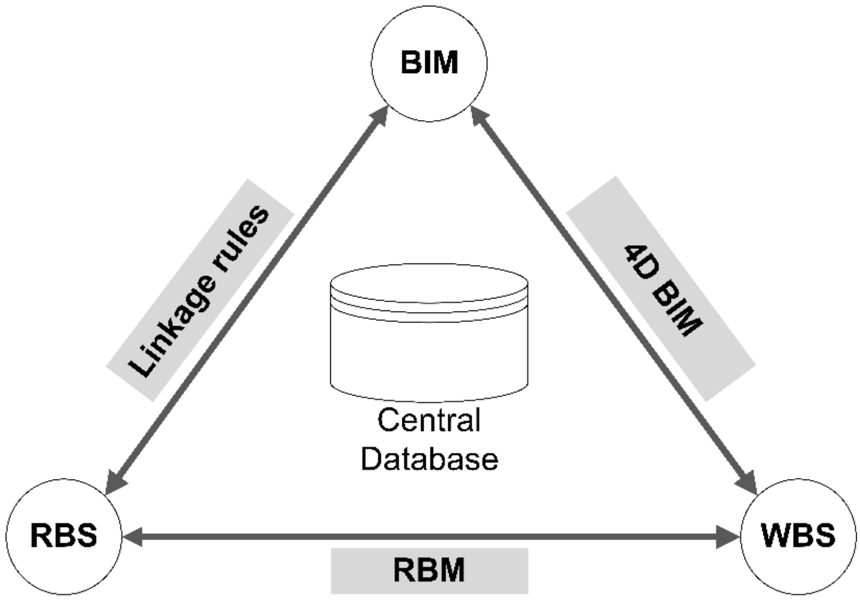



Figure 17. Conceptual model of linking RBM to BIM

## 6.2 Linkage rules between BIM and RBS

The motivation, methodology and findings of developing a tailored RBS for bridge projects and establishing the linkage between the resulting RBS and BIM were discussed in Sections 5.2 and 5.3. Generally 13 sub-models as the linkage rules between risks and BIM were developed. Apart from the structural risks and health and safety risks (as presented in Figure 15 and Figure 16), the rest of the sub-models of linkage are listed in Figures 18-28.

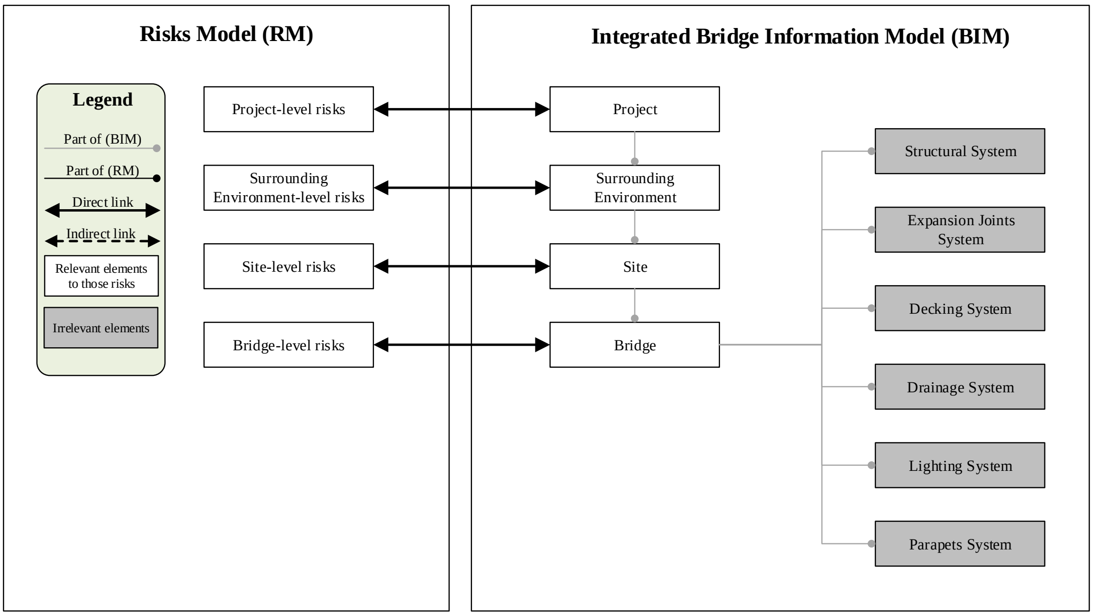



Figure 18. Basic relations between Risks and LOCs

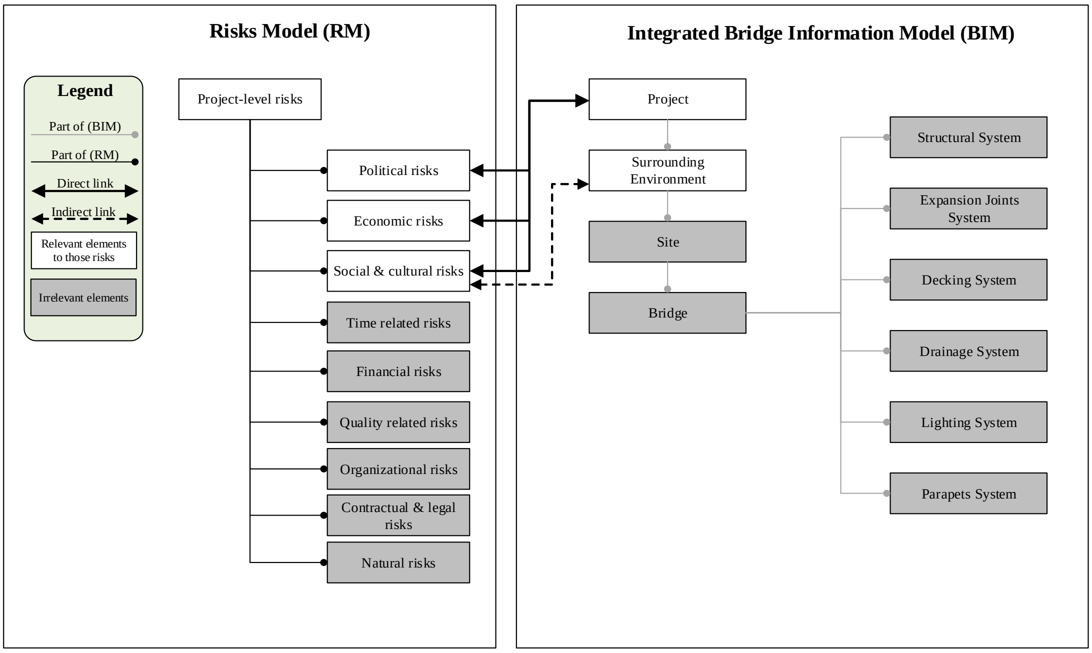



Figure 19. Linkage between political, economic, social and cultural risks and BIM

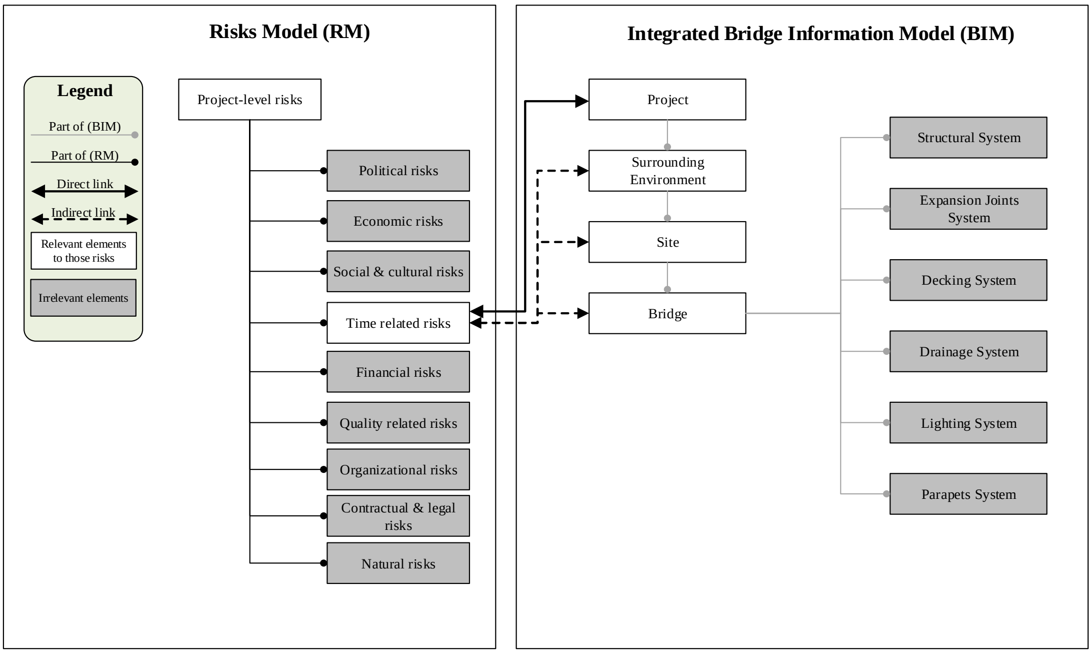



Figure 20. Linkage between time related risks and BIM

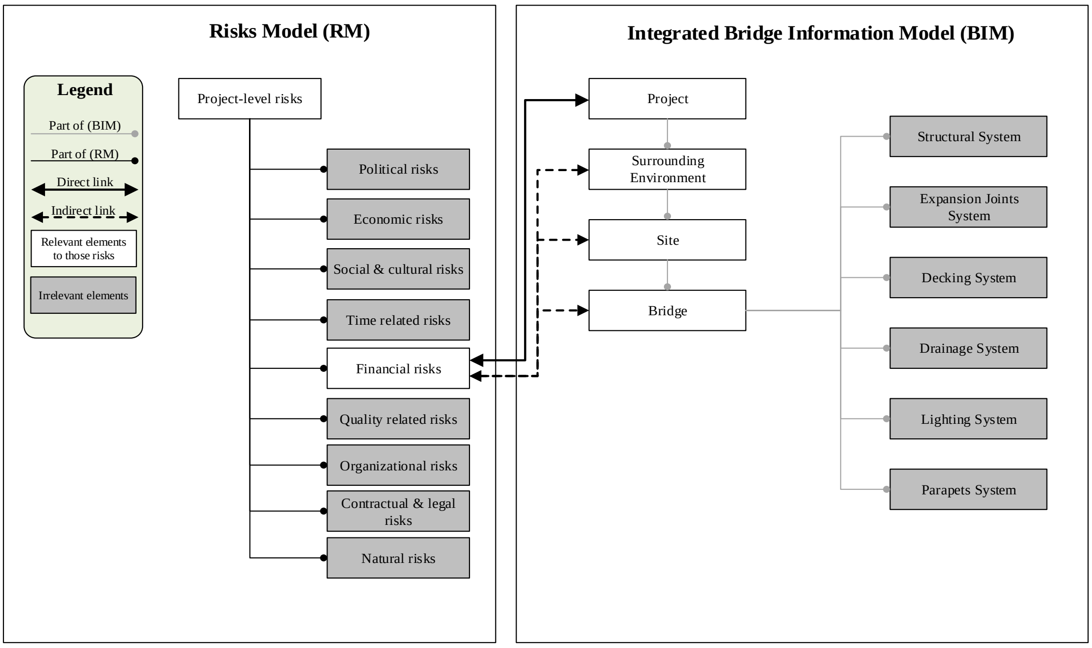



Figure 21. Linkage between financial risks and BIM

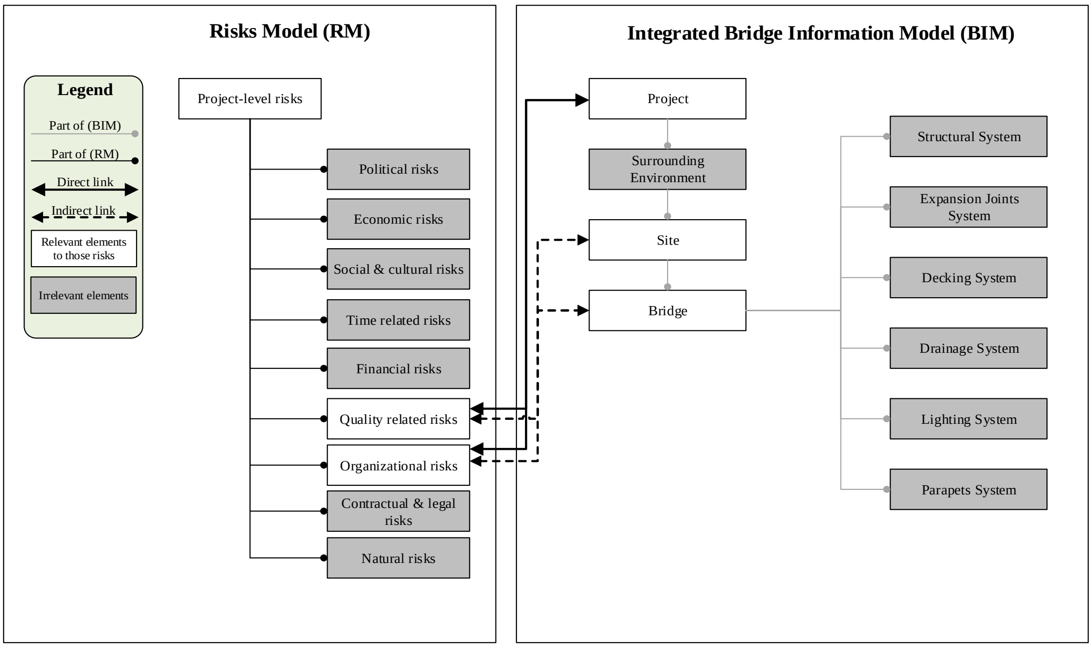



Figure 22. Linkage between quality and organisational risks and BIM

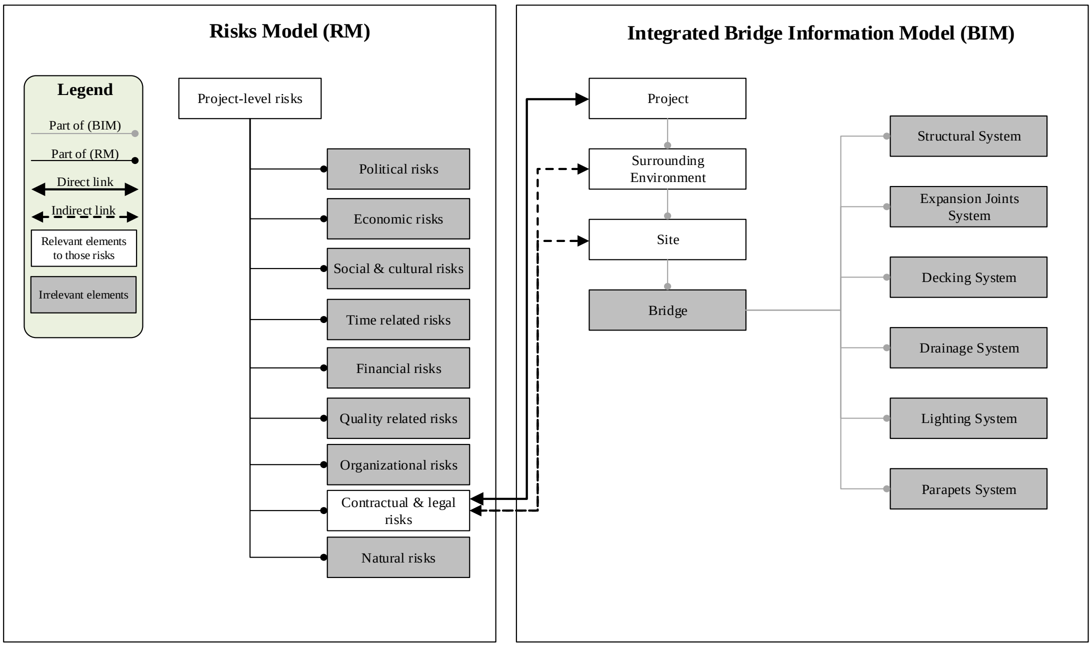



Figure 23. Linkage between contractual and legal risks and BIM

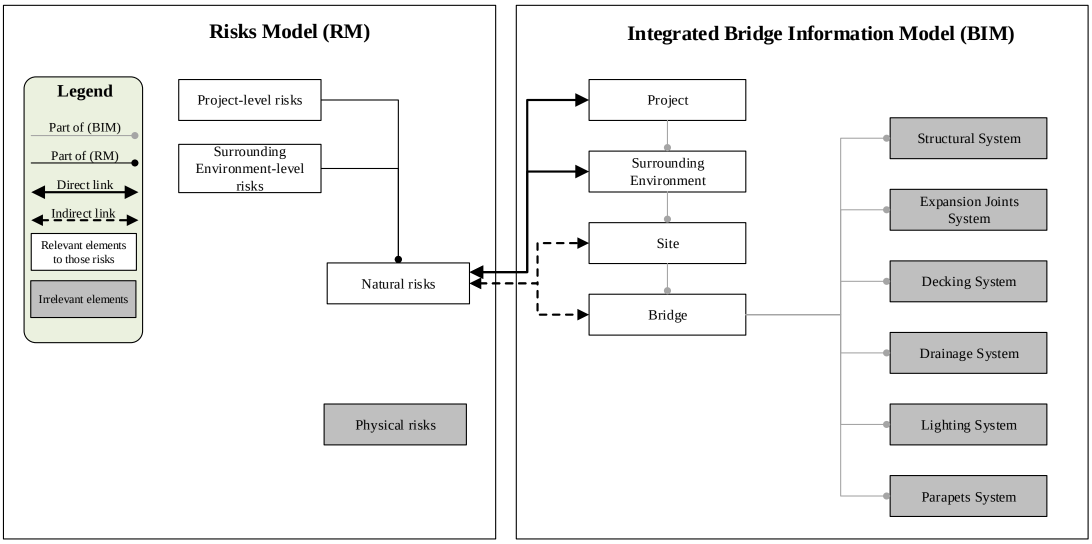



Figure 24. Linkage between natural risks and BIM

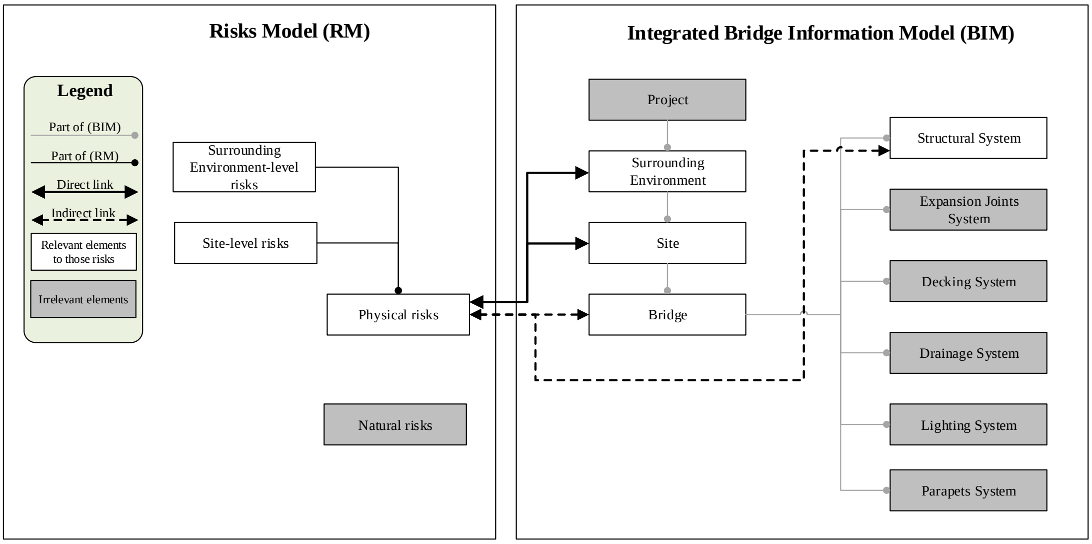



Figure 25. Linkage between physical risks and BIM

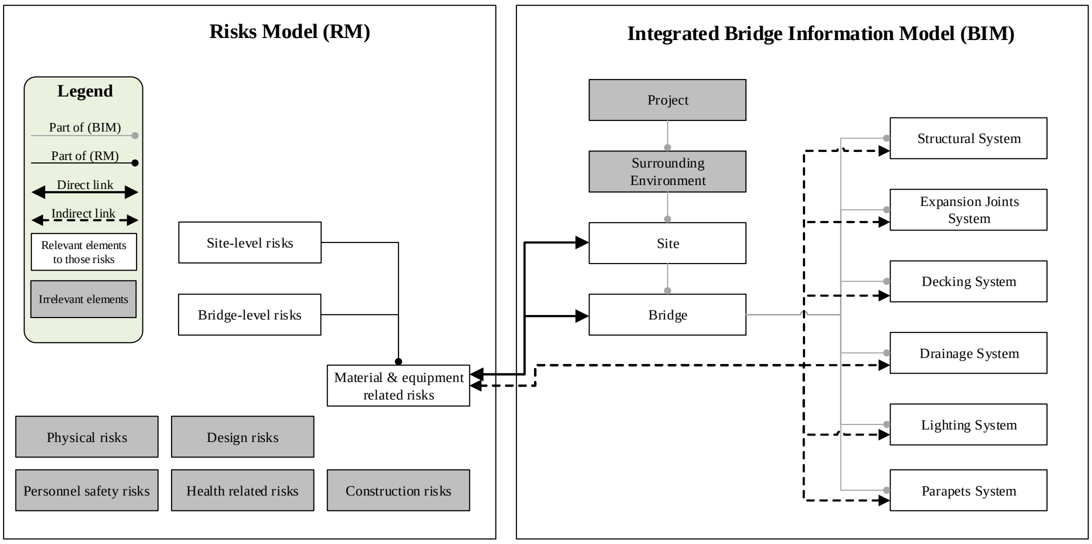



Figure 26. Linkage between material and equipment risks and BIM

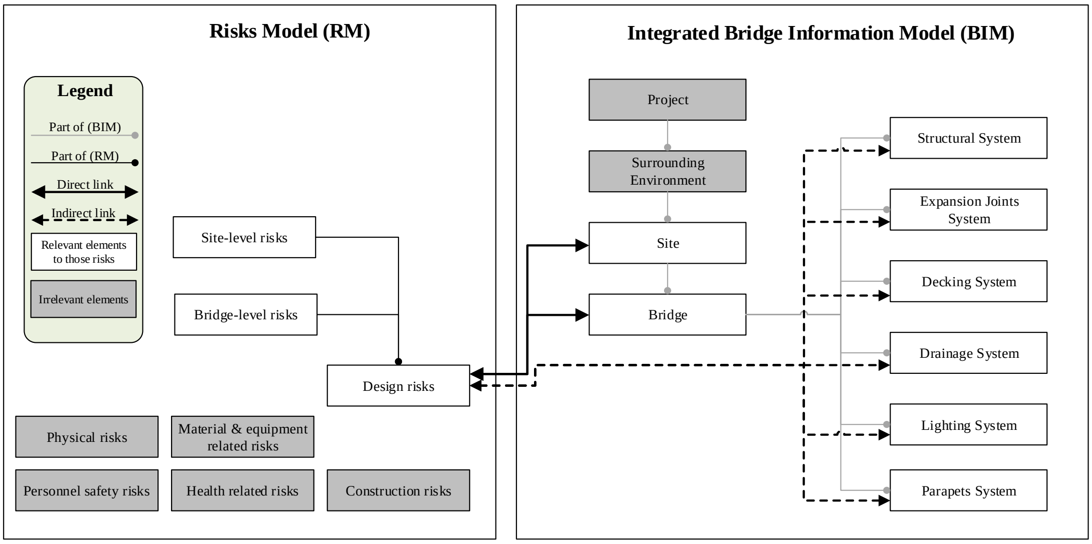



Figure 27. Linkage between design risks and BIM

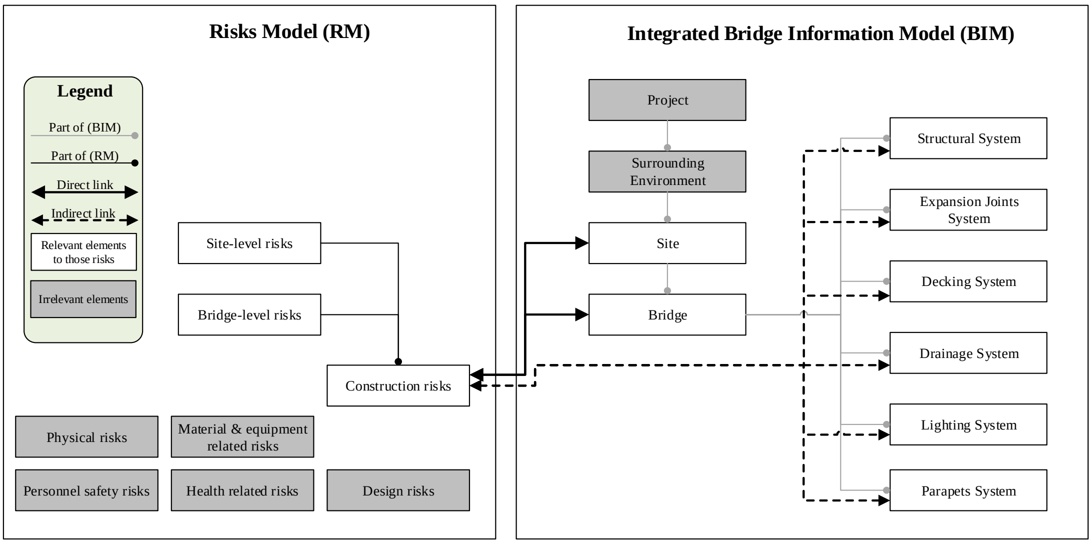



Figure 28. Linkage between construction risks and BIM

## 6.3 The RBM

As introduced in Section 5.3, a pre-defined RBS for bridge projects has been developed and the next step is to link the RBS to the WBS of a construction project to establish the RBM.

As for each construction project, the design, construction, project deliveries, and organisation structures are very different and there is no standardised method for defining its WBS (PMI, 2004), the formats and contents of WBS may vary from project to project and it is impossible to develop a standard WBS in this paper. Generally, there are two different ways for considering the development and use of a WBS for a particular project. Firstly, the WBS can be a breakdown representation for the different phases of the project lifecycle and it will help clients and project manager to have a global understanding of the potential risks during the whole project life cycle. Secondly, WBS can be used as a "construction schedule plan" for breaking down complex construction processes into small tasks. This paper employs WBS to form the latter perspective to prove the concept of applying RBS and WBS into 4D BIM.

## 6.4 Visualisation of risks in 4D BIM

As discussed in Section 6.1, WBS providing the time- and schedule- related information can be connected to 3D BIM for establishing the 4D BIM. Sections 5.3, 6.2 and 6.3 have demonstrated that risk information can be linked to BIM and WBS. Consequently, from a theoretical view, the identified risk information can be visualised in the 4D simulation according to the risks' affected WBS tasks. The benefit of visualisation of risks in 4D BIM can facilitate the global understanding how the identified risks may affect the dynamic process of a project and provide the possibility of review and check of the identified risk data in 4D environment.

## 6.5 System development and implementation

This section firstly introduces the architecture and choice of development environment of the proposed system, it then explains the system's core components and operation process. The introduction is followed by the demonstration of the practical development process and an illustrative example. The final sub-section addresses the benefits of this proposed system.

### 6.5.1 Prototype architecture and choice of development environment

In order to implement and test the feasibility of the proposed method, a BIM-based risk management system prototype has been developed through an open Application Programming Interface (API) which allows end users to manipulate the model, access and exchange data from BIM-based tools. The architecture of the proposed prototype tool is presented in Figure 29. Specifically, Autodesk Navisworks (Autodesk, 2016a) was chosen as the main BIM tool for implementation, which is a BIM-based project review software and provides functions such as model integration, clash detection, quantity take-off, site planning and 4D simulation. The main User Interface (UI) was developed as a Navisworks plugin using the C-sharp (C#) programming language based on the Microsoft .NET Framework. Meanwhile, the UI is linked to and could access and manipulate the data in a Microsoft SQL Server database through the Structural Query Language (SQL).

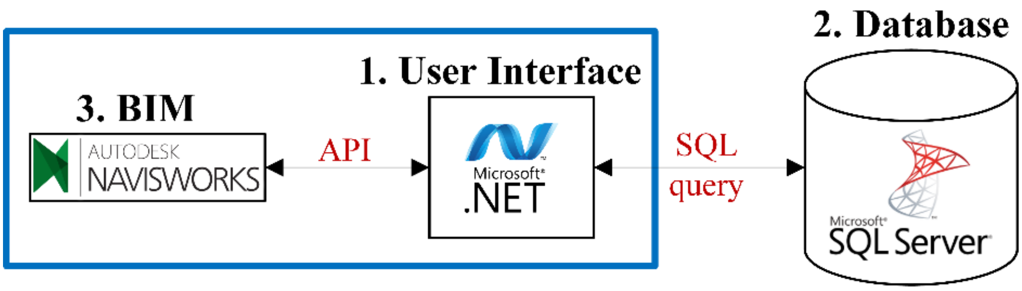



Figure 29. Architecture of risk management system

The reasons and benefits of choosing those tools to develop the system prototype are summarised as follows:

- Navisworks is one of the most popular 4D BIM-based tools which support a wide range of model formats from other mainstream BIM-based software including, for example, Industry Foundation Classes (IFC), Revit, Microstation, Rhino, ArchiCAD, CATIA. As the proposed system is embedded into Navisworks as a plugin, the system could be very easily implemented and tested using any model formats supported by Navisworks.

- Another benefit of using Navisworks is that the WBS information (i.e. schedule and time related 4D information) has already been embedded into the Timeliner module of Navisworks. This means there is no need to develop the WBS module additionally for the system prototype because the core purpose of the prototype development is to validate the proposed theory.

- The latest Navisworks API is based on the Microsoft .NET Framework, which is a software development platform for Microsoft Windows and allows users to access and manipulate databases, and develop applications for mobile devices, web pages and desktop. And C# is one of the main programming languages to support .NET Framework. Therefore the .NET Framework is the ideal platform to develop such a system prototype.

Although only Navisworks has been tested in this research, the proposed theory and methodology could be also applied to other BIM-based tools, which provide API facilities.

### 6.5.2 Core components and operation process of the system

The core components and operation process of the proposed BIM-based risk management system are presented in Figure 30. The system is made up of three major components: BIM, Plugin and Database. Here BIM is a representation of the 3D/4D model in the Navisworks environment. The concept of the triangular model has been built into the Plugin which provides the main UI to help end users to identify and manage risks, and save risk data to the database, and visualise identified risks in 4D BIM. The Database is used for storing identified risk information.

The operation process of this system can be explained as follows.

- When the 3D/4D model is ready in Navisworks, the plugin reads and loads the 4D time and schedule information to establish the WBS.

- According to the 3D/4D model and project-specific information, the next step is to combine the use of the pre-defined RBS and obtained WBS for risk identification, and at the same time identify the affected LOCs, Systems or Objects of BIM. The related data is then saved to the database and linked to BIM.

- The Visualise Risk in 4D module allows that the identified risk information can be obtained from the database and added into 4D BIM for visualisation.

- In addition, the Query & Manage module can be used for querying the risk data and highlight and visualise the risks in the 3D BIM environment.

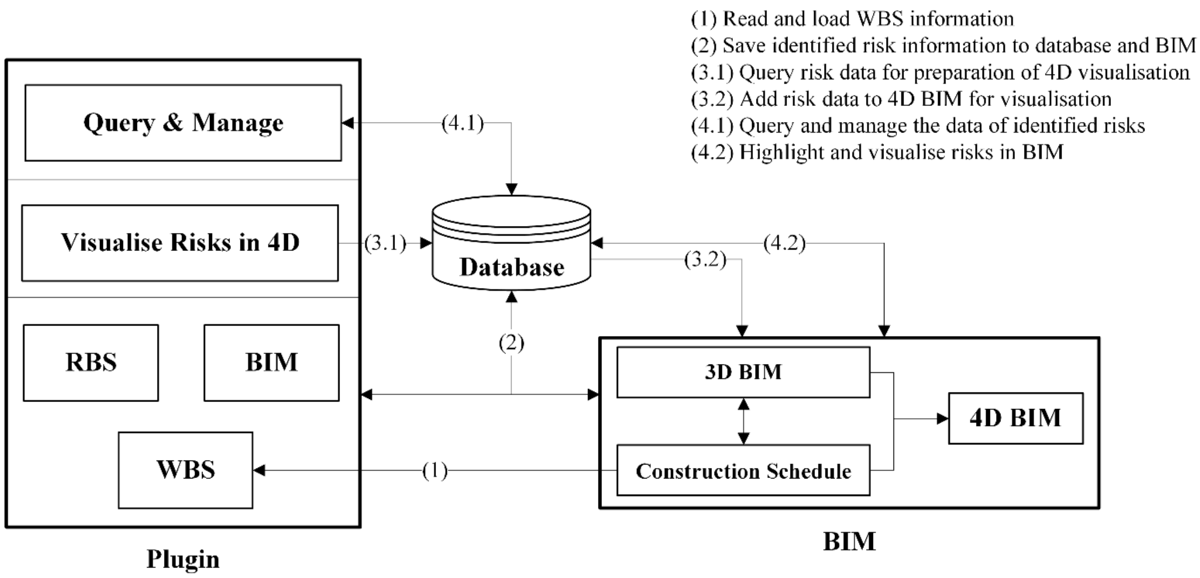



Figure 30. Core components and operation process of the system

### 6.5.3 Prototype development and implementation

The snapshot of the proposed system's UI is presented in Figure 31. This system is built in the Navisworks environment and consists of a plugin, 3D graphical model and Timeliner module. The Timeliner is the 4D module in Navisworks and can add time and schedule related information into a 3D model for schedule simulation. The system provides a helpful guide to assist end users to identify risks and enables the storage of the identified risk information into the database, and links the related information to BIM. In addition, risk information could be visualised in a 4D simulation in Navisworks for better communication. The details of the system implementation are described in the rest of this sub-section and an illustrative example is presented in Section 6.6.

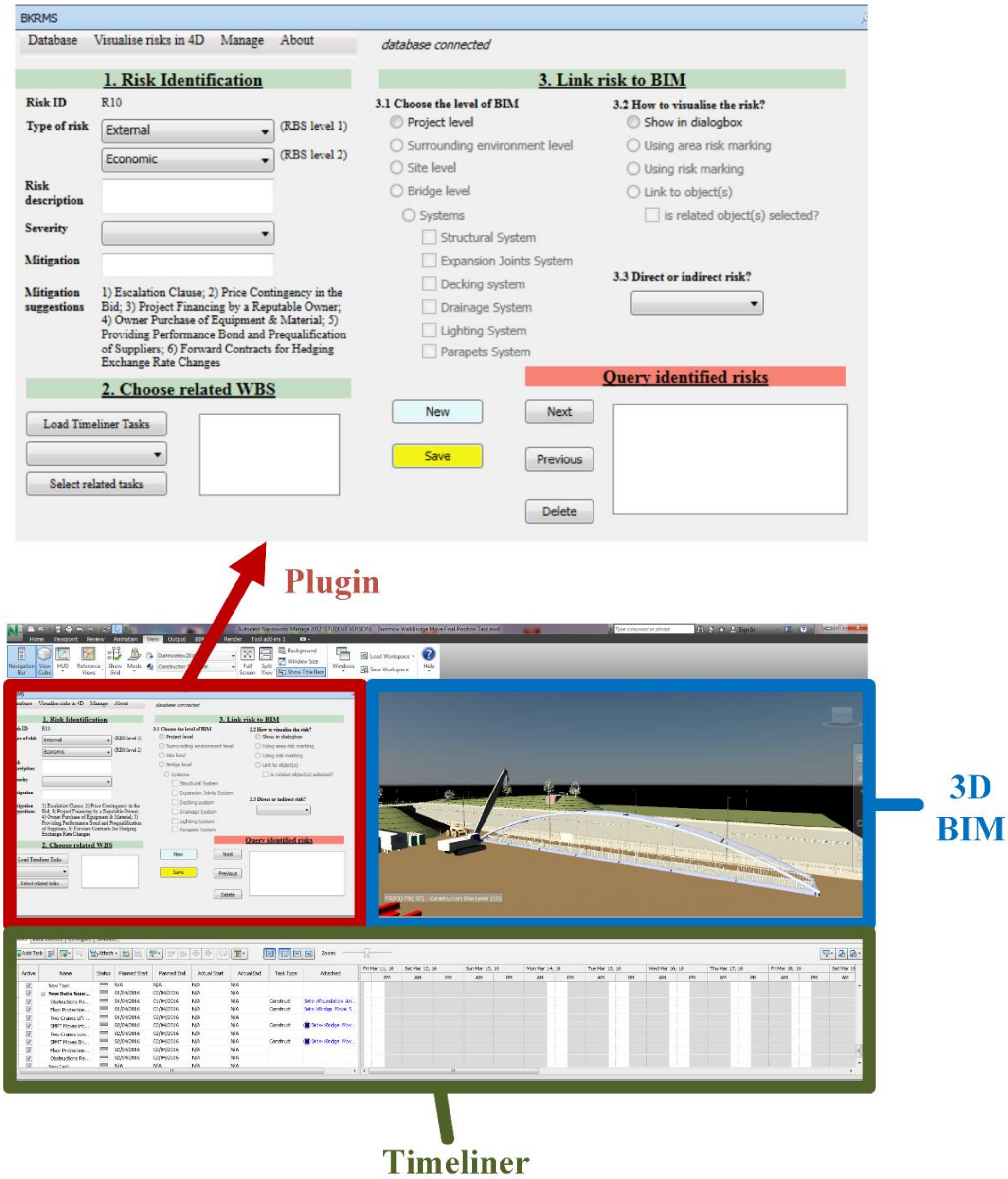



Figure 31. Snapshot of UI in Navisworks

As shown in Figure 31, there are four items in the UI's menu, i.e. "Database", "Visualise risks in 4D", "Manage", and "About". Specifically, "Database" is used to connect or disconnect the database and BIM; "Visualise risks in 4D" is designed for visualising and simulating all identified risk information in the 4D simulation of Navisworks; "Manage" enables users to review all identified risks by groups in tabular forms according to the LOCs of BIM; and "About" is to show the proposed system's information, e.g. version, developers, copyright.

#### *6.5.3.1 Documenting a risk record*

To save an identified risk record, there are generally three steps, i.e. "Risk Identification", "Choose tasks", and "Link risk to BIM".

- In the first step, Risk ID is generated automatically according to the existing risk records in the database to ensure each risk has a unique ID. Then the user needs to choose the "Type of risk" which is a simplified representation of the RBS and enter the customised description about the risk into the field "Risk description". After this, "Severity" provides three options (i.e. High, Medium, and Low) and the mitigation information could be filled into the "Mitigation" field. In addition, "Mitigation suggestions" is linked to a pre-defined table in the database and could show suggested information according to the chosen type of risk.

- In the second step, the Timeliner task information is read, loaded and the user can choose one or more tasks which are in association with the identified risk. This allows the system to generate a link between the particular risk and its affected tasks.

- The third step allows users to choose the particular LOCs and systems of BIM associated with the risk and select the most appropriate methods to visualise and link the risk information to BIM. To implement the linkage rules described in Section 4.2, a mechanism of disabling the irrelevant options was designed for the system. When a certain type of risk is selected, the system will enable the options of LOCs and systems which are associated with the chosen type of risk and disable the rest of the options according to the linkage rules. For example, if a “Design” risk is selected, only Site level, Bridge Level and Systems are available to choose (Figure 32). However, although the need for separation of BIM to meet different purposes of use has been recognised by some existing research (Kiviniemi, 2005, Fischer and Kam, 2002, Haymaker *et al.*, 2003), the total number of studies addressing the problem of separation of BIM is still very small and most BIM-based tools, including Navisworks, fail to support linking information to the LOCs and systems of BIM. Therefore this paper employs an alternative solution for visualising and linking risk information to the Navisworks model, which is summarised in Table 4. Specifically, for those risks which affect the Project level or are not easily linked to any objects in Navisworks (e.g. political and financial risks), information is just stored in the database and the user can review them in Navisworks through a tabular form. For those risks which influence the Surrounding Environment, Site or Bridge levels but cannot be easily linked to existing objects in Navisworks (e.g. natural risks, personnel health and safety risks), two different risk markings as shown in Figure 33 are appended to Navisworks as objects to store the related risk information and these can be visualised in the 4D simulation. The purpose of the “area marking” is to warn that the risk might influence a certain area. In addition, for those risks that are influencing the bridge level or systems and can be associated with particular model objects (e.g. structural risks), risk data is directly linked to the related objects which are visualised in a highlighted colour in the 4D simulation.

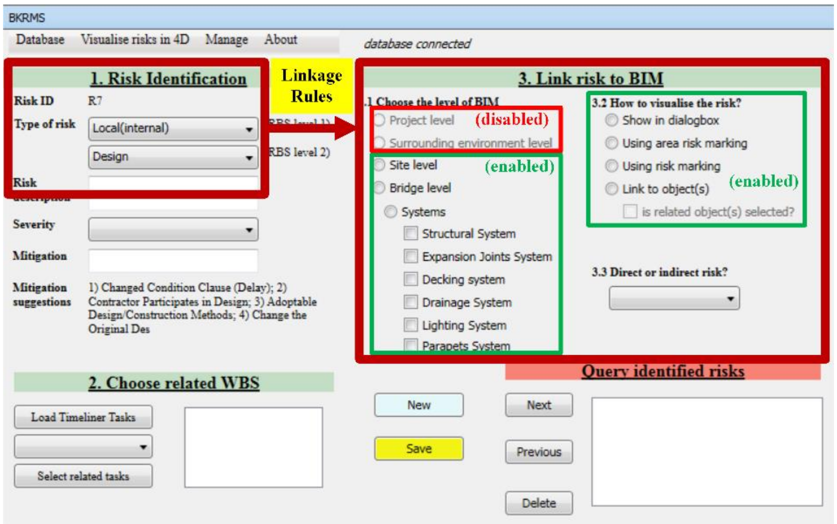
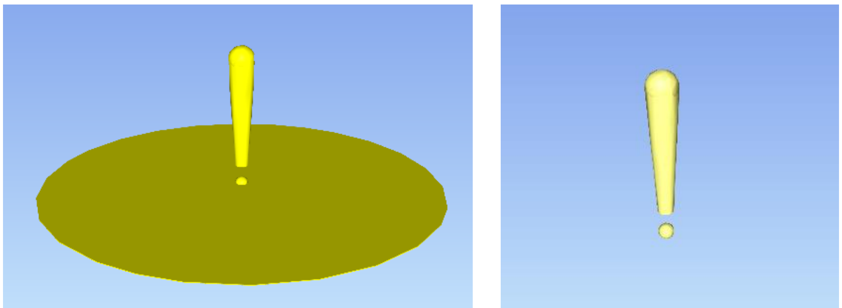



Figure 32. Example of implementing linkage rules

Table 4. Summary of methods of visualising and linking risk to BIM

| RBS | | Method of visualising and linking risk to BIM | | | |
|---|---|---|---|---|---|
| Level 1 | Level 2 | Tabular Form | Area Marking | Risk Marking | Existing objects |
| External | Political | √ | x | x | x |
| | Economic | √ | x | x | x |
| | Social & cultural | √ | x | x | x |
| Global (Internal) | Time | √ | x | x | x |
| | Organizational | √ | √ | √ | √ |
| | Quality | √ | √ | √ | √ |
| | Financial | √ | √ | √ | x |
| | Contractual & legal | √ | √ | √ | x |
| | Natural | x | √ | √ | x |
| Local (Internal) | Physical | √ | √ | √ | √ |
| | Personnel health | x | √ | √ | √ |
| | Personnel safety | x | √ | √ | √ |
| | Material & equipment | √ | √ | √ | √ |
| | Construction | √ | √ | √ | √ |
| | Design | √ | √ | √ | √ |
| | Structural | x | √ | √ | √ |



a) area marking b) regular marking

Figure 33. Risk markings used in this study

Although the proposed system has pre-defined a number of rules for visualising risks and linking risk data to Navisworks model as summarised in Table 4, there is no 100% correct answer on the choice of visualisation methods for risks and people can choose an appropriate method according to their own preference. For example, it could be assumed that a Rough Terrain Crane is working on the construction site and a risk might exist that construction workers may be hit by the crane. Users could either link that risk to the crane object directly or use an external marking.

#### *6.5.3.2 Visualise Risks in 4D BIM*

The strategy of visualisation of identified risks in 4D BIM is explained in this sub-section. The key solution used in this paper is to query the identified risks and their information from the database, and add them into the Timeliner Module automatically as part of the 4D schedule. Specifically, if an identified risk has been determined to use Tabular Form to visualise it in BIM, all these risks and their related information will be summarised in a tabular form when the module of Visualise risks in 4D begins. If a risk is linked to markings or objects, the markings and objects will be highlighted during the time period which is the same as the risk's affected WBS tasks.

#### 6.5.3.3 *Query and management of identified risks*

The purpose of the query and management module is to review and check the identified risks and how they may impact on the project. By using the “Previous” and “Next” buttons, the tool allows a review of all identified risks, and their related information will be shown in a textbox. Meanwhile, if the risk is linked to markings or objects, the related markings or objects will be highlighted. An example is presented in Figure 34.

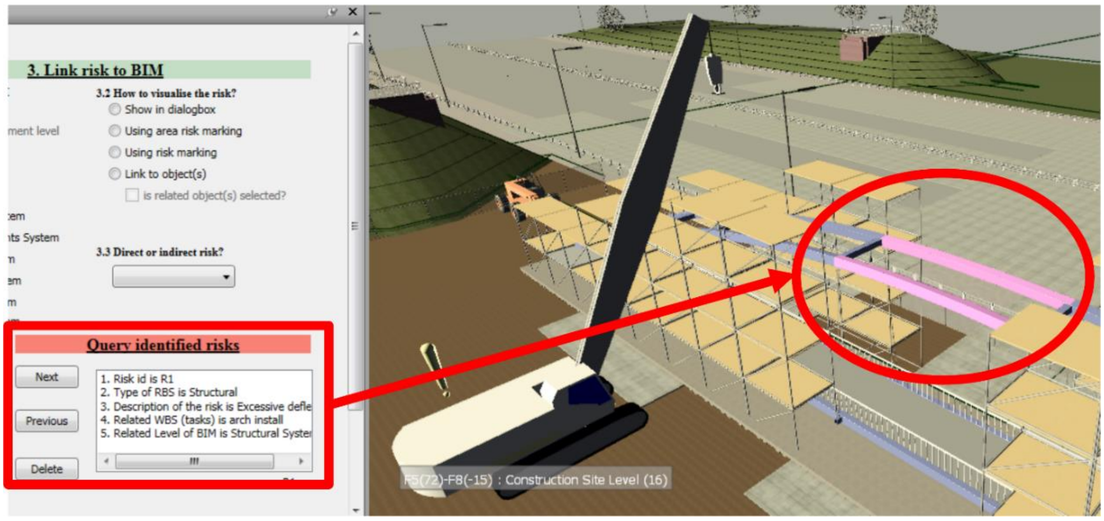



Figure 34. An example of highlighting risks in BIM

## 6.6 Illustrative case study

The purpose of this sub-section is to use a test model to illustrate the implementation process of the proposed risk management system. This study chose a standard steel footbridge in the UK for implementation and the overall implementation process is presented in Figure 35.

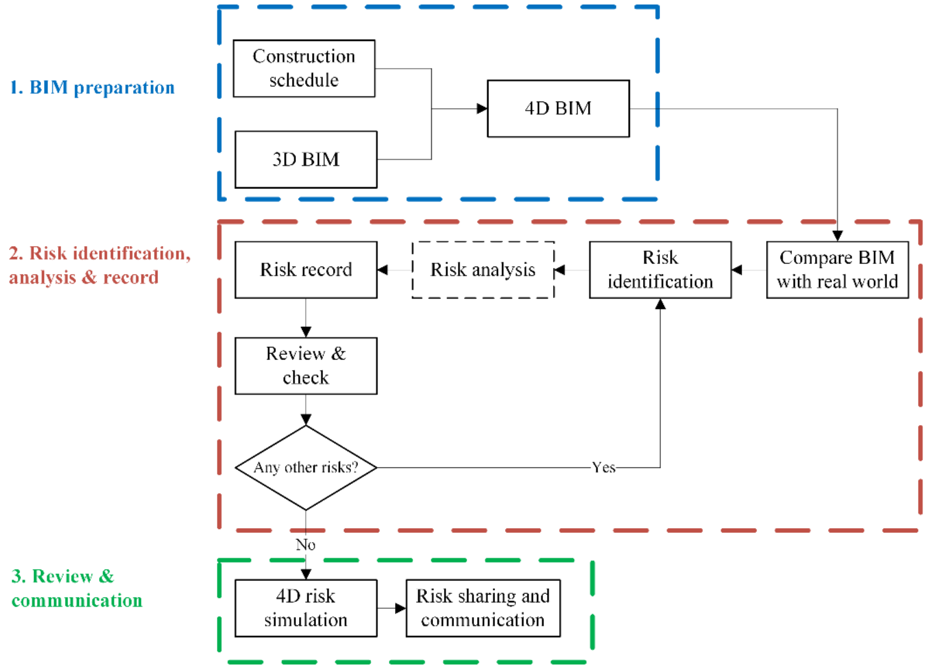



Figure 35. System implementation workflow

- **BIM preparation stage.** In this initial stage, a 3D graphical model was firstly created by Autodesk Revit by following the bridge's original 2D drawings. The Revit model was then exported to Navisworks as a 3D BIM and the construction schedule information of the bridge was brought into Navisworks to develop the 4D model. The construction of this bridge consists of three main processes: 1) on-site fabrication of the arch and deck, 2) construction of the bridge abutments, and 3) move and installation of the bridge structure. Snapshots of the simulation of the bridge construction in Navisworks are shown in Figure 36.

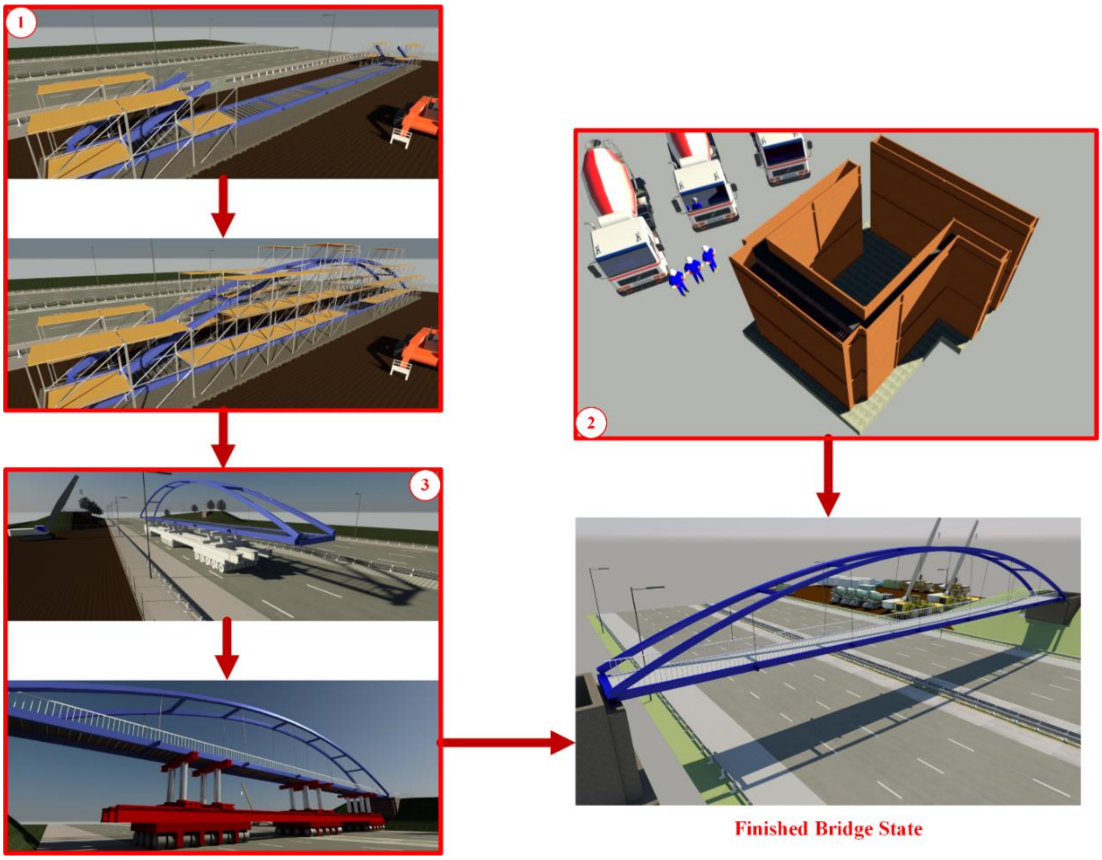



Figure 36. Simulation of the bridge construction

- **Risk identification, analysis and record**. With the 3D/4D BIM, the project, its surroundings, and construction schedule were reviewed in the computer-based virtual environment. The next step was to connect and compare the 3D/4D models with the real environment and situations to identify and analyse any potential risks. The proposed plugin tool was used for risk identification and information documentation. Then each risk was analysed and the best mitigation measures for that risk were investigated using the most appropriate methods, e.g. knowledge and experience, structural simulation, mathematical analysis. In this process, each risk's information was stored in the database and a link was established between the risk data and BIM. All identified risks were reviewed through the Query and Manage module to check if the risks and their related WBS tasks and BIM were correct. By taking the on-site fabrication of the arch and deck as an example, three risks relating to time, personnel safety and the structure respectively were identified in this process according to author's experience and were summarised in Table 5.

Table 5. Three identified risks visualised by different methods

| No. | RBS | Risk description | Severity | Mitigation | Level of BIM | Visualisation method |
|---|---|---|---|---|---|---|
| 1 | Time | Mechanical failure of the construction plant | High | Have standby plant ready to take over | Site | Tabular form |
| 2 | Personnel Safety | Workers may be hit by the moving crane | Medium | Safety education before implementation | Site | Marking |
| 3 | Structural | Excessive deflection | Medium | Strengthen monitoring and control when implementation | Bridge | Objects |

- **Review and communication**. After all foreseeable risks are identified, the module of Visualise Risks in 4D was used to review and check the identified risks in the 4D BIM environment, which can be used for risk communication both internally and externally. Using the three identified risks in the on-site fabrication process as an example, the system firstly popped up a new window showing the time-related risk in a "Tabular Form" and the other two risks were highlighted during the time period which is the same as the risks' affected WBS tasks in the 4D simulation (see Figure 37).

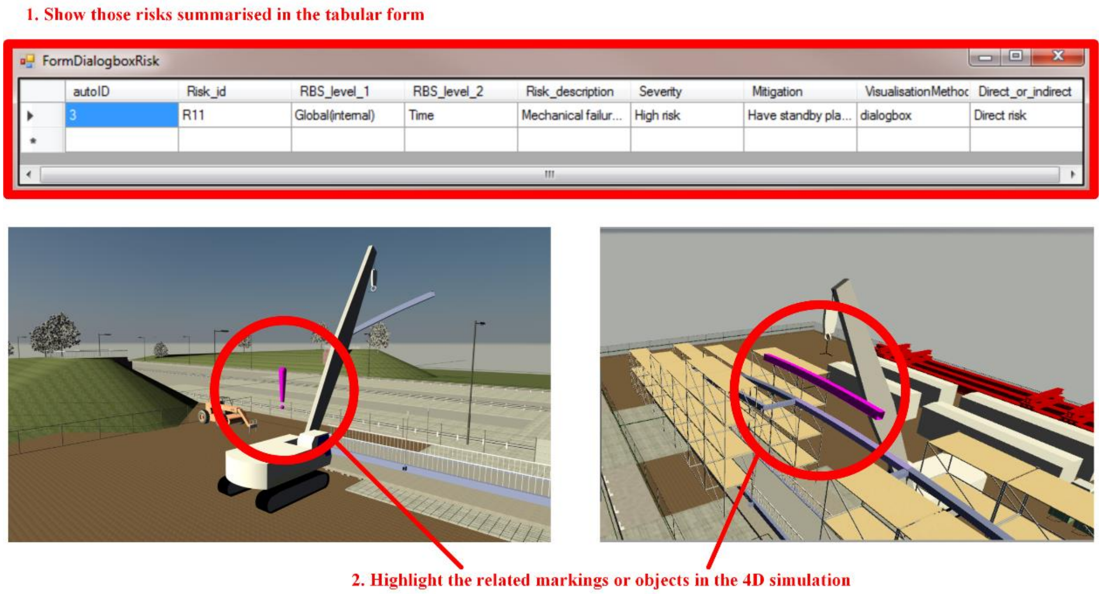



Figure 37. Visualise risks in 4D BIM

## 6.7 Benefits of system

Results show that traditional techniques such as RBS and RBM can be integrated into BIM as a whole for project risk management and the advantages of both traditional techniques and BIM could be combined. On one hand, pre-defined RBM and RBS could improve the risk understanding throughout the whole project and be used for risk identification, and information classification and management. On the other hand, BIM could not only act as an enriched data model which manages both information inside the model and risk information stored in the database but also facilitates the risk identification and communication through 3D visualisation and 4D construction sequence simulation.

It has also been demonstrated in this paper that an active linkage can be established between the risk information stored in an external database and BIM. Although only Navisworks is tested in this study, there is a growing number of BIM-based tools providing API and it could be seen in the future that a similar linkage can be also applied to other BIM-based tools and a collaborative risk management system could then be developed to support a multi-platform working environment.

This page is intentionally left blank

# Chapter 7. RETRIEVING SIMILAR RISK CASES FROM THE CBRL USING NLP TECHNIQUES

In order to improve the efficiency and performance of risk case retrieval, this chapter proposes an approach to integrate two NLP techniques, i.e. VSM and semantic query expansion, into CBR for risk case retrieval and outlines a new framework for the risk case retrieval system. A prototype system is developed with the Python programming language to support the implementation of the proposed method.

The rest of this chapter is organised as follows. The system architecture of the proposed Risk Case Retrieval System (RCRS) is presented in Section 7.1. The three main modules of this system, i.e. Risk Case Processing, Query Operation, and Retrieval Application, are illustrated in Sections 7.2, 7.3 and 7.4, respectively. In Section 7.5, a prototype system developed with Python and an illustrated example are described. Finally, the validation of this proposed system is addressed in Section 7.6.

## 7.1 System architecture of the Risk Case Retrieval System

The system architecture of the proposed RCRS is illustrated in Figure 38. The system consists of three major modules, i.e. (1) Risk case processing, (2) Query operation, and (3) Retrieval application. Firstly, the risk case processing module automatically extracts the textual information from a targeted collection of risk cases. It processes the collected textual information by a defined Sequence of Actions (SoA), i.e. tokenisation, converting all words into lowercase, lemmatisation, and removing stop words to establish a risk case content corpus. The SoA is a general approach in current NLP for processing textual documents (Jurafsky and Martin, 2009). Secondly, the query operation module reads and processes the given query by SoA. The processed query is prior scanned to match its expansion of related words in the pre-defined risk-related lexicon. The terms not found in the pre-defined risk-related lexicon are expanded by using synonyms in WordNet. Then the system scans the terms in both the original query and the expanded query, and removes those terms that do not exist in the risk case content corpus. Thirdly, the retrieval application module combines the queries and risk case corpus together and performs the query-document similarity calculations. After this, the system ranks all documents according to their similarity scores and finally returns, for example, the top 10 documents to the users.

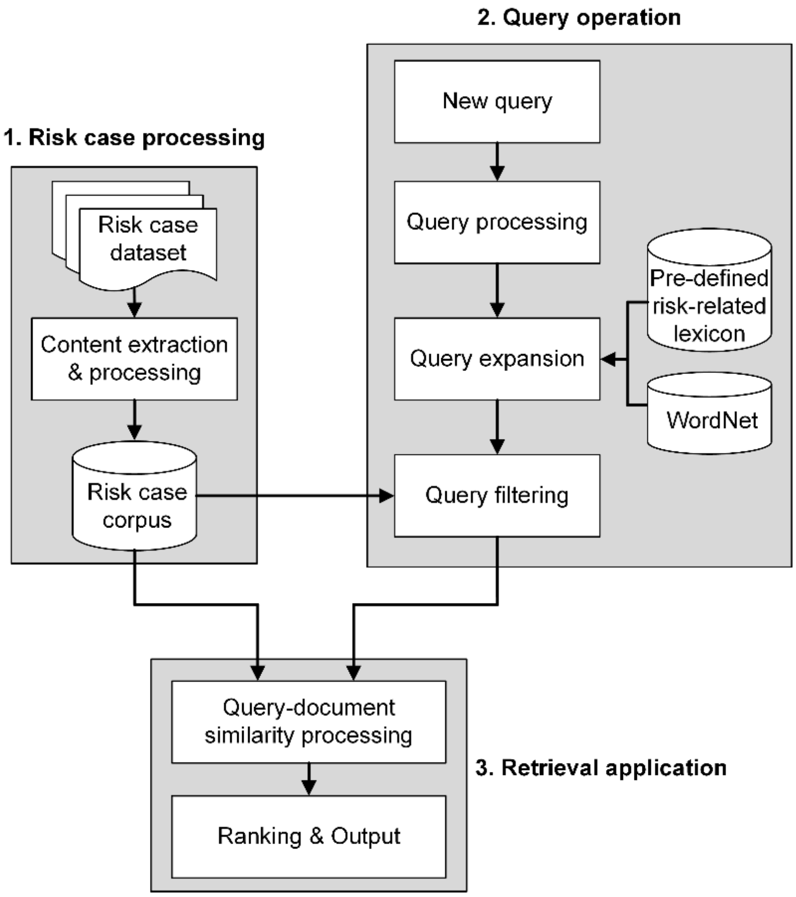



Figure 38. System architecture of RCRS

## 7.2 Risk case processing workflow

The first step in the risk case processing module is to collect risk cases through a web search method. In total 590 risk cases have been collected from major organisational and governmental construction accident databases, which is discussed in Section 3.3.4.

The second step is to extract the textural information from the collected reports and process them to be a risk case content corpus, which goes through the following processes:

**Tokenisation**: this is a process of chopping a document up into pieces (known as 'tokens') and discarding certain characters, such as punctuation (Manning *et al.*, 2008). An example is illustrated in Figure 39.

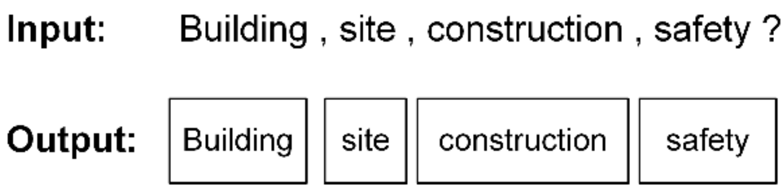



Figure 39. An example of tokenisation

**Converting words into lowercase**: this is a simple task to convert tokens into lowercase, which could improve the search results (Manning *et al.*, 2008). For instance, the term "Building" is converted to be "building".

**Lemmatisation**: it "usually refers to doing things properly with the use of a vocabulary and morphological analysis of words, normally aiming to remove inflectional endings only and to return the base or dictionary form of a word, which is known as the lemma" (Manning *et al.*, 2008). For example, the base form "walk" may appear as “walk”, “walked”, “walks”, or “walking” in the main text, and the process of lemmatisation is to convert those words to their base forms.

**Stop words removal**: stop words are those extremely common words which have little value in helping match documents (Manning *et al.*, 2008). Removal of those meaningless words could largely reduce the size of the collection and improve the retrieval efficiency. The stop words used in this study are presented in Table 6 which consists of two sub lists. The first list of stop words is identified by the Natural Language Toolkit (NLTK) (NLTK, 2016), which is a suite of libraries and programs for symbolic and statistical NLP for English written in the Python programming language (Perkins, 2014). The second list comes from a manual selection from the top 100 words that have the most occurrences in the risk case content corpus but are identified with little value. For example, ‘fig 1’ has extremely high occurrences in the whole risk case collection but its tokens (i.e. ‘fig’ and ‘1’) are of little help to the risk case retrieval. Because there are still some limitations in current NLP techniques (Hsu, 2013), some meaningless words are produced after Tokenisation, e.g. the symbol underline and the letter “j”. Removal of these manually selected meaningless words with the highest numbers of occurrence could effectively reduce the size of data and this method has been adopted in some previous studies, e.g. (Fan and Li, 2013).

**Establishing the risk case corpus**: corpus in the NLP context refers to a large collection of texts (Jurafsky and Martin, 2009) and this process is to combine the processed textual information into a corpus for further use in the query operation and retrieval application.

Table 6. Stop words used in this research

| **Stop words identified by NLTK** | | | | | **Manually selected stop words** |
|---|---|---|---|---|---|
| the | his | off | him | about | number |
| couldn | ain | with | doesn | re | 15 |
| shan | were | m | an | our | 20 |
| between | very | but | who | both | could |
| any | there | own | was | he | 14 |
| himself | while | for | during | this | 16 |
| a | hers | is | once | until | f |
| at | over | too | other | am | b |
| after | myself | just | ll | no | 12 |
| will | then | i | again | mightn | fig |
| ma | it | wasn | being | hadn | 11 |
| its | against | by | yourselves | through | _ |
| o | these | how | not | because | 0 |
| what | ve | them | can | out | e |
| don | her | in | up | if | would |
| does | are | from | on | mustn | also |
| didn | wouldn | under | having | below | j |
| most | theirs | down | of | shouldn | may |
| same | whom | only | each | aren | r |
| their | s | where | y | do | 10 |
| and | you | all | nor | isn | 9 |
| did | now | haven | herself | have | l |
| your | as | yourself | t | yours | c |
| which | won | into | should | above | 7 |
| further | itself | been | she | me | 1 |
| few | needn | d | ours | my | 6 |
| to | or | such | weren | here | 5 |
| so | why | had | than | more | 4 |
| they | before | some | that | themselves | 3 |
| those | be | we | hasn | | 2 |
| when | doing | ourselves | has | | |

## 7.3 Query operation process

A basic semantic similarity problem is often observed that terms of the original query are different to the ones used in the documents in describing the same semantics (Gong *et al.*, 2005). To deal with mismatching problem, a promising solution is to use query expansion (Gao *et al.*, 2015, Colace *et al.*, 2015, Gong *et al.*, 2005). In definition, query expansion is a process of reformulating or expanding a seed query using semantically related words (e.g. hyponyms, synonyms) to improve the retrieval performance of IR systems (Vechtomova and Wang, 2006). Many web IR efforts have adopted this approach and a common way is to extract the semantically related words from WordNet (Fellbaum, 1998, Gong *et al.*, 2005, Snasel *et al.*, 2005), a lexical database for the English language.

Because the collected risk cases are in different styles of expression by different individuals or organisations, the above problem also commonly exists in the risk case database, e.g. “structural failure” and “structure collapse”. Therefore this research integrates query expansion into the RCRS for this mismatching problem. However, WordNet is a relatively complete lexical database for the whole English environment and contains too much data which is not useful for the risk case retrieval context. For example, the synonyms of “failure” are “nonstarter”, “loser” and “unsuccessful person” which are not related to project risk management. In addition, no such dictionary or database has been found for defining the semantically related words in a risk management context. Hence, this study established a small risk-related lexicon to overcome this limitation and combines the use of this risk-related lexicon and WordNet.

The pre-defined risk-related lexicon is a dictionary consisting of 107 key words, which are most commonly used in the risk management context, and their expansion suggestions. An example is shown in Figure 40 and the full list of lexicon is presented in Appendix B. To develop the lexicon, three major steps were used. Firstly, the 107 key words (e.g. “building”, “risk”, “collapse”, “change”, “safety”) were manually selected from all risk factors in a risk database established in Section 5.2. The second step performed a deep learning approach to find out the most related words ( i.e. “Values” in Figure 40) of 107 key words by using Word2vec (Mikolov *et al.*, 2013a, Mikolov *et al.*, 2013b), a deep learning algorithm developed by a research group led by Tomas Mikolov at Google. Word2vec is an unsupervised learning tool for obtaining vector representations for words and could be used for finding the most similar or related words in an N-dimensional vector environment. The collected 590 risk cases were initially used for training but it was quickly realised the size of data was so small that the performance of calculation is not as good as the author expected. Then, the free and open Wikipedia content database (Wikipedia, 2016) is used as a supplement for calculating the most similar words. In the third step, similar words calculated by using both risk case content corpus and Wikipedia content database are gathered together and a manual selection process based on knowledge and experience is conducted to delete words that are not related to the risk management context.

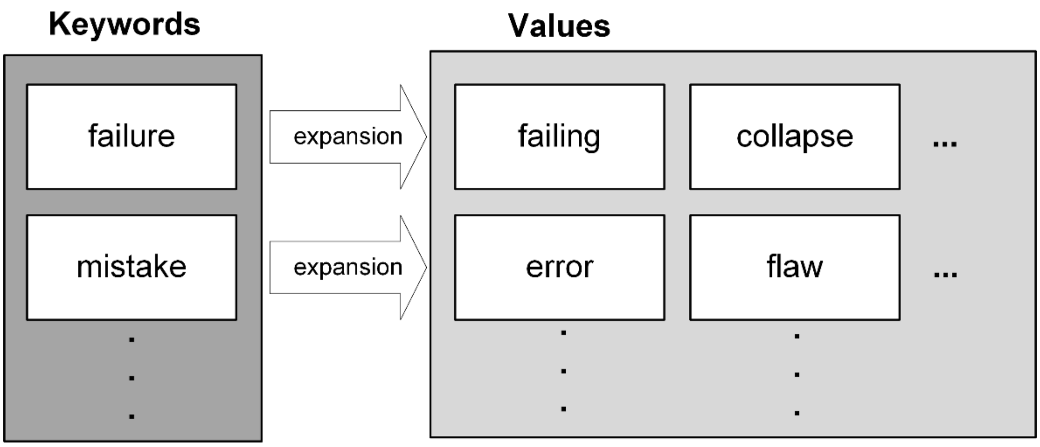



Figure 40. Example of risk-related lexicon

The work flow of query expansion is shown in Figure 41. Specifically, a new query is firstly read and processed by SoA. Secondly the processed query terms are prior scanned to match its expansion of related words in the pre-defined risk-related lexicon. If any terms are not found in the pre-defined risk-related lexicon, they are expanded by using synonyms in WordNet. After this, there are two queries, i.e. original query, expanded query. With the observation that original query could mostly reflect a user's need for case retrieval, this research keeps the original query and expanded query as two separate queries. Thirdly, the system scans the terms in both original query and expanded query, and removes terms that do not exist in the risk case content corpus. Lastly, the system outputs both refined original query and expanded query for further use in retrieval application.

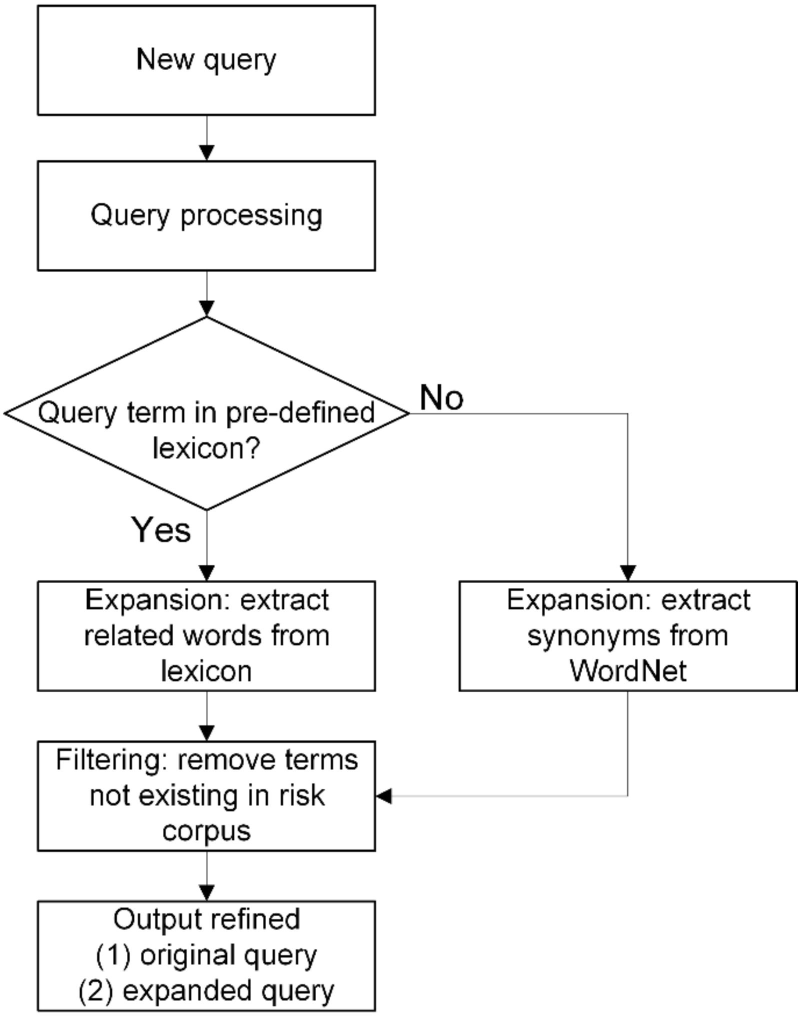



Figure 41. Work flow of query expansion

## 7.4 Retrieval application process

### 7.4.1 The classical VSM

In definition, the VSM is an algebraic model for representing textual documents as vectors of identifiers and assigning non-binary weights to index terms in queries and in documents, which is broadly used to compute the degree of similarity between each document and the query (Baeza-Yates and Ribeiro-Neto, 2011, Salton *et al.*, 1975, Sparck Jones, 1972). The classical VSM is described as follows (Baeza-Yates and Ribeiro-Neto, 2011):

Query $q$ and document $d_j$ can be represented as t-dimensional vectors, as shown in Equations (1) and (2). For the vector model, t is the total number of index terms and each dimension corresponds to a separate index term. The elements $w_{i,j}$ in each vector is the weight associated with a term-document pair $(k_i, d_j)$ and $w_{i,j} \geq 0$.

$$\vec{q} = (w_{1,q}, w_{2,q}, \dots, w_{t,q}) \tag{1}$$

$$\overrightarrow{d}_j = (w_{1,j}, w_{2,j}, \dots, w_{t,j}) \tag{2}$$

In the classical VSM, $w_{i,j}$ is known as the Term Frequency-Inverse Document Frequency (TF-IDF) weight. If the weight vector model for a document $d_j$ is $\overrightarrow{d_j}$, the document's TF-IDF weights can be quantified as:

$$w_{\mathrm{i,j}} = (1 + \log f_{i,j}) \times \log\left(\frac{N}{n_i}\right) \tag{3}$$

where $f_{i,j}$ is the frequency of index term $k_i$ in the document, $N$ is the total number of documents in the document set, and $n_i$ is the number of documents containing the term $k_i$.

Through using the VSM and TF-IDF model, the degree of similarity $sim(d_j, q)$ between the document $d_j$ and the query $q$ can be quantified as the cosine of the angle between the vectors $\vec{d_j}$ and $\vec{q}$:

$$sim(d_j, q) = \frac{\vec{d_j} \cdot \vec{q}}{|\vec{d_j}| \times |\vec{q}|} = \frac{\sum_{i=1}^{t} w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^{t} w_{i,j}^2} \times \sqrt{\sum_{i=1}^{t} w_{i,q}^2}} \tag{4}$$

where $|\vec{d_j}|$ and $|\vec{q}|$ are the norms of the document and query vectors, and $\vec{d_j} \cdot \vec{q}$ is the inner product of the document and query vectors.

### 7.4.2 The proposed score strategy and computational process

A number of existing studies (Snasel *et al.*, 2005, De Simone and Kazakov, 2005) have validated that query expansion could effectively improve the IR performance and a common method for query expansion is to use WordNet or other lexical databases. WordNet has pre-defined the basic semantic relationships between words, e.g. hypernym, synonym, hyponym. Two recent studies (Gong *et al.*, 2005, Gong and Cheang, 2004) pointed out these different semantic relations between words for query expansion will lead to different effects on the IR performance and an easy and effective approach to distinguish their effects is to give different weighting coefficients to the expanded terms.

After considering the effect of the expanded query $q_e$, this study takes the classical VSM as a starting point and proposes the following method to compute the similarity between the query and risk case:

$$score = sim(d_j, q_o) + \lambda \times sim(d_j, q_e) \tag{5}$$

where $\lambda$ is the coefficient for the effect of $q_e$ and $0 < \lambda < 1$, and this study takes $\lambda = 0.7$.

The reasons are discussed as follows:

- The basic assumption of this study is that the original query and expanded query will cause different effects on the retrieval results. The original query by the user could mostly reflect a user's searching need for the risk case retrieval, and expanded terms using pre-defined risk-related lexicon or WordNet are more or less different with the original query in semantics. Therefore an optimal solution to distinguish the effects of the original query and the expanded query is to keep the original query and expanded query as separate operations (i.e. two queries $q_o$ and $q_e$), and allocate different coefficients for them (Gong *et al.*, 2005). The expanded query $q_e$ can be considered as an additional interpretation for the original query $q_o$. If the coefficient for $q_o$ is 1, then it is clear that the coefficient for $q_e$ should be less than 1.

- As discussed in Section 7.3, this research combines the use of a pre-defined risk-related lexicon and synonyms in WordNet as the databases for query expansion. The suggested expansion terms in the risk-related lexicon are "synonyms" of the keyword in the project risk management context. Therefore, all expanded terms can be considered similarly as "synonyms" of the original query. A previous study by Gong *et al.* (2005) tested the performance of a web IR system using the different semantic relations between words of WordNet for query expansion, and demonstrated that the optimal value of coefficient for synonyms is 0.7. Hence this study takes $\lambda$ as 0.7 for practical implementation.

The computational process is illustrated as follows. Assume there are in total $k$ risk case documents in the risk case database, a term-document weighting matrix can be constructed as shown in Figure 42, where the two queries are extended as the last two "documents". For each risk case or document, the TF-IDF weights of all terms are presented in a row. If a document contains no specific term, then this term's weight in the document is 0.

$$
\begin{array}{c|ccccccccc}
 & \boldsymbol{Doc_1} & \boldsymbol{Doc_2} & \cdots & \boldsymbol{Doc_j} & \cdots & \boldsymbol{Doc_k} & \boldsymbol{q_o} & \boldsymbol{q_e} \\
\boldsymbol{Term_1} & W_{1,1} & W_{1,2} & \cdots & W_{1,j} & \cdots & W_{1,k} & W_{1,k+1} & W_{1,k+2} \\
\boldsymbol{Term_2} & W_{2,1} & W_{2,1} & \cdots & W_{2,j} & \cdots & W_{2,k} & W_{2,k+1} & W_{2,k+2} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\boldsymbol{Term_i} & W_{i,1} & W_{i,2} & \cdots & W_{i,j} & \cdots & W_{i,k} & W_{i,k+1} & W_{i,k+2} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\boldsymbol{Term_n} & W_{n,1} & W_{n,2} & \cdots & W_{n,j} & \cdots & W_{n,k} & W_{n,k+1} & W_{n,k+2}
\end{array}
$$

Figure 42. Term-document weighting matrix

For any document $d_j$, the similarity between the query $q$ and $d_j$ can be computed as:

$$
\begin{aligned}
score &= sim(d_j, q_o) + 0.7 \times sim(d_j, q_e) \\
&= \frac{\sum_{i=1}^{n} w_{i,j} \times w_{i,k+1}}{\sqrt{\sum_{i=1}^{n} w_{i,j}^2} \times \sqrt{\sum_{i=1}^{n} w_{i,k+1}^2}} + 0.7 \times \frac{\sum_{i=1}^{n} w_{i,j} \times w_{i,k+2}}{\sqrt{\sum_{i=1}^{n} w_{i,j}^2} \times \sqrt{\sum_{i=1}^{n} w_{i,k+2}^2}}
\end{aligned} \tag{6}
$$

Due to the combined effects of $q_o$ and $q_e$, the range of overall similarity is from 0 to 1.7.

## 7.5 System development and implementation

### 7.5.1 Prototype development

In order to fully implement the proposed RCRS, a prototype was developed using the Python programming language. Although other programming languages (e.g. R, Java) could have been used, this study chose Python because:

- Python is one of most widely used object-oriented programming languages with lots of features such as free and open source, easy syntax, and good extensibility. This means a Python program is easily read and understood by others and is highly extensible.

- A number of existing tools have been designed to support Python working with NLP, e.g. NLTK (NLTK, 2016), data mining and analysis, e.g. Scikit-learn (Scikit-learn, 2016). Therefore developing the prototype using Python could build on valuable previous work and avoid repeated modelling work.

### 7.5.2 Illustrative example

The purpose of this sub-section is to use the example of "Worker Fall from Height" to illustrate the computational process of the developed prototype system. The overall computational process is presented in Figure 43.

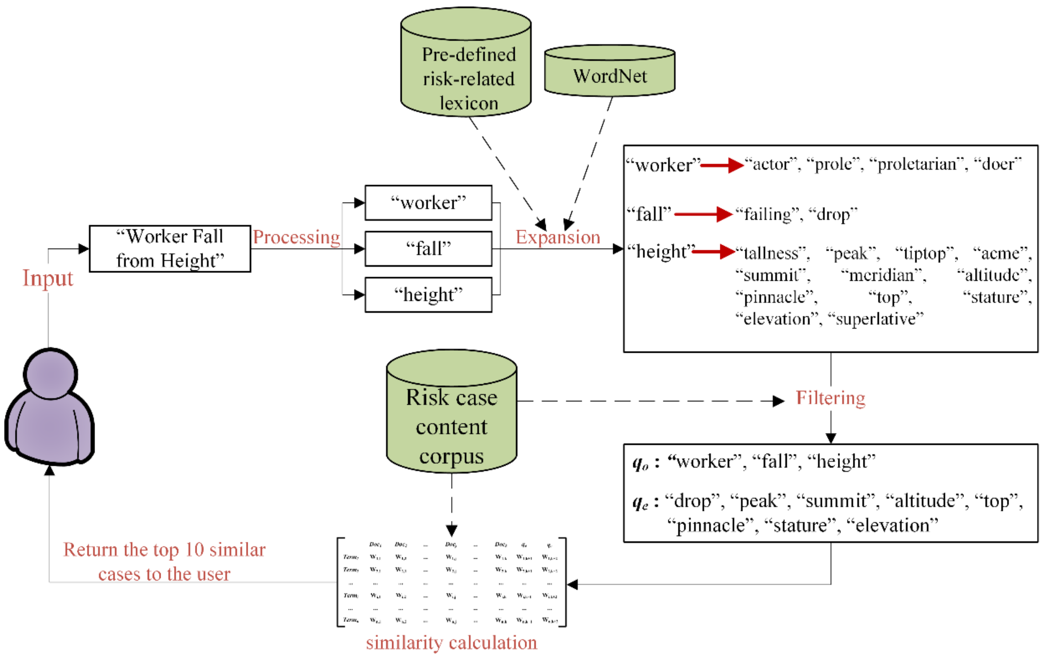



Figure 43. Computational process of retrieving "Worker Fall from Height" similar cases

The overall computational process can be described as follows:

- Before starting risk case retrieval, the system needs to read and process all the risk cases and establish a corpus for further use. As discussed in Section 3.2, a total of 590 risk cases have been collected. The system starts with extracting textual content from each risk case and getting the name list of all risk cases. After reading each case, the system processes its textual content through SoA, and saves the processed case in a temporary file. Then, all temporary files are read according to the sequence of name list and stored in a list where each risk case is a string.
- If a new query “Worker Fall from Height” is given by the user, the system first processes the query through SoA and obtains the tokens of original query, i.e. “worker”, “fall” and “height”. Then each token in the processed original query is prior scanned to find out its related words in the pre-defined lexicon. The terms not found in the pre-defined risk-related lexicon are expanded by using synonyms in WordNet. As only “fall” exists in the keyword list of pre-defined lexicon, the pre-defined lexicon is used for expansion of “fall” and the synonyms of WordNet is used for expansion of “worker” and “height”. The related words for “fall” are “falling” and “drop”. The related words for “worker” are “actor”, “prole”, “proletarian” and “doer”. And the related words for “height” are “tallness”, “peak”, “tiptop”, “acme”, “summit”, “meridian”, “altitude”, “pinnacle”, “top”, “stature”, “elevation” and “superlative”. Thirdly, the system filters the original query and expanded query by scanning the risk case content corpus and deleting those terms that do not appear in the corpus. After filtering, the original query is “worker”, “fall” and “height” and the expanded terms are “drop”, “peak”, “summit”, “altitude”, “top”, “pinnacle”, “stature” and “elevation”.

- In the third step, the processed original query and expanded query are first extended to the corpus as the last two strings in the list. Then the system performs the calculation of TF-IDF weights and establishes the corresponding term-document matrix (shown in Figure 42). Lastly the similarity between the query and each risk case is computed by using Equation (6) and the system returns the ranked top 10 similar risk cases to the end users. The result is shown in Table 7 and its snapshot in Python is presented in Figure 44.

Table 7. Top 10 similar cases of “Worker Fall from Height”

| Similarity | Title of risk case | Source | Number |
|---|---|---|---|
| 0.355807864882 | Young worker falls from third-storey balcony | WorkSafeBC | 30 |
| 0.350710609398 | Fall from roof with too much slack in lifeline | WorkSafeBC | 3 |
| 0.306337588766 | Hispanic laborer dies after falling through a second story floor opening | NIOSH | 5 |
| 0.286606375085 | Worker falls through roof insulation to concrete floor | WorkSafeBC | 27 |
| 0.282279911804 | Worker died after fall from steep-sloped roof | WorkSafeBC | 12 |
| 0.281084486537 | Worker entangled in chain falling from dismantled conveyor | WorkSafeBC | 13 |
| 0.278102714551 | Worker died after being submerged in flooded cranberry field | WorkSafeBC | 11 |
| 0.277708195414 | Workers seriously burned in flash fire | WorkSafeBC | 20 |
| 0.238392609973 | Hispanic worker falls from residential roof | NIOSH | 1 |
| 0.235168098338 | Workers fall when unsecured bin tips off elevated forks | WorkSafeBC | 19 |

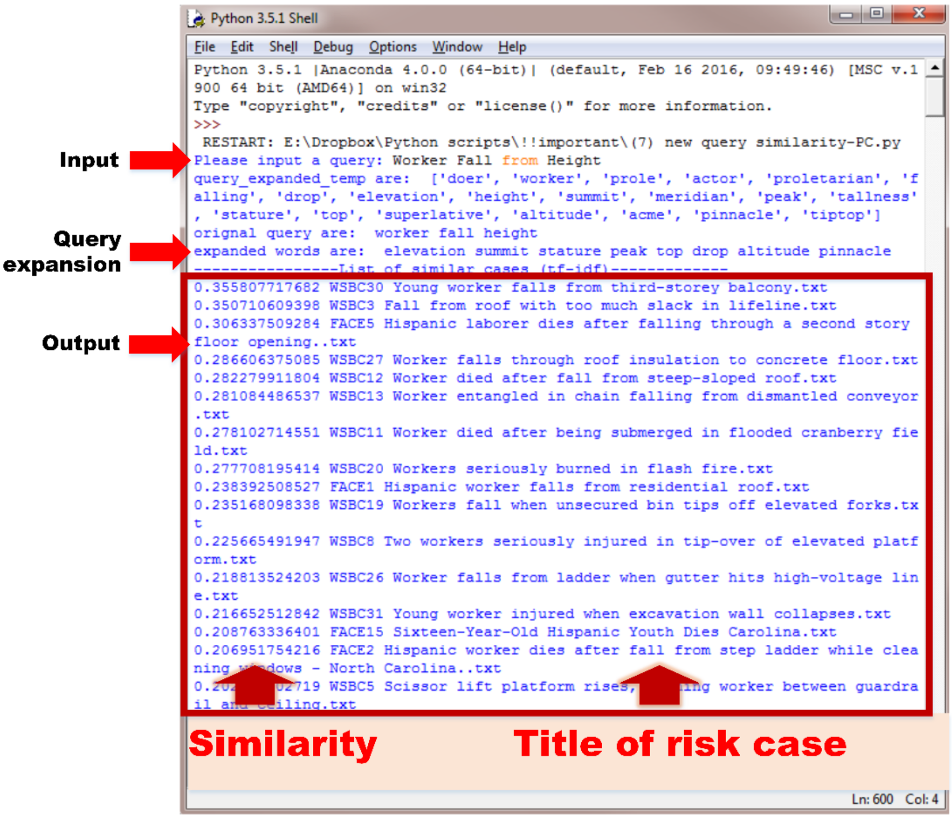



Figure 44. Snapshot of searching results of "Worker Fall from Height"

## 7.6 System testing

Although there are a number of matrices that have been proposed to evaluate and test IR systems, the most widely used are Precision, Recall and F score (Baeza-Yates and Ribeiro-Neto, 2011, Hsu, 2013, Tixier *et al.*, 2016) which can be calculated with the help of a simplified confusion matrix (Olson and Delen, 2008, Baeza-Yates and Ribeiro-Neto, 2011) shown in Table 8. There are four variables in the simplified confusion matrix, i.e. True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). Here the terms "positive" and "negative" means the expectation of a retrieval while the terms "true" and "false" refer to whether that expectation corresponds to the external judgment. In other words, TP means the number of relevant documents retrieved, FP means the number of irrelevant documents retrieved, FN means the number of relevant documents not retrieved, and TN means the number of irrelevant documents not retrieved.

Table 8. Confusion matrix

| | **Relevant** | **Not relevant** |
|---|---|---|
| **Retrieved** | True Positive (TP) | False Positive (FP) |
| **Not retrieved** | False Negative (FN) | True Negative (TN) |

Precision refers to the fraction of retrieved documents that is relevant and is used to measure the percentage of relevant documents in all retrieved documents, i.e.

$$Precision = \frac{TP}{TP+FP} \times 100\% \tag{7}$$

Recall is defined as the fraction of relevant documents that has been retrieved and used for measuring the percentage of retrieved documents in all relevant documents, i.e.

$$recall = \frac{TP}{TP+FN} \times 100\% \tag{8}$$

Another measure called F is the harmonic mean of $precision$ and $recall$ and is defined as follows:

$$F = \frac{precision \times recall}{precision + recall} \times 100\% \tag{9}$$

It is noticed that Precision, Recall, and F value are commonly used for evaluating the whole retrieval system and it requires an accurate boundary between “retrieved” and “not retrieved” to calculate the three measures. Here determining the threshold (or cut-off) is extremely important and its value could in large degree affect the evaluation results of an IR system. However, there is a need to point out determining the threshold value in an IR system is complex and needs a large number of experiments, which is not within the scope of this study. Unlike web-scale IR, the information in the construction industry is relatively small-scale and domain-specific and a common method to evaluate the performance of an IR system for construction projects is through testing a number of samples and setting user experience based threshold value, e.g. (Hsu, 2013, Fan and Li, 2013). Besides, with the observation that in the real working environment engineers often expect to obtain the needed information within a limited amount of time (Kazi, 2005) and the top 10-20 cases would by nature have the most value to the end users (Fan and Li, 2013), the proposed RCRS is designed to return the most similar top 10 cases. Hence, this study also evaluated the percentage of relevant risk cases among the top 10 similar documents, which is defined as Precision at 10 (P@10):

$$P@10 = \frac{number\ of\ relevant\ documents\ in\ top\ 10}{10} \times 100\% \tag{10}$$

In order to test and evaluate the proposed RCRS, this study took the threshold value as 0.1 from preliminary system use experience and the testing procedure consists of the following steps:

- Firstly, a set of key terms (e.g. “bridge”, “fall”, “collapse”, “construction”) that are relevant to the scope of collected risk cases were selected for making up 10 testing queries. The queries were divided into 3 groups, i.e. “type of risk”, “object + type of risk”, and “object + type of risk + project phase”, to simulate the real situations of case retrieval. The “type of risk” group contains three queries, i.e. “fall from height”, “flood risk”, “design error”. The “object + type of risk” group consists of 5 queries, i.e. “flood risk of bridge”, “worker fall from height”, “tower crane collapse”, “bridge failure”, “worker injury”. The “object + type of risk + project phase” group contains two queries, i.e. “worker die in construction” and “structure collapse in demolition”;

- Secondly, each testing query was inputted into the RCRS for query-document matching and the corresponding output was recorded in an Excel table. As this study took an experience-based threshold (or cut-off) value 0.1, those documents with the similarity score over 0.1 were classified into the "retrieved" group while those documents with the similarity score which is less than 0.1 were classified to the "not retrieved" group;

- Thirdly, because the similarity value for those documents containing no terms of original and expanded queries is 0, then those documents were determined to be irrelevant directly. Then the results were carefully reviewed to determine if a risk case is relevant to the query by quickly reading and understanding each document and analysing the relationship between the query and the document. If a document is determined to be relevant to the query, the value "1" was labelled for that document in Excel. Otherwise, the value "0" was given. Then, TP, FP, FN, TN and P@10 were calculated.

- In the last step, the calculation of Precision, Recall, and F value for each testing retrieval was performed and the testing results are shown in Table 9.

Table 9. Testing results

| No. | Testing query | Number of retrievals | | | | Performance | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | TP | FP | FN | TN | Precision | Recall | F | P@10 |
| 1 | fall from height | 18 | 1 | 18 | 553 | 94.7% | 50.0% | 65.5% | 90% |
| 2 | flood risk | 11 | 5 | 0 | 574 | 68.8% | 100.0% | 81.5% | 100% |
| 3 | design error | 22 | 4 | 6 | 558 | 84.6% | 78.6% | 81.5% | 100% |
| 4 | flood risk of bridge | 11 | 30 | 0 | 549 | 26.8% | 100.0% | 42.3% | 100% |
| 5 | worker fall from height | 25 | 10 | 2 | 553 | 71.4% | 92.6% | 80.6% | 90% |
| 6 | tower crane collapse | 18 | 23 | 0 | 549 | 43.9% | 100.0% | 61.0% | 70% |
| 7 | bridge failure | 42 | 16 | 3 | 529 | 72.4% | 93.3% | 81.6% | 100% |
| 8 | worker injury | 32 | 3 | 18 | 537 | 91.4% | 64.0% | 75.3% | 100% |
| 9 | worker die in construction | 30 | 1 | 11 | 548 | 96.8% | 73.2% | 83.3% | 100% |
| 10 | structure collapse in demolition | 16 | 34 | 0 | 540 | 32.0% | 100.0% | 48.5% | 100% |

The search results show that generally the proposed RCRS is capable of retrieving relevant risk cases from the database for a specified query. In particular, the results of P@10 are excellent, mostly 100% (7 of 10). Only one testing query had 70% of P@10, which also is satisfactory result. Therefore the top 10 cases returned by the system are valuable to the user. The high percentage of P@10 can be explained by the term frequency being an important factor in computing the TF-IDF weights and a document containing as many query terms as possible is easier to obtain a high similarity score. Although the Precision score for several queries are relatively low, this does not mean the retrieval results were not good. For example, for the “flood risk of bridge” query, 41 results were retrieved and only 11 were determined to be similar to the query. Two reasons could explain this problem: first, there are a very small number of “flood” related samples in the risk case database; second is because the threshold value 0.1 in this case is too small and the expanded terms were producing some “noise”. But from its P@10 score, it can be seen that the top 10 were all similar to the query and nearly all valuable documents were ranked. Therefore simply increasing the threshold value for some queries could improve the search results. In addition, some researchers (Hsu, 2013, Tixier *et al.*, 2016) also claim that there are still some technical limitations in the current NLP, which lead to the conclusion that the search results cannot be perfect. For example, the “flood risk” here is an entity but the system failed to read it as an entity and split it into two separate terms “flood” and “risk” for consideration.

This page is intentionally left blank

# Chapter 8. DISCUSSION AND CONCLUSIONS

In this chapter, evaluations and suggestions of future work about this PhD study from five industry experts are first described in Section 8.1. Then the theoretical contributions and practical significance are summarised in Sections 8.2 and 8.3, respectively. These are followed by the summary of limitations of this study and suggested future research in Section 8.4. Finally, the summary of conclusions of the whole research is presented in Section 8.5.

## 8.1 Comments from industry experts

In order to evaluate the practical value of this research in addressing the current challenges in BIM-based risk management, a group of leading industry experts whose expertise is BIM, structural safety, or project risk management was invited to participate in semi-structured interviews. Each interview lasted for 1-2 hours and consisted of the following three main processes. Firstly, the background to the study was introduced, and the research questions, methodologies, findings, and contributions were explained. Then, a free discussion was conducted to help the participants to have a deeper understanding about this research. Thirdly, the participants were asked to complete a questionnaire to answer the following four questions:

- Question1: What is your profession and how long have your worked in this area?
- Question 2: Do you think the observed knowledge gaps in this research are correct? If yes, why do the observed gaps exist in the industry from your

perspective? If no, could you please help point out what existing documents have covered them?

- Question 3: Do you think that the proposed theory and tool prototype have the potential to address the observed problem? And why?

- Question 4: Do you have any comments and suggestions about the proposed theory and tool, and for future research in the area of BIM-based risk management?

The group of industry experts included: Alastair Soane of Structural Safety, Gordon Crick of the Health and Safety Executive (HSE), Benedict Wallbank of Viewpoint Construction Software Ltd., David Philp of AECOM and UK BIM Task Group, and Martin Simpson of Arup. The interview documents and the experts' statements are included in Appendix C.

It can be concluded from the experts' comments that the identified knowledge gaps in current BIM-based project risk management are correct and this research was seen as a valuable contribution in addressing those gaps through the linkage approach. Their suggestions for the future research include, for example, research to extract knowledge automatically and quickly for decisions making and risk analysis, and research to extend the proposed approach to link site information, task documents, etc., within a collaborative data environment.

## 8.2 Theoretical contributions

Through a critical review of literature, it was observed that two main knowledge gaps exist in current BIM-based risk management: 1) very few theories exist that can explain how BIM could be aligned with traditional techniques and integrated into traditional work processes, and 2) current BIM solutions have very limited support on risk information sharing and communication during the project development process. The objective of this PhD research is to overcome these limitations and develop a methodology aligning traditional techniques with BIM to improve the systematic risk management and information sharing during the project development process.

As the main scientific contribution, the PhD research proposes a novel method of integrating two traditional techniques (i.e. RBS and CBR) into BIM through an active link, and outlines a new framework for a BKRMS. The theoretical contribution can be divided up and summarised in the following five sub-sections.

### 8.2.1 A Tailored RBS for bridge projects

One purpose of the PhD research is to address the current theoretical gap in integrating knowledge and experience into BIM for risk management of bridge projects by developing a tailored RBS and formalising an active link between the resulting RBS and BIM. In the first stage of the research, a tailored RBS for bridge projects was developed, which is the theoretical basis for developing the linkage between the resulting RBS and BIM afterwards. The tailored RBS developed in this research can be claimed as a knowledge contribution for the following reason.

Construction projects are varying in type, materials, construction methods, etc., and different people and organisations have different understanding about risks. Mehdizadeh *et al.* (2013) stated that currently there is no consensus on the standards or general methods of developing a RBS; however, the RBS to be developed should satisfy the particular purposes and requirements. This study took bridge engineering as the example and described a process to develop a tailored RBS for managing risks for bridge projects through a data mining approach (Jun Lee and Siau, 2001, Gargano and Raggad, 1999) and a further analysis of relationships between different categories of risks and LOCs of BIM. In particular, the places of different types of risk in the RBS were classified according to their relationship with the four LOCs of BIM, e.g. structure-related risks are related to bridge-level while the financial risks are related to the project-level. Although this study only focused on bridge projects, the approach of developing a tailored RBS can be generally applied to other types of project.

### 8.2.2 Conceptual separation of BIM into LOCs and Systems

With the observation that risks can affect and have impacts on a project and BIM differently, this study defined the conceptual separation of the integrated bridge information model into four LOCs and six technical systems for risk management purpose based on analysis of the IFC specification, a critical review of previous studies and the author's project experience. Some researchers, e.g. Fischer and Kam (2002) and Haymaker *et al.* (2003), realised that the data and its structure in a complete BIM are extremely complex and there is a need for division of the integrated BIM to meet particular needs. However, there are no existing studies that have defined the LOCs and technical systems of an integrated bridge information model for risk management purposes. Therefore, from this perspective, the conceptual separation of an integrated bridge information model into LOCs and systems is a piece of new knowledge. The separation is to help understand the relationship between different risks and the particular LOCs and systems associated with the risk, and a theoretical basis for establishing the linkage rules between RBS and BIM.

### 8.2.3 Linkage rules between RBS and BIM

Built on the tailored RBS and conceptual separation of BIM, this research developed a set of linkage rules between the resulting RBS and LOCs and technical systems of BIM. The rules consist of a basic linkage model and 13 sub-models. The basic linkage model explains the direct linked relationship between the four groups of risk in the RBS (project-level, surrounding environment-level, site-level and bridge-level) and the four LOCs in the integrated bridge information model (project, surrounding environment, site, bridge). And the 13 sub-models of linkage describe the specific relationships between each type of risk and the LOCs and systems of BIM. The feasibility of such a linkage approach to address the information sharing and communication in the design process has been illustrated by Kiviniemi (2005).

The linkage models provide evidence that RBS as a traditional technique can be integrated into BIM for project risk management through an active link between the RBS and BIM. On one hand, the tailored RBS as a hierarchical structure can be used for categorising and managing data in the knowledge-based risk database and could provide a global view on project risks. On the other hand, through linking risk information to the BIM, risks can be visualised, reviewed and managed in BIM throughout a project lifecycle. This proposed method merges the RBS with BIM as an integrated approach, takes advantage of both methods and could effectively facilitate identification, analysis, communication, and decision making of risks.

### 8.2.4 An approach of linking RBM to 4D BIM for project risk management

As a key theoretical contribution, this PhD research proposes a novel approach for systematic project risk management and information sharing through linking 4D BIM with RBM, and developed a new framework for a 4D BIM and RBM based risk management system. Established on the linkage rules, the proposed method and framework successfully describe how RBS, WBS, and BIM can be linked to each other and developed to be a risk management system, and further illustrate why risk information should be linked to BIM, how the linkage rules can be implemented, and how risk information could be visualised in 4D BIM for better communication. This proposed approach can be used as a helpful guide for integrating BIM into traditional risk management processes. Hartmann *et al.* (2012) conducted a case study to align BIM with traditional risk management process from a "technology-pull" perspective and highlighted that one significant advantage of such a method is the minimal disruption to the existing work process. The concept and feasibility of this approach was validated and tested by the tool prototype developed based on Navisworks 2017 and Microsoft SQL Server 2014.

### 8.2.5 A method of integrating NLP into CBR for risk case retrieval

The literature shows that CBR is a process of learning from the past, which could facilitate previous knowledge and experience to be effectively used for risk management in new projects. In the CBR cycle, RETRIEVE is the first and the most important step (De Mantaras *et al.*, 2005, Goh and Chua, 2009b). A commonly used traditional way for assessing the similarity between user need and risk cases is through attaching attribute labels to each risk case document and allocating different weights to those attributes (Kolodner, 1993, Karim and Adeli, 2003, Lu *et al.*, 2013). However, as discussed in Section 4.3, some challenges still exist: 1) traditional methods are very limited in scope, 2) a large amount of pre-processing or preparation work is needed, and 3) very few studies have been found to be capable of addressing the challenge of semantic similarity. In order to overcome the current challenges of RETRIEVE in

CBR, this PhD research analysed the potential and benefits of integrating NLP into CBR for risk case retrieval. The idea was motivated by some recent research that has introduced NLP into textual information management into the construction industry, e.g. retrieval of CAD drawings (Hsu, 2013), retrieval of relevant information for assisting decision making (Lv and El-Gohary, 2016a, Lv and El-Gohary, 2016b), injury report content analysis (Tixier *et al.*, 2016), and document clustering (Al Qady and Kandil, 2014). It can be seen that the application of NLP into textual documents analysis and management in the construction industry is a new and promising trend. Some recent studies even extended the use of NLP into BIM for automated code checking (Zhang and El-Gohary, 2017), processing building information (Beetz *et al.*, 2009), retrieving online BIM resources (Gao *et al.*, 2015), etc.

A number of recent studies (Fan and Li, 2013, Hsu, 2013) successfully used the classical VSM for IR and document management, and discussed that the semantic similarity is still a huge challenge in any current application of NLP in the construction industry. To partially overcome this gap, this PhD research outlines a framework of combining the use of semantic query expansion and VSM for retrieval of similar risk cases, and develops a system prototype with Python to support the proposed approach. The testing results show the proposed system could quickly and effectively retrieve and rank valuable risk cases if a query is specified. Through implementing the proposed system, end users could quickly find out risk cases that are valuable references to the new situations or problems and embed the knowledge and experience of previous accidents into daily work. Any new cases could be added into the risk case database flexibly for retrieval without pre-processing work. In addition, because this system prototype is written with Python, the RCRS could also be easily integrated into software written by other programming languages. As an example of its practical contributions, the proposed approach can be embedded into some online risk case databases, e.g. Structural-Safety and NIOSH, as a semantic searching engine. In the future, the proposed approach can be also expanded for a wider management of engineering documents and information.

## 8.3 Practical implications

The main practical significance of this PhD research is that the proposed approaches and frameworks enable the development of BIM-based risk management software to improve the risk identification, analysis, and information management during the project development process, which can be expanded upon as follows:

- BIM provides a new way of design, management and communication, and allows the project team to easily implement risk identification and analysis on their daily work. Through establishing the linkage between RBS and BIM, two main practical advantages become possible: 1) this solution can take advantage of both the traditional method and BIM for managing risks. On the one hand, RBS enables risk information to be stored in a proper structure, used and communicated effectively. On the other hand, some features of BIM such as 3D visualisation and 4D construction scheduling can facilitate the risk identification, analysis and communication at an early stage. Through this linkage, risks at different levels could be linked to the particular LOC and technical system in BIM for visualisation and management; and 2) risk information sharing and communication could be effectively improved by managing fragmented risk data using RBS and linking risk information to BIM. These data linked to BIM could provide important evidence for risk management and decision making at key stages. A practical example is that, when project information is being transferred between different people or forwarded to the next phase, project participants (e.g. client, principal designers, sub designers, and contractors) could check and review the attached information for identifying potential risks and seeking possible mitigation measures. Furthermore, conducting a design review is a legal requirement in the UK for identifying and mitigating any foreseeable health and safety risks (HSE, 2015).

- This study outlined a new framework for a BIM and RBM based risk management system and developed a tool prototype to support the proposed method. The proposed tool enables users to identify potential project risks with the help of 3D/4D BIM as well as RBM. The illustrated case study shows how risk information is stored in a central database and related information could be visualised and linked to BIM. Although the proposed theory was only tested in Navisworks, benefits and potential of the proposed system can be summarised as: 1) the risk of information loss because of updates or changes of the BIM and information transfer between different platforms should be reduced or avoided; and 2) there is the potential to develop a collaborative risk management system to support information sharing and collaboration in a multi-platform environment.

- This research proposed a framework of combining the use of semantic query expansion and VSM for risk case retrieval, and developed a system prototype with Python to support the proposed approach. The testing results showed the proposed system could quickly and effectively retrieve and rank valuable risk cases when a query is specified. Through implementing the proposed system, end users could quickly find out risk cases that are valuable references to the new situations or problems and embed the knowledge and experience of previous accidents into daily work. Any new cases could be added into the risk case database flexibly for retrieval without pre-processing work. In addition, because the risk case retrieval system was written with Python, the system could also be easily embedded or integrated into software written by other programming languages.

## 8.4 Limitations and suggested future research

Suggestions for future research can be primarily divided into the following two categories: 1) research to improve the risk analysis capability and implementation usability of the linkage approach, 2) research to further investigate how previous knowledge and lessons can be used effectively to support the risk identification, analysis and decision making process during the dynamic design, construction and maintenance stages through NLP, and 3) research to expand the method to other types of projects.

### 8.4.1 Improvement of the linkage approach

Firstly, RBM which combines RBS with WBS was chosen as a core traditional technique in this study because there is a linked relationship between RBS and BIM and at the same time 4D BIM and WBS share the same time- and schedule- related information. As a hierarchical structure, the RBS also is beneficial to improving the understanding of different project risks and the risk communication. However, Cagliano *et al.* (2015) pointed out that RBM and RBS are most suitable for risk identification and communication, and they have only limited use in risk analysis, evaluation and treatment. As a result, other risk analysis, evaluation and treatment techniques (e.g. sensitivity analysis) or professional simulations (e.g. structural analysis) may be needed to assist the use of the proposed risk management system. As BIM can be seen as an object-oriented data enriched model, it is noted that some researchers have taken advantage of the data in BIM for cost simulation and management (Smith, 2014) and structural analysis (Zhang and Hu, 2011). Therefore, one recommendation for future research is to enrich the risk analysis capability of the risk management system through integrating other appropriate techniques and methods.

Secondly, the proposed linkage approach and tool prototype were tested using Navisworks, a widely used 4D BIM based tool for project management. However, in practice, the information flow and exchange during a construction project lifecycle are often very complicated and a number of different tools and methods are used. Moreover, these tools are developed by different software vendors for different disciplines, and some of them do not provide API. Therefore, another recommendation for future research is to investigate how the proposed linkage method and prototype could be extended to support other BIM-based software (e.g. design tools) and how a collaborative risk management environment could be established to support multi-platforms in practice.

Thirdly, a straightforward case study was tested in this research and it is difficult to judge the real practical value of the proposed method and tool. To address this limitation, recommendations for future research could be: 1) to further develop easy-to-use software and test the proposed theory on both small and large projects, and 2) to investigate how the proposed theory could be integrated into existing workflows.

### 8.4.2 Use of previous knowledge and lessons

This research proposes a new framework of integrating two NLP techniques (VSM and query expansion) into CBR to facilitate the quick and accurate retrieval of previous knowledge and experience; however, Aamodt and Plaza (1994) pointed out that the implementation cycle of CBR contains four main processes, RETRIEVE, REUSE, REVISE, and RETAIN (known as 'the four REs'), and this research only focused on the RETRIEVE process (risk case retrieval). As a result, in the long term there is a need to investigate how human knowledge and experience on project risks can be implemented effectively in the full CBR cycle to support risk identification, analysis and decision making during the dynamic project development process.

In addition, some limitations also exist in the proposed approach for risk case retrieval. These limitations and the corresponding recommendations for future research are discussed as follows:

- First, the proposed system is limited in case retrieval within the internal risk case database and the total number of collected risk cases is still relatively small. As described in Section 7.2, due to the limited time only 590 risk cases covering 7 types of risk were collected. The reasons are: 1) the main purpose of this study is developing a general approach (i.e. proof of concept) based on NLP for risk case retrieval instead of establishing a complete risk case database; and 2) there are relatively few detailed reports on those risks that are not so dangerous or fatal, e.g. financial loss, time overrun. However, the limited size of the database will influence the retrieval results and practical applicability. For example, if a user query is "time overrun" and the database contains no risk cases about "time overrun", it will be difficult for the system to return the desired results to the user. Therefore, future research may consider: 1) how to enrich the risk case database, 2) how to formulate case retrieval guidelines to the end user according to the distribution of risk cases; 3) how to extend the proposed system for risk case retrieval in external databases and online resources.

- Secondly, the semantic similarity problem is still a huge challenge within the state-of-the-art research of NLP (Jurafsky and Martin, 2009), and the query expansion approach adopted by this study can only address a limited proportion of the problem. In particular, the proposed system combines the use of a pre-defined risk-related lexicon and WordNet to deal with the word mismatching problem of case retrieval. However, the pre-defined lexicon only contains explanations of 107 key terms in the project risk management domain and is not a complete dictionary. To overcome the shortcoming of the pre-defined lexicon, WordNet is used as an important supplementary source. However, because WordNet is a large lexical database for the English language and is not specially designed for risk management, this study found some terms expanded by WordNet are not related to project risks and have little, or no value in risk case retrieval. Moreover, it can be seen that human language is still extremely complex and difficult for computers to understand and process. For example, Caldas and Han (Caldas *et al.*, 2002) made use of IR and text mining for automatic classification of project documents but found the results were not perfect due to the multiple meanings of words. In addition, as discussed in Section 7.3, although the pre-defined lexicon and WordNet can be used for explanation of a single term, it is still difficult for computer to process the word groups. Hence, one short-term recommendation for future research may be to establish a comprehensive lexicon for project risk management which includes the definition of the linked relationships of common word groups. From a long-term perspective, future research may apply the state-of-the-art techniques of NLP into addressing the semantic similarity problem in both risk case retrieval and other fields.

- Thirdly, the proposed system has not been put into use and validated in real practice. For better implementation of the proposed approach, the prototype system needs to be further developed as a tool with easy-to-use user interface and checked by different scenarios. In addition, as the proposed system was designed to return the most similar 10 risk cases to the user and the testing results are satisfactory, when conducting the preliminary testing this paper checked the results manually and did not study the best value of the threshold. Although a number of matrices (e.g. Precision, Recall, F and P@10) could be used for evaluating an IR system, nearly all of them require a clear boundary of "retrieved" and "not retrieved", and "relevant" and "not relevance". The threshold value is often used to divide the returned results into "retrieved" and "not retrieved"; however, Al Qady and Kandil (2014) pointed out the best threshold value normally lies between 0.05 and 0.95, and determining the best value needs a large number of experiments. Furthermore, the relevance is by nature often continuous instead of binary, which leads to the difficulty of determining if a retrieved document is relevant or not (Kekäläinen, 2005, Janes, 1991). Hence, future research may further study the threshold value and relevance problem, and test and improve the proposed approach and system in real practice.

### 8.4.3 Expansion of the research to other project types

Bridges were chosen as the project type within the scope of this research to study the RBS, linkage rules and case study. Future research may apply the basic methods and principles to other project types, e.g. buildings, roads, railways and power plants. For example, different types of risk may be present for other project types and future research may expand the tailored RBS and linkage rules to cover other type of projects. In addition, the concepts of separation of BIM, and the linkage rules may be used in the Common Data Environment (CDE) and for the expansion of open BIM standards.

## 8.5 Summary of conclusions

Utilising BIM and BIM-related digital technologies to manage project risks has been a growing research interest in the AEC industry. Successful use of these technologies requires a comprehensive understanding of the fundamentals, general processes, techniques of risk management and the relationship between the new and traditional methods.

The literature shows the implementation of traditional risk management is still a manual undertaking, the assessment is heavily reliant on experience and mathematical analysis, and the decision making is frequently based on knowledge and experience based intuition, which leads to a decreased efficiency in the real environment. To improve the above situation, some standards or governmental documents (e.g. ISO 31010:2009, CDM regulations) put emphasis on foreseeable risks being identified and mitigated at an early stage and risk information should be documented and updated during the development process of a project. This is where BIM could be of help. BIM can not only be used as a systematic risk management tool in the development process, but also can act as a core data generator and platform to allow other BIM-based tools to carry out further risk analysis. However, it was observed that BIM-based risk management has not been widely used in the real workplace because of the following obstacles: 1) very few theories exist that can explain how BIM can be aligned with traditional techniques and integrated into existing processes for project risk management. Hartmann *et al.* (2012) highlighted that one significant benefit of addressing the gap is that there would be little disruption of existing work practices; and 2) current BIM solutions have very limited support on risk information sharing and communication during the project development process (Han *et al.*, 2008, Frewer, 2004).

To overcome these limitations, this PhD thesis proposes a novel method of integrating two traditional techniques (i.e. RBS and CBR) into BIM for risk management and information sharing through an active link, and outlines a new framework for a BKRMS. The core motivations behind the proposed solution are: 1) a tailored RBS could be used as a knowledge-based approach to classify, store and manage the information of a risk database in a proper structure and risk information in RBS could be linked to BIM for review, visualisation and communication; and 2) a CBRL contains a number of risk case documents written in everyday language and previous knowledge and experience stored in those documents could contribute to avoiding similar risks in new situations during the project lifecycle.

The scope of this research was limited to bridge projects; however, the basic methods and principles could be also applied to other AEC projects. In order to achieve the research objectives, three main steps were used.

In the first stage, with the observation that different types of risk have impacts on different levels of a project, this research analysed the conceptual division of BIM and the linkage between different types of risk and BIM. Specifically, an integrated bridge information model is firstly divided into four LOCs and six technical systems based on analysis of the IFC specification, a critical review of previous studies and the author's project experience. Then, a knowledge-based risk database containing various types of risk for bridge projects was developed through data collection, risk mining, and assessment and translation. Risk data referred to here is mined from academic publications, risk assessment reports of real bridge projects, and related standards. Built on the results in the risk database, a tailored RBS was developed to categorise and manage this risk information and performed on a theoretical basis for establishing the linkage between the resulting RBS and BIM. Lastly, an overall framework and 13 sub models were established for the linking the RBS with the four LOCs and the six systems of BIM.

In the next stage, to further implement the linkage rules, this research developed a novel method to link BIM, RBS, and WBS as a whole, and outlined a new risk management framework based on 4D BIM with a RBM. A tool prototype was developed based on Navisworks and Microsoft SQL Server and the proposed theory and tool were tested through a selected case study. The results show, on one hand, that risk identification could be facilitated through 3D/4D BIM, and on the other hand, that RBM as a traditional technique could be used as a guide to improve the understanding of project risks and the management of risk information.

In the third stage of the PhD research, to facilitate the use of previous knowledge and experience for risk management during the project development process, an approach of combining the use of two NLP techniques (i.e. VSM and semantic query expansion) for risk case retrieval was proposed and a new framework for the risk case retrieval system was developed. The VSM could represent textual documents as vectors of identifiers and assigning TF-IDF weights to index terms in both queries and documents, which can be used to compute the degree of similarity between documents and the query. While the query expansion could solve the mismatching problem of terms that have the same semantic meanings through expanding the original query using related terms defined in a pre-defined risk-related lexicon and synonyms in WordNet. A prototype system was developed using Python to implement and test the proposed approach. Through implementing the proposed system, textual content information is firstly extracted from the risk case dataset and processed to generate a content corpus. After a query is inputted by the user, then the system starts to read and process the query, combines the use of a pre-defined risk-related lexicon or WordNet to expand the original query, and filters out the query terms that do not exist in the content corpus. Lastly the system gathers original query, expanded query and content corpus together for query-document similarity computing and returns the top 10 similar risk cases to the user. The preliminary test results have demonstrated successfully the system's capacity of automatically retrieving similar risk cases.

The proposed solution would push risk management a step forward by aligning traditional methods with BIM to systematically support the development process of a project. Firstly, the developed knowledge-based risk database, RBS, and the conceptual linkage model not only can be implemented manually as effective tools for understanding and managing project risks but have a practical value for developing BIM-based risk management software. Secondly, the 4D BIM and RBM based risk management system allows not only the storage of risk information in a central database but also to link the related risk information in the BIM model for review, visualisation and simulation. In addition, the proposed method and tool prototype provide both theoretical and technical evidence of the potential for developing a collaborative risk management system to support multi-platforms and the project development processes. Thirdly, although there are still some limitations of applying current NLP technology into engineering textual information management, using a NLP supported system to manage risk cases could effectively facilitate the risk identification and communication, and information management.

This page is intentionally left blank

# REFRENCES


Aamodt, A. and Plaza, E. (1994), "Case-based reasoning: Foundational issues, methodological variations, and system approaches", *Artificial Intelligence Communications,* Vol. 7 No. 1, pp. 39-59.

Abderrahim, M., Garcia, E., Diez, R. and Balaguer, C. (2005), "A mechatronics security system for the construction site", *Automation in Construction,* Vol. 14 No. 4, pp. 460-66.

Al Qady, M. and Kandil, A. (2014), "Automatic clustering of construction project documents based on textual similarity", *Automation in Construction,* Vol. 42, pp. 36-49.

Alaeddini, A. and Dogan, I. (2011), "Using Bayesian networks for root cause analysis in statistical process control", *Expert Systems with Applications,* Vol. 38 No. 9, pp. 11230-43.

Alasuutari, P. (2010), "The rise and relevance of qualitative research", *International journal of social research methodology,* Vol. 13 No. 2, pp. 139-55.

Aleshin, A. (2001), "Risk management of international projects in Russia", *International Journal of Project Management*, Vol. 19 No. 4, pp. 207-22.

Allread, B.S. (2009), Real-time Pro-active Safety in Construction, MSc Dissertation, Georgia Institute of Technology, Atlanta, USA.

Atkins (2006), *Walton Bridge Flood Risk Assessment,* Atkins, Surrey, UK.

Autodesk (2016a), "Navisworks", available at: http://www.autodesk.com/products/navisworks/overview (accessed 10 October 2016).

Autodesk (2016b), "Revit", available at: http://www.autodesk.com/products/revit-family/overview (accessed 07 November 2016).

Azhar, S. (2011), "Building information modeling (BIM): Trends, benefits, risks, and challenges for the AEC industry", *Leadership and Management in Engineering*, Vol. 11 No. 3, pp. 241-52.

Babbie, E.R. (2015), *The practice of social research,* Cengage Learning, Boston, MA.

Baeza-Yates, R. and Ribeiro-Neto, B. (2011), *Modern information retrieval : the concepts and technology behind search*, second ed., Addison Wesley, Harlow, UK.

Bai, X. (2011), "Predicting consumer sentiments from online text", *Decision Support Systems,* Vol. 50 No. 4, pp. 732-42.

Balachandran, M., Rosenman, M. and Gero, J. (1991), "A knowledge-based approach to the automatic verification of designs from CAD databases", paper presented at *Artificial Intelligence in Design*, Oxford, UK.

Banaitiene, N. and Banaitis, A. (2012), "Risk Management in Construction Projects", in Banaitiene, N. (Ed.), *Risk Management – Current Issues and Challenges*, Intech, Rijeka, Croatia, pp. 429-48.

Bansal, V. (2011), "Application of geographic information systems in construction safety planning", *International Journal of Project Management,* Vol. 29 No. 1, pp. 66-77.

Bansal, V.K. and Pal, M. (2007), "Potential of geographic information systems in building cost estimation and visualization", *Automation in Construction,* Vol. 16 No. 3, pp. 311-22.

Bar-Hillel, Y. (1960), "The present status of automatic translation of languages", *Advances in computers,* Vol. 1, pp. 91-163.

Baskerville, R.L. and Wood-Harper, A.T. (1996), "A critical perspective on action research as a method for information systems research", *Journal of information Technology,* Vol. 11 No. 3, pp. 235-46.

Becerik-Gerber, B., Jazizadeh, F., Li, N. and Calis, G. (2011), "Application areas and data requirements for BIM-enabled facilities management", *Journal of Construction Engineering and Management,* Vol. 138 No. 3, pp. 431-42.

Beetz, J., Van Leeuwen, J. and De Vries, B. (2009), "IfcOWL: A case of transforming EXPRESS schemas into ontologies", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing,* Vol. 23 No. 01, pp. 89-101.

Benjaoran, V. and Bhokha, S. (2010), "An integrated safety management with construction management using 4D CAD model", *Safety Science,* Vol. 48 No. 3, pp. 395-403.

Bhatla, A., Choe, S.Y., Fierro, O. and Leite, F. (2012), "Evaluation of accuracy of as-built 3D modeling from photos taken by handheld digital cameras", *Automation in Construction,* Vol. 28, pp. 116-27.

Borrmann, A., Hyvärinen, J. and Rank, E. (2009), "Spatial constraints in collaborative design processes", in *International Conference on Intelligent Computing in Engineering (ICE09) in Berlin, Germany, 15-17 July 2009,* pp. 1-8.

Bryde, D., Broquetas, M. and Volm, J.M. (2013), "The project benefits of Building Information Modelling (BIM)", *International Journal of Project Management,* Vol. 31 No. 7, pp. 971-80.

buildingSMART (2016a), "buildingSMART - International home of openBIM", available at: http://www.buildingsmart-tech.org/ (accessed 06 December 2016).

buildingSMART (2016b), "ongoing infrastructure projects", available at: http://www.buildingsmart-tech.org/infrastructure/projects (accessed 4 January 2016).

Cagliano, A.C., Grimaldi, S. and Rafele, C. (2015), "Choosing project risk management techniques. A theoretical framework", *Journal of Risk Research,* Vol. 18 No. 2, pp. 232-48.

Caldas, C.H. and Soibelman, L. (2003), "Automating hierarchical document classification for construction management information systems", *Automation in Construction,* Vol. 12 No. 4, pp. 395-406.

Caldas, C.H., Soibelman, L. and Han, J. (2002), "Automated classification of construction project documents", *Journal of Computing in Civil Engineering,* Vol. 16 No. 4, pp. 234-43.

Carter, G. and Smith, S.D. (2006), "Safety hazard identification on construction projects", *Journal of Construction Engineering and Management,* Vol. 132 No. 2, pp. 197-205.

Chapman, R.J. (2001), "The controlling influences on effective risk identification and assessment for construction design management", *International Journal of Project Management,* Vol. 19 No. 3, pp. 147-60.

Chen, L. and Luo, H. (2014), "A BIM-based construction quality management model and its applications", *Automation in Construction,* Vol. 46, pp. 64-73.

Cheng, J.C. and Ma, L.Y. (2013), "A BIM-based system for demolition and renovation waste estimation and planning", *Waste management*, Vol. 33 No. 6, pp. 1539-51.

Cheng, M.Y., Ko, C.H. and Chang, C.H. (2002), "Computer-aided DSS for safety monitoring of geotechnical construction", *Automation in Construction*, Vol. 11 No. 4, pp. 375-90.

Cheng, T. and Teizer, J. (2014), "Modeling tower crane operator visibility to minimize the risk of limited situational awareness", *Journal of Computing in Civil Engineering*, Vol. 28 No. 3, pp. 04014004.

Cheung, S.O., Cheung, K.K. and Suen, H.C. (2004), "CSHM: Web-based safety and health monitoring system for construction management", *Journal of Safety Research,* Vol. 35 No. 2, pp. 159-70.

Chi, N.-W., Lin, K.-Y. and Hsieh, S.-H. (2014), "Using ontology-based text classification to assist Job Hazard Analysis", *Advanced Engineering Informatics,* Vol. 28 No. 4, pp. 381-94.

Chiu, C.-T., Hsu, T.-H., Wang, M.-T. and Chiu, H.-Y. (2011), "Simulation for steel bridge erection by using BIM tools", in *28th International Symposium on Automation and Robotics in Construction (ISARC) in Seoul, Republic of Korea, 29 June - 2 July, 2011*, International Association for Automation and Robotics in Construction (IAARC), pp. 560-63.

Choi, H.-H. and Mahadevan, S. (2008), "Construction project risk assessment using existing database and project-specific information", *Journal of Construction Engineering and Management,* Vol. 134 No. 11, pp. 894-903.

Chowdhury, G.G. (2003), "Natural language processing", *Annual review of information science and technology,* Vol. 37 No. 1, pp. 51-89.

Clough, R.H., Sears, G.A. and Sears, S.K. (2000), *Construction project management,* John Wiley & Sons, New York.

Colace, F., De Santo, M., Greco, L. and Napoletano, P. (2015), "Weighted word pairs for query expansion", *Information Processing & Management,* Vol. 51 No. 1, pp. 179-93.

Colding, J. (2007), "'Ecological land-use complementation'for building resilience in urban ecosystems", *Landscape and urban planning,* Vol. 81 No. 1, pp. 46-55.

Cooke, T., Lingard, H., Blismas, N. and Stranieri, A. (2008), "ToolSHeDTM: the development and evaluation of a decision support tool for health and safety in construction design", *Engineering, Construction and Architectural Management,* Vol. 15 No. 4, pp. 336-51.

Coombes, H. (2001), *Research using IT,* Palgrave Macmillan, Houndmills, UK.

Creswell, J.W. (2013), *Research design: Qualitative, quantitative, and mixed methods approaches*, Sage publications, Thousand Oaks, California.

Cunningham, P. (2009), "A taxonomy of similarity mechanisms for case-based reasoning", *IEEE Transactions on Knowledge and Data Engineering,* Vol. 21 No. 11, pp. 1532-43.

De Mantaras, R.L., McSherry, D., Bridge, D., Leake, D., Smyth, B., Craw, S., Faltings, B., Maher, M.L., Cox, M.T. and Forbus, K. (2005), "Retrieval, reuse, revision and retention in case-based reasoning", *The Knowledge Engineering Review,* Vol. 20 No. 03, pp. 215-40.

De Simone, T. and Kazakov, D. (2005), "Using wordnet similarity and antonymy relations to aid document retrieval", in *Proceedings of Recent Advances in Natural Language Processing (RANLP) in Borovets, Bulgaria, 21-23 September 2005*.

Delis, E.A. and Delis, A. (1995), "Automatic fire-code checking using expert-system technology", *Journal of Computing in Civil Engineering,* Vol. 9 No. 2, pp. 141-56.

Deshpande, A., Azhar, S. and Amireddy, S. (2014), "A Framework for a BIM-based Knowledge Management System", *Procedia Engineering,* Vol. 85, pp. 113-22.

Dey, P.K. (2002), "Project risk management: a combined analytic hierarchy process and decision tree approach", *Cost Engineering,* Vol. 44 No. 3, pp. 13-27.

Dey, P.K. (2010), "Managing project risk using combined analytic hierarchy process and risk map", *Applied Soft Computing,* Vol. 10 No. 4, pp. 990-1000.

Dikmen, I., Birgonul, M., Anac, C., Tah, J. and Aouad, G. (2008), "Learning from risks: A tool for post-project risk assessment", *Automation in Construction,* Vol. 18 No. 1, pp. 42-50.

Ding, L.Y., Zhong, B.T., Wu, S. and Luo, H.B. (2016), "Construction risk knowledge management in BIM using ontology and semantic web technology", *Safety Science,* Vol. 87, pp. 202-13.

Ding, L.Y. and Zhou, C. (2013), "Development of web-based system for safety risk early warning in urban metro construction", *Automation in Construction,* Vol. 34, pp. 45-55.

Dossick, C.S. and Neff, G. (2011), "Messy talk and clean technology: communication, problem-solving and collaboration using Building Information Modelling", *The Engineering Project Organization Journal,* Vol. 1 No. 2, pp. 83-93.

Easterby-Smith, M., Thorpe, R. and Lowe, A. (1991), *Management research: an introduction,* Sage Publications, London.

Eastman, C., Lee, J., Jeong, Y. and Lee, J. (2009), "Automatic rule-based checking of building designs", *Automation in Construction,* Vol. 18 No. 8, pp. 1011-33.

Eastman, C., Teicholz, P., Sacks, R. and Liston, K. (2011), *BIM handbook: A guide to building information modeling for owners, managers, designers, engineers and contractors,* John Wiley & Sons, Hoboken, New Jersey.

Eastman, C.M. (1999), *Building product models: computer environments, supporting design and construction,* CRC press, Boca Raton, Florida.

El-Sayegh, S.M. (2008), "Risk assessment and allocation in the UAE construction industry", *International Journal of Project Management,* Vol. 26 No. 4, pp. 431-38.

Esmaeili, B. and Hallowell, M. (2012), "Attribute-based risk model for measuring safety risk of struck-by accidents", in *Construction Research Congress in West Lafayette, Indiana, 21-23 May 2012,* American Society of Civil Engineers, New York, pp. 289-98.

Fan, H. and Li, H. (2013), "Retrieving similar cases for alternative dispute resolution in construction accidents using text mining techniques", *Automation in Construction,* Vol. 34, pp. 85-91.

Fan, L. (2012), *Bridge Engineering: Part I (In Chinese),* China Communications Press, Beijing, China.

Fang, D.P., Li, M.G., Fong, P.S.W. and Shen, L.Y. (2004), "Risks in Chinese construction market - Contractors' perspective", *Journal of Construction Engineering and Management,* Vol. 130 No. 6, pp. 853-61.

Fellbaum, C. (1998), *WordNet: an electronic lexical database, Language, speech, and communication,* MIT Press, Cambridge, Mass.

Fenves, S.J. (1966), "Tabular decision logic for structural design", *Journal of the Structural Division,* Vol. 92 No. 6, pp. 473-90.

Fenves, S.J., Garrett, J., Kiliccote, H., Law, K. and Reed, K. (1995), "Computer representations of design standards and building codes: US perspective", *The International Journal of Construction Information Technology,* Vol. 3 No. 1, pp. 13-34.

Fiatech (2013), *An Overview of Existing BIM Standards and Guidelines: A Report to Fiatech AutoCodes Project*, Austin, USA.

Fischer, M. and Kam, C. (2002), *PM4D final report,* CIFE, Stanford University, Stanford, California.

Flanagan, R. and Norman, G. (1993), *Risk management and construction,* Blackwell Scientific, Oxford.

Forbes, D., Smith, S. and Horner, M. (2008), "Tools for selecting appropriate risk management techniques in the built environment", *Construction Management and Economics,* Vol. 26 No. 11, pp. 1241-50.

Forsythe, P. (2014), "Proactive construction safety systems and the human factor", *Proceedings of the Institution of Civil Engineers: Management, Procurement and Law,* Vol. 167 No. 5, pp. 242-52.

Frewer, L. (2004), "The public and effective risk communication", *Toxicology Letters,* Vol. 149 No. 1–3, pp. 391-97.

Gable, G.G. (1994), "Integrating case study and survey research methods: an example in information systems", *European journal of information systems*, Vol. 3 No. 2, pp. 112-26.

Gambatese, J.A., Behm, M. and Hinze, J.W. (2005), "Viability of designing for construction worker safety", *Journal of Construction Engineering and Management,* Vol. 131 No. 9, pp. 1029-36.

Gao, G., Liu, Y.-S., Wang, M., Gu, M. and Yong, J.-H. (2015), "A query expansion method for retrieving online BIM resources based on Industry Foundation Classes", *Automation in Construction,* Vol. 56, pp. 14-25.

Gargano, M.L. and Raggad, B.G. (1999), "Data mining-a powerful information creating tool", *OCLC Systems & Services: International digital library perspectives,* Vol. 15 No. 2, pp. 81-90.

Garrett Jr, J.H. and Fenves, S.J. (1987), "A knowledge-based standards processor for structural component design", *Engineering with Computers,* Vol. 2 No. 4, pp. 219-38.

Given, L.M. (2008), *The Sage encyclopedia of qualitative research methods,* Sage Publications, Thousand Oaks, California.

Goedert, J.D. and Meadati, P. (2008), "Integrating construction process documentation into building information modeling", *Journal of Construction Engineering and Management,* Vol. 134 No. 7, pp. 509-16.

Goh, Y.M. and Chua, D. (2009a), "Case-based reasoning approach to construction safety hazard identification: adaptation and utilization", *Journal of Construction Engineering and Management,* Vol. 136 No. 2, pp. 170-78.

Goh, Y.M. and Chua, D. (2009b), "Case-based reasoning for construction hazard identification: case representation and retrieval", *Journal of Construction Engineering and Management,* Vol. 135 No. 11, pp. 1181-89.

Goldenhar, L.M., Moran, S.K. and Colligan, M. (2001), "Health and safety training in a sample of open-shop construction companies", *Journal of Safety Research,* Vol. 32 No. 2, pp. 237-52.

Gong, Z. and Cheang, C.W. (2004), "An implementation of web image search engines", in *Proceedings of International Conference on Asian Digital Libraries in Shanghai, China, 13-17 December 2004,* Springer, pp. 355-67.

Gong, Z., Cheang, C.W. and Hou, U.L. (2005), "Web query expansion by WordNet", in *Proceedings of 16th International Conference on Database and Expert Systems Applications in Copenhagen, Denmark, 22-26 August 2005,* Springer, pp. 166-75.

Graphisoft (2016), "ArchiCAD", available at: http://www.graphisoft.com/archicad/ (accessed 07 November 2016).

Grilo, A. and Jardim-Goncalves, R. (2010), "Value proposition on interoperability of BIM and collaborative working environments", *Automation in Construction,* Vol. 19 No. 5, pp. 522-30.

Gu, N. and London, K. (2010), "Understanding and facilitating BIM adoption in the AEC industry", *Automation in Construction,* Vol. 19 No. 8, pp. 988-99.

Guo, H., Li, H., Chan, G. and Skitmore, M. (2012), "Using game technologies to improve the safety of construction plant operations", *Accident Analysis & Prevention,* Vol. 48, pp. 204-13.

Guo, H.L., Li, H. and Li, V. (2013), "VP-based safety management in large-scale construction projects: A conceptual framework", *Automation in Construction,* Vol. 34, pp. 16-24.

Hadikusumo, B.H.W. and Rowlinson, S. (2002), "Integration of virtually real construction model and design-for-safety-process database", *Automation in Construction,* Vol. 11 No. 5, pp. 501-09.

Hadikusumo, B.H.W. and Rowlinson, S. (2004), "Capturing safety knowledge using design-for-safety-process tool", *Journal of Construction Engineering and Management,* Vol. 130 No. 2, pp. 281-89.

Han, C.S., Kunz, J.C. and Law, K.H. (2002), "Compliance analysis for disabled access", in McIver, W.J. and Elmagarmid, A.K. (Eds.), *Advances in Digital Government: Technology, Human Factors, and Policy*, Kluwer Academic Publishers, Springer, New York, NY, pp. 149-62.

Han, S.H., Kim, D.Y., Kim, H. and Jang, W.-S. (2008), "A web-based integrated system for international project risk management", *Automation in Construction,* Vol. 17 No. 3, pp. 342-56.

Hardin, B. (2011), *BIM and construction management: proven tools, methods, and workflows,* John Wiley & Sons, Indianapolis, Indiana.

Harispe, S., Ranwez, S., Janaqi, S. and Montmain, J. (2015), "Semantic similarity from natural language and ontology analysis", *Synthesis Lectures on Human Language Technologies,* Vol. 8 No. 1, pp. 1-254.

Hartmann, T., Gao, J. and Fischer, M. (2008), "Areas of application for 3D and 4D models on construction projects", *Journal of Construction Engineering and Management,* Vol. 134 No. 10, pp. 776-85.

Hartmann, T., Van Meerveld, H., Vossebeld, N. and Adriaanse, A. (2012), "Aligning building information model tools and construction management methods", *Automation in Construction,* Vol. 22, pp. 605-13.

Haymaker, J., Suter, B., Kunz, J. and Fischer, M. (2003), *PERSPECTORS: Automating the construction and coordination of multidisciplinary 3d design representations*, Technical Report 145, CIFE, Stanford University, Stanford, California.

HBC (2008), *Mersey Gateway Design and Access Statement*, Halton Borough Council (HBC), Halton, UK.

Heesom, D. and Mahdjoubi, L. (2004), "Trends of 4D CAD applications for construction planning", *Construction Management and Economics,* Vol. 22 No. 2, pp. 171-82.

Hillson, D., Grimaldi, S. and Rafele, C. (2006), "Managing project risks using a cross risk breakdown matrix", *Risk management*, Vol. 8 No. 1, pp. 61-76.

Ho, S.P., Tserng, H.P. and Jan, S.H. (2013), "Enhancing knowledge sharing management using BIM technology in construction", *The Scientific World Journal,* Vol. 2013, pp. 170498.

Holzmann, V. and Spiegler, I. (2011), "Developing risk breakdown structure for information technology organizations", *International Journal of Project Management,* Vol. 29 No. 5, pp. 537-46.

Howell, K.E. (2012), *An introduction to the philosophy of methodology,* Sage Publications, London.

HSE (2015), *The Construction (Design and Management) Regulations,* Health and Safety Executive (HSE), London, UK.

Hsu, J.-Y. (2013), "Content-based text mining technique for retrieval of CAD documents", *Automation in Construction,* Vol. 31, pp. 65-74.

Hu, Z. and Zhang, J. (2011), "BIM- and 4D-based integrated solution of analysis and management for conflicts and structural safety problems during construction: 2. Development and site trials", *Automation in Construction,* Vol. 20 No. 2, pp. 167-80.

ILO (2005), *Fact Sheet on Safety at Work*, International Labour Organization (ILO), Geneva, Switzerland.

Imhof, D. (2004), Risk assessment of existing bridge structures, PhD Thesis, University of Cambridge, Cambridge, UK.

Irizarry, J. and Karan, E.P. (2012), "Optimizing location of tower cranes on construction sites through GIS and BIM integration", *Journal of Information Technology in Construction,* Vol. 17, pp. 361-66.

ISO (2009), *Risk management - Risk assessment techniques*, IEC/FDIS 31010:2009, International Organization for Standardization (ISO), Geneva, Switzerland.

Janes, J.W. (1991), "The binary nature of continuous relevance judgments: A study of users' perceptions", *Journal of the American Society for Information Science,* Vol. 42 No. 10, pp. 754-56.

Ji, Y., Borrmann, A., Beetz, J. and Obergriesser, M. (2013), "Exchange of Parametric Bridge Models Using a Neutral Data Format", *Journal of Computing in Civil Engineering,* Vol. 27 No. 6, pp. 593-606.

Jonassen, D.H. and Hernandez-Serrano, J. (2002), "Case-based reasoning and instructional design: Using stories to support problem solving", *Educational Technology Research and Development,* Vol. 50 No. 2, pp. 65-77.

Joyce, R. and Houghton, D. (2014), "Briefing: Building information modelling and the law", *Proceedings of the ICE-Management, Procurement and Law,* Vol. 167 No. 3, pp. 114-16.

Jun Lee, S. and Siau, K. (2001), "A review of data mining techniques", *Industrial Management & Data Systems,* Vol. 101 No. 1, pp. 41-46.

Jurafsky, D. and Martin, J.H. (2009), *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition, Prentice Hall series in artificial intelligence,* second ed., Prentice Hall, New Jersey.

Kaminetzky, D. (2001), *Design and construction failures: Lessons from forensic investigations,* Mcgraw-Hill, New York.

Karim, A. and Adeli, H. (2003), "CBR Model for Freeway Work Zone Traffic Management", *Journal of Transportation Engineering,* Vol. 129 No. 2, pp. 134-45.

Kartam, N.A. and Kartam, S.A. (2001), "Risk and its management in the Kuwaiti construction industry: a contractors' perspective", *International Journal of Project Management,* Vol. 19 No. 6, pp. 325-35.

Kazi, A.S. (2005), *Knowledge management in the construction industry: A socio-technical perspective,* IGI Global, Hershey, Pennsylvania.

Kekäläinen, J. (2005), "Binary and graded relevance in IR evaluations—comparison of the effects on ranking of IR systems", *Information Processing & Management,* Vol. 41 No. 5, pp. 1019-33.

Khoshgoftaar, T.M. and Lanning, D.L. (1995), "A neural network approach for early detection of program modules having high risk in the maintenance phase", *Journal of Systems and Software,* Vol. 29 No. 1, pp. 85-91.

Khribi, M.K., Jemni, M. and Nasraoui, O. (2008), "Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval", in *Proceedings of 2008 Eighth IEEE International Conference on Advanced Learning Technologies in Santander, Cantabria, 1-5 July 2008*, IEEE, New Zealand, pp. 241-45.

Kim, C., Haas, C.T., Liapi, K.A., McLaughlin, J., Teizer, J. and Bosche, F. (2004), "Rapid human-assisted, obstacle avoidance system using sparse range point clouds", in *Proceedings of Engineering, Construction and Operations in Challenging Environments: Earth and Space 2004 in New York, March 7-10, 2004,* American Society of Civil Engineers, New York, pp. 115-22.

Kim, J.B., Clayton, M.J. and Yan, W. (2011), "Parametric Form-Based Codes: Incorporation of Land-use Regulations into Building Information Models", in *Parametricism (SPC) ACADIA Regional 2011 Conference in Lincoln, Nebraska, United States*, pp. 217-23.

Kim, T.W., Kavousian, A., Fischer, M. and Rajagopal, R. (2012), *Improving Facility Performance Prediction by Formalizing an Activity-Space-Performance Model,* Stanford University, Stanford, California.

Kimmance, A.G. (2002), An integrated product and process information modelling system for on-site construction, PhD, Loughborough University, Loughborough, UK.

Kiviniemi, A. (2005), Requirements management interface to building product models, PhD Thesis, Stanford University, Stanford, California.

Kiviniemi, M., Sulankivi, K., Kähkönen, K., Mäkelä, T. and Merivirta, M.-L. (2011), *BIM-based safety management and communication for building construction,* VTT, Finland.

Kolodner, J. (1993), *Case-Based Reasoning,* Morgan Kaufmann, San Mateo, CA.

Koo, B. and Fischer, M. (2000), "Feasibility study of 4D CAD in commercial construction", *Journal of Construction Engineering and Management,* Vol. 126 No. 4, pp. 251-60.

Kumar, V. and Viswanadham, N. (2007), "A CBR-based decision support system framework for construction supply chain risk management", in *Proceedings of 2007 IEEE International Conference on Automation Science and Engineering in Scottsdale, AZ, Sep 22-25, 2007,* IEEE, New York, pp. 980-85.

Kunz, J. and Fischer, M. (2007), "CIFE Research Questions and Methods - How CIFE Does Academic Research for Industrial Sponsors", available at: https://web.stanford.edu/class/cee320/CEE320A/ResMethods012307.pdf (accessed 15 July 2017).

Kwak, Y.H. and Stoddard, J. (2004), "Project risk management: lessons learned from software development environment", *Technovation*, Vol. 24 No. 11, pp. 915-20.

Laakso, M. and Kiviniemi, A. (2012), "The IFC standard - A review of history, development, and standardization", *Journal of Information Technology in Construction*, Vol. 17, pp. 134-61.

Lee, G., Park, H.K. and Won, J. (2012), "D3 City project — Economic impact of BIM-assisted design validation", *Automation in Construction*, Vol. 22, pp. 577-86.

Lee, J.K., Eastman, C.M., Lee, J., Kannala, M. and Jeong, Y.S. (2010), "Computing walking distances within buildings using the universal circulation network", *Environment and Planning B: Planning and Design*, Vol. 37 No. 4, pp. 628-45.

Lee, J.M. (2010), Automated checking of building requirements on circulation over a range of design phases, PhD Thesis, Georgia Institute of Technology, Atlanta, Georgia.

Lee, S.-I., Bae, J.-S. and Cho, Y.S. (2012), "Efficiency analysis of Set-based Design with structural building information modeling (S-BIM) on high-rise building structures", *Automation in Construction*, Vol. 23, pp. 20-32.

Li, H., Chan, G. and Skitmore, M. (2013), "Integrating real time positioning systems to improve blind lifting and loading crane operations", *Construction Management and Economics*, Vol. 31 No. 6, pp. 596-605.

Li, Q.F., Zhang, P. and Fu, Y.C. (2013), "Risk identification for the construction phases of the large bridge based on WBS-RBS", *Research Journal of Applied Sciences, Engineering and Technology*, Vol. 6 No. 9, pp. 1523-30.

Lincoln, Y.S., Lynham, S.A. and Guba, E.G. (2011), "Paradigmatic controversies, contradictions, and emerging confluences, revisited", *The Sage handbook of qualitative research*, Vol. 4, pp. 97-128.

Liu, G., Shen, Q., Li, H. and Shen, L. (2004), "Factors constraining the development of professional project management in China's construction industry", *International Journal of Project Management*, Vol. 22 No. 3, pp. 203-11.

Liu, W.P., Guo, H.L., Li, H. and Li, Y. (2014), "Using BIM to Improve the Design and Construction of Bridge Projects: A Case Study of a Long-span Steel-box Arch Bridge Project", *International Journal of Advanced Robotic Systems*, Vol. 11, pp. 125.

Locke, L.F., Silverman, S.J. and Spirduso, W.W. (2009), *Reading and understanding research*, Sage Publications, Thousand Oaks, California.

Lu, Y., Li, Q. and Xiao, W. (2013), "Case-based reasoning for automated safety risk analysis on subway operation: Case representation and retrieval", *Safety Science*, Vol. 57, pp. 75-81.

Lv, X. and El-Gohary, N.M. (2016a), "Enhanced context-based document relevance assessment and ranking for improved information retrieval to support environmental decision making", *Advanced Engineering Informatics*, Vol. 30 No. 4, pp. 737-50.

Lv, X. and El-Gohary, N.M. (2016b), "Semantic annotation for supporting context-aware information retrieval in the transportation project environmental review domain", *Journal of Computing in Civil Engineering,* Vol. 30 No. 6, pp. 04016033.

Manning, C.D., Raghavan, P. and Schütze, H. (2008), *Introduction to Information Retrieval,* Cambridge University Press, New York.

Markowski, A.S. and Mannan, M.S. (2008), "Fuzzy risk matrix", *Journal of hazardous materials,* Vol. 159 No. 1, pp. 152-57.

Marzouk, M. and Hisham, M. (2014), "Implementing earned value management using Bridge Information Modeling", *Ksce Journal of Civil Engineering,* Vol. 18 No. 5, pp. 1302-13.

McElwee, N. (2007), *At-risk children & youth: Resiliency explored*, The Haworth Integrative Healing Press, New York, NY, USA.

McLaughlin, J., Sreenivasan, S.V., Haas, C. and Liapi, K. (2004), "Rapid Human-Assisted Creation of Bounding Models for Obstacle Avoidance in Construction", *Computer-Aided Civil and Infrastructure Engineering,* Vol. 19 No. 1, pp. 3-15.

Mehdizadeh, R., Breysse, D., Taillandier, F. and Niandou, H. (2013), "Dynamic and multi perspective risk management in construction with a special view to temporary structures", *Civil Engineering and Environmental Systems,* Vol. 30 No. 2, pp. 115-29.

Mikolov, T., Chen, K., Corrado, G. and Dean, J. (2013a), *Efficient estimation of word representations in vector space,* http://arxiv.org/abs/1301.3781.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J. (2013b), *Distributed representations of words and phrases and their compositionality,* http://arxiv.org/abs/1310.4546.

Miller, R., Lessard, D.R., Michaud, P. and Floricel, S. (2001), *The strategic management of large engineering projects: Shaping institutions, risks, and governance,* MIT press, Massachusetts, USA.

Mitchell, D. (2012), "5D BIM: Creating cost certainty and better buildings", in *Proceedings of RICS COBRA 2012 (the annual RICS international research conference) in Las Vegas, Nevada, USA,* pp. 1-9.

Molenaar, K.R., Ashley, D.B. and Diekmann, J.E. (2006), *Guide to risk assessment and allocation for highway construction management,* Federal Highway Administration, US Deportment of Transportation, Washington, D.C.

Motamedi, A., Hammad, A. and Asen, Y. (2014), "Knowledge-assisted BIM-based visual analytics for failure root cause detection in facilities management", *Automation in Construction,* Vol. 43, pp. 73-83.

Motawa, I. and Almarshad, A. (2013), "A knowledge-based BIM system for building maintenance", *Automation in Construction,* Vol. 29, pp. 173-82.

Murty, M.N. and Jain, A.K. (1995), "Knowledge-based clustering scheme for collection management and retrieval of library books", *Pattern recognition,* Vol. 28 No. 7, pp. 949-63.

Myers, M.D. and Avison, D. (2002), *Qualitative research in information systems: a reader,* Sage, London.

Newman, I. and Benz, C.R. (1998), *Qualitative-quantitative research methodology: Exploring the interactive continuum,* Southern Illinois University Press (SIU) Press, Carbondale, Illinois.

NIBS. (2007), *United States National Building Information Modeling Standard, Version 1 - Part 1: Overview, Principles, and Methodologies,* National Institute of Building Sciences (NIBS), Washington, DC.

NIOSH (2016), "National Institute for Occupational Safety and Health (NIOSH) database", available at: http://www.cdc.gov/niosh/ (accessed 25 March 2016).

NLTK (2016), "Natural Language Toolkit", available at: www.nltk.org (accessed 5 April 2016).

OECD. (2002), *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development, The Measurement of Scientific and Technological Activities,* Organisation for Economic Co-operation and Development (OECD).

Olson, D.L. and Delen, D. (2008), *Advanced data mining techniques,* Springer-Verlag Berlin Heidelberg, Germany.

OSHA (2016), "Occupational Safety and Health Administration (OSHA) database", available at: www.osha.gov (accessed 1 April 2016).

OSHC (2008), *Occupational Safety and Health Statistics Bulletin Issue No. 7,* Occupational Safety and Health Council (OSHC), Hong Kong.

Pereira, F.C., Rodrigues, F. and Ben-Akiva, M. (2013), "Text analysis in incident duration prediction", *Transportation Research Part C: Emerging Technologies,* Vol. 37, pp. 177-92.

Perkins, J. (2014), *Python 3 Text Processing with NLTK 3 Cookbook,* second ed., Packt Publishing Ltd, Birmingham, UK.

PMI. (2004), *A guide to the project management body of knowledge (PMBOK Guide),* Project Management Institute (PMI), Newtown, Pennsylvania.

Porwal, A. and Hewage, K.N. (2013), "Building Information Modeling (BIM) partnering framework for public construction projects", *Automation in Construction,* Vol. 31, pp. 204-14.

Qi, J., Issa, R.R.A., Hinze, J. and Olbina, S. (2011), "Integration of safety in design through the use of building information modeling", in *Proceedings of 2011 ASCE International Workshop on Computing in Civil Engineering in Miami, FL, United States, 19-22 June 2011,* American Society of Civil Engineers, pp. 698-705.

Rajabifard, A., Williamson, I. and Kalantari, M. (2012), *A National Infrastructure for Managing Land Information—Research Snapshot,* The University of Melbourne, Victoria, Australia.

Rasdorf, W.J. and Lakmazaheri, S. (1990), "Logic-based approach for modeling organization of design standards", *Journal of Computing in Civil Engineering,* Vol. 4 No. 2, pp. 102-23.

Ryall, M.J., Parke, G.A. and Harding, J.E. (2000), *The manual of bridge engineering,* Thomas Telford, London, UK.

Sacks, R. and Barak, R. (2008), "Impact of three-dimensional parametric modeling of buildings on productivity in structural engineering practice", *Automation in Construction,* Vol. 17 No. 4, pp. 439-49.

Sacks, R., Rozenfeld, O. and Rosenfeld, Y. (2009), "Spatial and temporal exposure to safety hazards in construction", *Journal of Construction Engineering and Management,* Vol. 135 No. 8, pp. 726-36.

Salton, G., Wong, A. and Yang, C.-S. (1975), "A vector space model for automatic indexing", *Communications of the ACM,* Vol. 18 No. 11, pp. 613-20.

Schank, R.C. (1983), *Dynamic memory: A theory of reminding and learning in computers and people,* Cambridge University Press, New York.

Schank, R.C., Kass, A. and Riesbeck, C.K. (2014), *Inside case-based explanation,* Psychology Press, New York.

Scikit-learn (2016), "Scikit-learn toolkit", available at: www.scikit-learn.org/stable/ (accessed 15 April 2016).

Shim, C.-S., Lee, K.-M., Kang, L.S., Hwang, J. and Kim, Y. (2012), "Three-Dimensional Information Model-Based Bridge Engineering in Korea", *Structural Engineering International,* Vol. 22 No. 1, pp. 8-13.

Sigmund, Z. and Radujković, M. (2014), "Risk Breakdown Structure for Construction Projects on Existing Buildings", *Procedia - Social and Behavioral Sciences,* Vol. 119, pp. 894-901.

Smith, M.F. (1991), *Software prototyping: adoption, practice and management, McGraw-Hill software engineering series,* McGraw-Hill, London.

Smith, P. (2014), "BIM & the 5D Project Cost Manager", *Procedia - Social and Behavioral Sciences,* Vol. 119, pp. 475-84.

Snasel, V., Moravec, P. and Pokorny, J. (2005), "WordNet ontology based model for web retrieval", in *Proceedings of International Workshop on Challenges in Web Information Retrieval and Integration in Tokyo, Japan, 8-9 April 2005,* IEEE, pp. 220-25.

Soibelman, L., Wu, J., Caldas, C., Brilakis, I. and Lin, K.-Y. (2008), "Management and analysis of unstructured construction data types", *Advanced Engineering Informatics,* Vol. 22 No. 1, pp. 15-27.

Solihin, W. (2004), *Achieving Automated Code Checking with Ease - a white paper,* novaCITYNETS Pte Ltd, Singapore.

Sparck Jones, K. (1972), "A statistical interpretation of term specificity and its application in retrieval", *Journal of documentation,* Vol. 28 No. 1, pp. 11-21.

Steuer, J. (1992), "Defining virtual reality: Dimensions determining telepresence", *Journal of communication,* Vol. 42 No. 4, pp. 73-93.

Structural-Safety (1997), "Scour and flood damage to bridges", available at: http://www.structural-safety.org/publications/view-report/?report=2157 (accessed 31 December 2015).

Structural-Safety (2016), "Structural Safety database", available at: http://www.structural-safety.org/ (accessed 20 March 2016).

Sulankivi, K., Zhang, S., Teizer, J., Eastman, C.M., Kiviniemi, M., Romo, I. and Granholm, L. (2013), "Utilization of BIM-based Automated Safety Checking in Construction Planning", in *Proceedings of the 19th International CIB World Building Congress in Brisbane, Australia, 5-9 May 2013*, Queensland University of Technology, Brisbane, Australia, pp. 1-11.

Suresh, P., Babar, A. and Raj, V.V. (1996), "Uncertainty in fault tree analysis: a fuzzy approach", *Fuzzy Sets and Systems*, Vol. 83 No. 2, pp. 135-41.

Synchro (2016), "Synchro PRO", available at: https://synchroltd.com/synchro-pro/ (accessed 07 November 2016).

Tah, J. and Carr, V. (2001), "Knowledge-based approach to construction project risk management", *Journal of Computing in Civil Engineering*, Vol. 15 No. 3, pp. 170-77.

Tah, J.H.M., Carr, V. and Howes, R. (1999), "Information modelling for case-based construction planning of highway bridge projects", *Advances in Engineering Software*, Vol. 30 No. 7, pp. 495-509.

Tam, C.M., Zeng, S.X. and Deng, Z.M. (2004), "Identifying elements of poor construction safety management in China", *Safety Science*, Vol. 42 No. 7, pp. 569-86.

Tang, P., Anil, E., Akinci, B. and Huber, D. (2011), "Efficient and effective quality assessment of as-is building information models and 3D laser-scanned data", in *Proceedings of the ASCE International Workshop on Computing in Civil Engineering in Miami, Florida, United States, 19-22 June 2011*, American Society of Civil Engineers, New York, pp. 486-93.

Teizer, J. (2008), "3D range imaging camera sensing for active safety in construction", *Journal of Information Technology in Construction*, Vol. 13, pp. 103-17.

Teizer, J., Allread, B.S., Fullerton, C.E. and Hinze, J. (2010), "Autonomous pro-active real-time construction worker and equipment operator proximity safety alert system", *Automation in Construction*, Vol. 19 No. 5, pp. 630-40.

Teizer, J., Caldas, C.H. and Haas, C.T. (2007), "Real-time three-dimensional occupancy grid modeling for the detection and tracking of construction resources", *Journal of Construction Engineering and Management*, Vol. 133 No. 11, pp. 880-88.

Teizer, J., Cheng, T. and Fang, Y. (2013), "Location tracking and data visualization technology to advance construction ironworkers' education and training in safety and productivity", *Automation in Construction*, Vol. 35, pp. 53-68.

Tekla (2016), "Tekla Structures", available at: https://www.tekla.com/uk/products/tekla-structures (accessed 07 November 2016).

Tixier, A.J.-P., Hallowell, M.R., Rajagopalan, B. and Bowman, D. (2016), "Automated content analysis for construction safety: A natural language processing system to extract precursors and outcomes from unstructured injury reports", *Automation in Construction,* Vol. 62, pp. 45-56.

Vechtomova, O. and Wang, Y. (2006), "A study of the effect of term proximity on query expansion", *Journal of Information Science,* Vol. 32 No. 4, pp. 324-33.

Vico Software (2017), "Vico Control", available at: http://www.vicosoftware.com/products/vico_control/control-features--benefits-2009-copy (accessed 6 March 2017).

Volk, R., Stengel, J. and Schultmann, F. (2014), "Building Information Modeling (BIM) for existing buildings—Literature review and future needs", *Automation in Construction,* Vol. 38, pp. 109-27.

Vrouwenvelder, A., Holicky, B., Tanner, C., Lovegrove, D. and Canisius, E. (2001), "Risk assessment and risk communication in civil engineering", in *Safety, Risk, Reliability - Trends in Engineering in Malta*, CIB Report Publication, pp. 885-90.

Wang, B., Li, H., Rezgui, Y., Bradley, A. and Ong, H.N. (2014), "BIM based virtual environment for fire emergency evacuation", *The Scientific World Journal,* Vol. 2014, pp. 589016.

Wang, M.T. and Chou, H.Y. (2003), "Risk allocation and risk handling of highway projects in Taiwan", *Journal of Management in Engineering,* Vol. 19 No. 2, pp. 60-68.

Wardhana, K. and Hadipriono, F.C. (2003), "Analysis of recent bridge failures in the United States", *Journal of performance of constructed facilities,* Vol. 17 No. 3, pp. 144-50.

Whyte, J. (2000), Virtual reality applications in the house-building industry, PhD Thesis, Loughborough University.

Whyte, J. (2002), *Virtual reality and the built environment,* Architectural Press, Oxford, UK.

Wikipedia (2016), "Wikipedia content database (English)", available at: meta.wikimedia.org/wiki/Data_dumps (accessed 10 April 2016).

Williams, T.M. (1996), "The two-dimensionality of project risk", *International Journal of Project Management,* Vol. 14 No. 3, pp. 185-86.

WorkSafeBC (2016), "WorkSafeBC database", available at: www.worksafebc.com/en (accessed 28 March 2016).

Wu, H., Tao, J., Li, X., Chi, X., Li, H., Hua, X., Yang, R., Wang, S. and Chen, N. (2013), "A location based service approach for collision warning systems in concrete dam construction", *Safety Science,* Vol. 51 No. 1, pp. 338-46.

Yabuki, N. and Li, Z. (2006), "Development of new IFC-BRIDGE data model and a concrete bridge design system using multi-agents", in *Proceedings of Intelligent Data Engineering and Automated Learning – IDEAL 2006 in Burgos, Spain, September 20-23, 2006*, Springer-Verlag, Berlin, Germany, pp. 1259-66.

Yang, Q. and Xu, X. (2004), "Design knowledge modeling and software implementation for building code compliance checking", *Building and Environment,* Vol. 39 No. 6, pp. 689-98.

Yi, K.-J. and Langford, D. (2006), "Scheduling-based risk estimation and safety planning for construction projects", *Journal of Construction Engineering and Management,* Vol. 132 No. 6, pp. 626-35.

Yu, H. (2009), A knowledge based system for construction health and safety competence assessment, PhD Thesis, University of Wolverhampton, Wolverhampton, England.

Zayed, T., Amer, M. and Pan, J. (2008), "Assessing risk and uncertainty inherent in Chinese highway projects using AHP", *International Journal of Project Management,* Vol. 26 No. 4, pp. 408-19.

Zhang, J. and El-Gohary, N.M. (2017), "Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking", *Automation in Construction,* Vol. 73, pp. 45-57.

Zhang, J.P. and Hu, Z.Z. (2011), "BIM- and 4D-based integrated solution of analysis and management for conflicts and structural safety problems during construction: 1. Principles and methodologies", *Automation in Construction,* Vol. 20 No. 2, pp. 155-66.

Zhang, L., Wu, X., Skibniewski, M.J., Zhong, J. and Lu, Y. (2014), "Bayesian-network-based safety risk analysis in construction projects", *Reliability Engineering & System Safety,* Vol. 131, pp. 29-39.

Zhang, S.J., Teizer, J., Lee, J.K., Eastman, C.M. and Venugopal, M. (2013), "Building Information Modeling (BIM) and Safety: Automatic Safety Checking of Construction Models and Schedules", *Automation in Construction,* Vol. 29, pp. 183-95.

Zhang, X., Deng, Y., Li, Q., Skitmore, M. and Zhou, Z. (2016), "An incident database for improving metro safety: The case of shanghai", *Safety Science,* Vol. 84, pp. 88-96.

Zhao, J., Cui, L., Zhao, L., Qiu, T. and Chen, B. (2009), "Learning HAZOP expert system by case-based reasoning and ontology", *Computers & Chemical Engineering,* Vol. 33 No. 1, pp. 371-78.

Zhao, J.J. and Tonias, D.E. (2012), *Bridge engineering : design, rehabilitation, and maintenance of Modern Highway Bridges,* McGraw-Hill Professional, New York.

Zhou, W., Heesom, D., Georgakis, P., Nwagboso, C. and Feng, A. (2009), "An interactive approach to collaborative 4D construction planning", *Journal of Information Technology in Construction,* Vol. 14, pp. 30-47.

Zhou, W., Whyte, J. and Sacks, R. (2012), "Construction safety and digital design: A review", *Automation in Construction,* Vol. 22, pp. 102-11.

Zou, P.X.W., Zhang, G. and Wang, J. (2007), "Understanding the key risks in construction projects in China", *International Journal of Project Management,* Vol. 25 No. 6, pp. 601-14.

# APPENDICES

## Appendix A. Knowledge-based Risk Database

### A.1 External risks

Table 10. Knowledge-based risk database – external risks

| Risk Category | Risk Factor | Risk Description | Possible Mitigation Strategy | Reference |
|---|---|---|---|---|
| Economic | Inflation | Price inflation of construction materials; Monetary inflation; Unanticipated local inflation and interest rates due to immature local economic and banking systems; Increase of wages and welfare | 1) Escalation Clause;<br>2) Price Contingency in the Bid;<br>3) Project Financing by a Reputable Owner;<br>4) Owner Purchase of Equipment & Material;<br>5) Providing Performance Bond and Prequalification of Suppliers;<br>6) Forward Contracts for Hedging Exchange Rate Changes | (Casey, 1979, Mustafa and Al-Bahar, 1991, Al‐Bahar and Crandall, 1990, Kangari, 1995, Tang *et al.*, 2007, Tah *et al.*, 1993, Li *et al.*, 2013, Hastak and Shaked, 2000, Zou *et al.*, 2007, Wang *et al.*, 2004, Fang *et al.*, 2004, Kartam and Kartam, 2001, Wang and Chou, 2003) |
| | Currency | Rate fluctuation; devaluation; difficulty in converting foreign currency | | (Casey, 1979, Mustafa and Al-Bahar, 1991, Al‐Bahar and Crandall, 1990, Kangari, 1995, Tang *et al.*, 2007, Tah *et al.*, 1993, Wang *et al.*, 2004, Ibrahim, 2011, Zayed *et al.*, 2008) |
| | National and international impacts | Changes by International Associations such as OPEC | | (Casey, 1979, Mustafa and Al-Bahar, 1991, Al‐Bahar and Crandall, 1990, Kangari, 1995, Tang *et al.*, 2007, Tah *et al.*, 1993, Hastak and Shaked, 2000) |
| | Inadequate market demand | Inadequate forecast of market demand; Owners' unreasonable upfront capital demand | | (Casey, 1979, Mustafa and Al-Bahar, 1991, Al‐Bahar and |

| | | | | |
|---|---|---|---|---|
| | | | | Crandall, 1990, Kangari, 1995, Tang *et al.*, 2007, Tah *et al.*, 1993, Fang *et al.*, 2004, Wang *et al.*, 2004) |
| Political | Inappropriate use of power | Bureaucracy; Lack of legal judgment reinforcement; Problems of the construction examination and approval procedure; Delay or refusal of project approval and permit by local government; Delay in design and regulatory approval; Delays in preparation of submittals; Delays in obtaining no object certificates (NOCs) from authorities; Nationalization or Expropriation; Inadequate claim administration; Unnecessary and unjust Government influence on disputes; Unnecessary and unjust intervention; Unwelcome attitude toward foreign investor and profit; Failure in obtaining fair import/export quota allocation from local government; Government incentives; unexpected disputes or strikes; disruption to power or utilities supplies; Bribes and corruption | - | (Casey, 1979, Mustafa and Al-Bahar, 1991, El-Sayegh and Mansour, 2015, Tang *et al.*, 2007, Fang *et al.*, 2004, Andi, 2006, Tah *et al.*, 1993, Bing *et al.*, 2005, Rezakhani, 2012, Hastak and Shaked, 2000, Al‐Bahar and Crandall, 1990, Zayed *et al.*, 2008, Kangari, 1995, Wang *et al.*, 2004, Li *et al.*, 2013, Ibrahim, 2011, Kartam and Kartam, 2001, Zou *et al.*, 2007) |
| | Stability of nation | Public disorder and war; Frequent changes in government; agitation for change of government or disputes between political parties or different organs of the state; Unstable relation to neighbouring countries or regions | | (Casey, 1979, Mustafa and Al-Bahar, 1991, El-Sayegh and Mansour, 2015, Tang *et al.*, 2007, Fang *et al.*, 2004, Andi, 2006, Tah *et al.*, 1993, Bing *et al.*, 2005, Zayed *et al.*, 2008, Wang *et al.*, 2004, Al‐Bahar and Crandall, 1990, Li *et al.*, 2013, Kartam and Kartam, 2001) |
| | Changes or deficiencies in laws and regulations | <u>1) Deficiencies:</u><br>Imperfection of safety or labour law; Stringent regulation which will have an impact on construction firms' poor attention to environmental issues; Regional protection policy;<br><u>2) Changes or variations:</u><br>Local government's inconsistent application of new regulations and laws; mandatory joint venture (JV); mandatory technology transfer; differential taxation of foreign firms; Embargoes; changes in government or policy; changes to supply of oil and energy | | (Casey, 1979, Mustafa and Al-Bahar, 1991, El-Sayegh and Mansour, 2015, Tang *et al.*, 2007, Fang *et al.*, 2004, Andi, 2006, Tah *et al.*, 1993, Bing *et al.*, 2005, Rezakhani, 2012, Hastak and Shaked, 2000, Wang *et al.*, 2004, Al‐Bahar and Crandall, 1990, Kangari, 1995, Li *et al.*, 2013, Wang and Chou, 2003) |

| | | | | |
|---|---|---|---|---|
| | Third party pressure | Pressure from environment protection group | | (Casey, 1979, Mustafa and Al-Bahar, 1991, El-Sayegh and Mansour, 2015, Tang *et al.*, 2007, Fang *et al.*, 2004, Andi, 2006, Tah *et al.*, 1993, Bing *et al.*, 2005) |
| Social & Cultural | Cultural difference and conflicts | Interaction of foreign management with local contractors; Differences in work culture, education, values, language, racial prejudice, etc., between foreign and local partners | - | (Bing *et al.*, 2005, El-Sayegh and Mansour, 2015, Wang *et al.*, 2004, Zayed *et al.*, 2008) |
| | Social issues | 1) Corrupt social morality;<br>2) Criminal acts;<br>3) Corporate fraud:<br>Unexpected increases in turnover, unexpected resignation of financial adviser, letter of credits with 'unreasonably round figures', intentional or unintentional negligence either by auditors, bankers or creditors | | (Bing *et al.*, 2005, Wang *et al.*, 2004, El-Sayegh and Mansour, 2015, Li *et al.*, 2013) |
| | Serious competition | Competition from other international investors, developers or contractors; competition in limited market volume | | (Bing *et al.*, 2005, Zayed *et al.*, 2008, Wang *et al.*, 2004) |

## A.2 Internal risks

Table 11. Knowledge-based risk database – internal risks

| Risk Category | Risk Factor | Risk Description | Possible Mitigation Strategy | Reference |
|---|---|---|---|---|
| Structural | Temporary structure damage or collapse | *<u>1) Design:</u>*<br>Defective or inappropriate selection of needed temporary structures; Design of temporary structures does not comply with codes and standards; Design process does not include sensitivity study to assess the impact of natural hazards; Designed temporary structures are not compatible with natural constraints of the site; Designed temporary structures are not executable; Poor design check by consultant;<br>*<u>2) Construction:</u>*<br>Poor preliminary assessment and evaluation of possible forms of contract for execution of temporary structures;<br>Allocation of risks related to temporary structures is not clear; Damage to installed temporary structures during the work due to poor performance of workers; Error in execution of temporary structures due to poor performance of experts and workers; Use of overweight material or equipment; Impropriate (deep) excavations; Scaffolds and traffic protection risks;<br>*<u>3) Operation:</u>*<br>Poor maintenance of temporary structures;<br>*<u>4) Management:</u>*<br>Consultant is not informed about the changes in temporary structures; Contractor with lack of needed temporary structures; Disregarding to sequential and staged | 1) at planning stage, any assessment of risk should identify 'safety critical' considerations;<br>2) strengthen appropriate supervision and post-concreting checks;<br>3) method statement compiled by a contractor;<br>4) long-term risk strategy;<br>5) temporary Motorway barriers for protecting scaffolds;<br>6) designers (and CDM co-ordinators) regard getting adequate strength as their key functions;<br>7) strengthen regular check and inspection | (Casey, 1979, El-Sayegh and Mansour, 2015, Tang *et al.*, 2007, Fang *et al.*, 2004, Wang and Chou, 2003, Mehdizadeh *et al.*, 2013, Structural-Safety, 1997c, Structural-Safety, 2015l, Structural-Safety, 1980, Structural-Safety, 2013, Structural-Safety, 2004, Structural-Safety, 2015q, Structural-Safety, 1974, Structural-Safety, 2015p, Structural-Safety, 2006b, Structural-Safety, 2015k, TranSystems, 2012) |

| | | | | |
|---|---|---|---|---|
| | | activities of remedial action by contractor; Failure of temporary structures due to irregular control and inspection (e.g. ground anchors, walls); | | |
| | Permanent structure damage or collapse | <u>*1) Design:*</u><br>Design process does not include sensitivity study to assess the impact of natural hazards; Defective or inappropriate selection of needed temporary structures by designers; Design of permanent structures does not comply with codes and standards; Poor design check by consultant;<br><u>*2) Construction:*</u><br>Damage to installed permanent structures during the work due to poor performance of workers; Error in execution of permanent structures due to poor performance of experts and workers; Use of overweight material or equipment; Impropriate (deep) excavations; Natural forces (e.g. scour and flood);<br><u>*3) Operation:*</u><br>Irregular or inadequacy of inspection of permanent structures; Poor maintenance of permanent structures (e.g. concrete half joints); Natural forces (e.g. scour and flood); fatigue of steel; Thaumasite sulphate attack; Dynamic behaviour of bridges under pedestrian loading<br><u>*4) Management:*</u><br>Disregarding to sequential and staged activities of remedial action by contractor; | | (Al‐Bahar and Crandall, 1990, Casey, 1979, El-Sayegh and Mansour, 2015, Tang *et al.*, 2007, Fang *et al.*, 2004, Wang and Chou, 2003, Mehdizadeh *et al.*, 2013, Structural-Safety, 1980, Structural-Safety, 2013, Structural-Safety, 2004, Structural-Safety, 2015q, Structural-Safety, 2007, Structural-Safety, 2015g, Structural-Safety, 2015m, Structural-Safety, 2006a, Structural-Safety, 2015f, Structural-Safety, 2006b, Structural-Safety, 2015c, Structural-Safety, 2015d, Structural-Safety, 2008c, Structural-Safety, 2008a, Structural-Safety, 2009a, Structural-Safety, 2015j, Structural-Safety, 2015e, Structural-Safety, 2009b, Structural-Safety, 1999e, Structural-Safety, 2015a, Structural-Safety, 2015o, Structural-Safety, 2015r, Structural-Safety, 2008b, Structural-Safety, 1997b, Structural-Safety, 1997h, Structural-Safety, 1997f, Structural-Safety, 1997d, Structural-Safety, 1999c, Structural-Safety, 1999b, Structural-Safety, 1999d, Structural-Safety, 1999f, Structural-Safety, 2000, SEStran, 2008, Thames-Water, 2013) |

| | | | | |
|---|---|---|---|---|
| | Auxiliary elements damage or collapse | Unsafe light poles or advertisement board bases; Unsafe lamp posts | | (Structural-Safety, 1999g, Structural-Safety, 1999h, Structural-Safety, 2006b) |
| | Durability and safety | Potential risks of post-tensioned concrete bridge; predicted effects of increased traffic loads; fatigue of steel; Thaumasite sulphate attack; dynamic behaviour of bridges under pedestrian loading; washing-out of timber structure | | (Structural-Safety, 1997g, Structural-Safety, 1997a, Structural-Safety, 1997d, Structural-Safety, 1997e, Structural-Safety, 1999a, Structural-Safety, 1999f, Structural-Safety, 2000, SEStran, 2008) |
| Personnel Safety | Falling from height | Working at height; Disregard to different warning signs; Inappropriate personnel protective equipment; Incompetency of personnel working with tower crane, ground freezing system | 1) strengthen safety training | (Rezakhani, 2012, HSE, 2015, Mehdizadeh *et al.*, 2013) |
| | Being struck or crushed | Vehicles and other transport; Disregard to different warning signs; Inappropriate personnel protective equipment; Incompetency of personnel working with tower crane, ground freezing system | | (Structural-Safety, 1999g, Structural-Safety, 1999h, Rezakhani, 2012, HSE, 2015, Mehdizadeh *et al.*, 2013) |
| | Electrocution | Power cables and electrical installations; Disregard to different warning signs; Inappropriate personnel protective equipment; Incompetency of personnel working with tower crane, ground freezing system | | (Rezakhani, 2012, HSE, 2015, Mehdizadeh *et al.*, 2013) |
| | Fire | Disregard to different warning signs; Inappropriate personnel protective equipment; Incompetency of personnel working with tower crane, ground freezing system | | (Rezakhani, 2012, HSE, 2015, Mehdizadeh *et al.*, 2013) |
| | Collapse | Collapse of excavations; Collapse of structures (e.g. walls, cranes, scaffolds); Disregard to different warning signs; Inappropriate personnel protective equipment; Incompetency of personnel working with tower crane, ground freezing system | | (Structural-Safety, 1999g, Structural-Safety, 1999h, Rezakhani, 2012, HSE, 2015, Mehdizadeh *et al.*, 2013, Structural-Safety, 2015l, Structural-Safety, 1997c, Structural-Safety, 2015q, Structural-Safety, 1974, |

| | | | | |
|---|---|---|---|---|
| | | | | Structural-Safety, 2008a, Structural-Safety, 2015j) |
| | Exposure | Exposure to building dusts; Exposure to asbestos; Disregard to different warning signs; Inappropriate personnel protective equipment; Incompetency of personnel working with tower crane, ground freezing system | | (Rezakhani, 2012, HSE, 2015, Mehdizadeh *et al.*, 2013) |
| | Others | Unsuitable protection measures of construction safety; Inadequate safety measures or unsafe operations; Unavailability of sufficient professionals and managers; Labour dispute and strike; Poor coordination; Disregard to different warning signs; Careless barge driving; Anti-social behaviours | | (Bing *et al.*, 2005, Rezakhani, 2012, El-Sayegh and Mansour, 2015, Tang *et al.*, 2007, Ibrahim, 2011, Kartam and Kartam, 2001, Mustafa and Al-Bahar, 1991, Kangari, 1995, Andi, 2006, Zou *et al.*, 2007, HSE, 2015, Casey, 1979, Al - Bahar and Crandall, 1990, Fang *et al.*, 2004, Wang and Chou, 2003, Li *et al.*, 2013, Hastak and Shaked, 2000, Mehdizadeh *et al.*, 2013, Choi and Mahadevan, 2008, Therrien, 2011, SEStran, 2008) |
| Personnel Health | Musculo-skeletal | This is one of the most common causes of ill health; lifting heavy weights | 1) Designers should consider lifting (e.g. choice of unit size), operating space, and the ergonomics of relevant activities; 2) Designers can obtain useful advice from contractors and suppliers of equipment. | (HSE, 2015) |
| | Noise-induced hearing loss and hand-arm and whole body vibration | Current Regulations require significant reductions in the level of exposure to workers from those previously tolerated. If noisy or vibration-prone activities unnecessarily result from the design, this may result in additional project costs. | | (HSE, 2015) |
| | Dermatitis and other skin-related problems | Designers should consider whether there are alternatives to materials or processes which cause particular problems. | | (HSE, 2015) |
| | Asbestos-related diseases | This is a major issue on refurbishment projects. Influence can be exerted through adequate information provision and careful consideration of survey information and the management plan. | | (HSE, 2015) |

| | | | | |
|---|---|---|---|---|
| | Others | Project specific, e.g. presence of vermin and bird excreta, specific materials, dusts, sprays, contaminated land, lead. | | (HSE, 2015) |
| Design | Unqualified or defective design | Insufficient planning; Incomplete design scope; Difficult for construction; Improper site estimation; Improper material use; Lack of experience and knowledge in design; Inadequate specifications; Defective or inappropriate selection of needed temporary structures by designers; Design of temporary structures does not comply with codes and standards; Disregard to different warning signs; Poor assessment and evaluation of different options for temporary structure; Poor communication of designer and contractor to control the suitability and constructability of temporary structures; Poor design check by consultant; Unproven technical design of temporary structures accepted by consultant; Using inadequate software for design of temporary structures; Poor design due to disregarding to the proximity of existing services; Insufficient or incomplete Detailing | 1) Changed Condition Clause (Delay);<br>2) Contractor Participates in Design;<br>3) Adoptable Design/ Construction Methods;<br>4) Changes to the Original Design | (Bing *et al.*, 2005, Li *et al.*, 2013, El-Sayegh and Mansour, 2015, Al－Bahar and Crandall, 1990, Mustafa and Al-Bahar, 1991, Hastak and Shaked, 2000, Kangari, 1995, Andi, 2006, Ibrahim, 2011, Zayed *et al.*, 2008, Tang *et al.*, 2007, Kartam and Kartam, 2001, Wang *et al.*, 2004, Mehdizadeh *et al.*, 2013, Choi and Mahadevan, 2008, Mulholland and Christian, 1999, Structural-Safety, 2015h, Structural-Safety, 2015r) |
| | Errors and mistakes | Carelessness; Lack of experience and knowledge in design; Inadequate specifications; Incorrect quantity calculation; competence; Designed temporary structures are not executable; Incorrect definition of type and quantity of needed temporary structures; Poor communication of designer and contractor to control the suitability and constructability of temporary structures; Poor design check by consultant; Unproven technical design of temporary structures accepted by consultant; Using inadequate software for design of temporary structures | | (Bing *et al.*, 2005, Al－Bahar and Crandall, 1990, Mustafa and Al-Bahar, 1991, Hastak and Shaked, 2000, Kangari, 1995, Andi, 2006, Ibrahim, 2011, Zayed *et al.*, 2008, Tang *et al.*, 2007, Kartam and Kartam, 2001, Wang *et al.*, 2004, Mehdizadeh *et al.*, 2013, Mulholland and Christian, 1999, Structural-Safety, 2015o, Structural-Safety, 2015b, Structural-Safety, 2015h) |
| | Design changes and rework | Incorrect definition of type and quantity of needed temporary structures; Poor assessment and evaluation of different options for | | (Bing *et al.*, 2005, Rezakhani, 2012, Mustafa and Al-Bahar, 1991, Kangari, 1995, Fang *et al.*, 2004, |

| | | | | |
|---|---|---|---|---|
| | | temporary structure; Poor communication of designer and contractor to control the suitability and constructability of temporary structures; Poor design check by consultant; Poor design due to disregarding to the proximity of existing services | | Andi, 2006, Ibrahim, 2011, Mehdizadeh *et al.*, 2013) |
| | Delays of design works | Low productivity; Work order change; Delays in design and regulatory approval; Poor assessment and evaluation of different options for temporary structure; Poor communication of designer and contractor to control the suitability and constructability of temporary structures; Poor design check by consultant | | (Bing *et al.*, 2005, Rezakhani, 2012, Hastak and Shaked, 2000, Zayed *et al.*, 2008, Kartam and Kartam, 2001, Wang and Chou, 2003, Ibrahim, 2011, Mulholland and Christian, 1999) |
| Construction | Deviation between design and construction | Defective design and errors; Disregarding to sequential and staged activities of remedial action by contractor; Use of overweight material or equipment | 1) Contingency in the Bid;<br>2) Insurance for Liability from accidents;<br>3) Contract Clause for Time Extension Due to Delays;<br>4) Safety and Training Programmes from Employees;<br>5) Planning Procurement Activities in Advance | (Tah *et al.*, 1993, Al‐Bahar and Crandall, 1990, Bing *et al.*, 2005, Fang *et al.*, 2004, Hastak and Shaked, 2000, Kangari, 1995, Ibrahim, 2011, Casey, 1979, Tang *et al.*, 2007, Kartam and Kartam, 2001, El-Sayegh and Mansour, 2015, Andi, 2006, Zayed *et al.*, 2008, Mehdizadeh *et al.*, 2013, Choi and Mahadevan, 2008) |
| | Construction errors | Lack of appreciation of basic stability of structures | | (Structural-Safety, 2015b, Structural-Safety, 2015k) |
| | Inadequate construction planning | Half-baked consideration on the actual condition of the construction site; Unfamiliar with the design drawings and design intention; Insufficient site information and unforeseeable circumstances underground; Unreasonable personnel organization and arrangement; Unreasonable materials and unreasonable equipment allocation; Lack of knowledge and experience; Disregarding to sequential and staged activities of remedial action by contractor | | (Bing *et al.*, 2005, Li *et al.*, 2013, Tang *et al.*, 2007, Zayed *et al.*, 2008, Mehdizadeh *et al.*, 2013, Therrien, 2011) |
| | Improper construction method and scheme | Unfeasible construction methods; Lack of knowledge and experience | | (Li *et al.*, 2013, Tang *et al.*, 2007, Zayed *et al.*, 2008, El-Sayegh and |

| | | | | |
|---|---|---|---|---|
| | | | | Mansour, 2015, Ibrahim, 2011, Wang and Chou, 2003) |
| | Construction changes and delay | Third party delays; Delay of drawing supply; changes in work; Owner changes; Construction delay; Delayed site access; Late drawings and instructions; Delays in material supply; Improper intervention | | (Rezakhani, 2012, Fang *et al.*, 2004, Hastak and Shaked, 2000, Kangari, 1995, Ibrahim, 2011, Casey, 1979, Tang *et al.*, 2007, Kartam and Kartam, 2001, El-Sayegh and Mansour, 2015, Al - Bahar and Crandall, 1990, Andi, 2006, Zayed *et al.*, 2008, Mehdizadeh *et al.*, 2013) |
| | Poor construction quality | Unqualified workmanship and skills; Improper material use; Violating construction standards; Cutting corners | | (Casey, 1979, Li *et al.*, 2013, El-Sayegh and Mansour, 2015, Andi, 2006, Mustafa and Al-Bahar, 1991, Al - Bahar and Crandall, 1990, Zayed *et al.*, 2008, Choi and Mahadevan, 2008) |
| | Construction cost overrun | Cost of tests and samples | | (Bing *et al.*, 2005, Casey, 1979, Mulholland and Christian, 1999) |
| | Construction condition risks | Operation at or under river level; existing traffic; Unforeseeable or bad site conditions | | (Bing *et al.*, 2005, Li *et al.*, 2013, Tang *et al.*, 2007, Andi, 2006, Mustafa and Al-Bahar, 1991, Al - Bahar and Crandall, 1990, Kangari, 1995, Kartam and Kartam, 2001, Castles and Parish, 2011, Transport for London, 2013) |
| | Low construction productivity | Obsolete technology and practices by local partners; poor skills or inadequate supervision; Shortage of skilled and unskilled workers; Foreign firms face difficulties in hiring and keeping suitable and valuable employees; Insufficient labour; Productivity of equipment | | (Tah *et al.*, 1993, Casey, 1979, El-Sayegh and Mansour, 2015, Kangari, 1995, Kartam and Kartam, 2001, Mustafa and Al-Bahar, 1991, Ibrahim, 2011, Wang *et al.*, 2004, Andi, 2006, Hastak and Shaked, 2000, Zayed *et al.*, 2008, Mulholland and Christian, 1999) |
| | Improper project management | Improper project budgeting; Inadequate project organisation structure; Incompetence of local project team; Incompetence of subcontractor | | (Tang *et al.*, 2007, Wang *et al.*, 2004) |

| | | | | |
|---|---|---|---|---|
| | Failure to identify defects | Insufficient inspections | | (Ibrahim, 2011, Tang *et al.*, 2007) |
| | Environmental risks | Environmental pollution during due to poor inspection of temporary structures; Adverse weather conditions | | (El-Sayegh and Mansour, 2015, Hastak and Shaked, 2000, Al - Bahar and Crandall, 1990, Ibrahim, 2011, Kartam and Kartam, 2001, Rezakhani, 2012, Mehdizadeh *et al.*, 2013, Therrien, 2011, Transport for London, 2013, WSP, 2013, Gardenbridgetrust, 2014) |
| Material and equipment | Availability of resources or equipment | Short supply; Over-consumption in transportation, storage and construction; Restriction of the local transportation; | | (Tah *et al.*, 1993, Casey, 1979, Hastak and Shaked, 2000, Kangari, 1995, Tang *et al.*, 2007, Kartam and Kartam, 2001, Li *et al.*, 2013, Zayed *et al.*, 2008, Fang *et al.*, 2004, Ibrahim, 2011, Wang and Chou, 2003) |
| | Breakdown or failure | Premature failure of facility; Construction machinery breakdown and the power fault; Installation errors and debugging errors of the construction equipment; | | (Tah *et al.*, 1993, Casey, 1979, Hastak and Shaked, 2000, Kangari, 1995, Tang *et al.*, 2007, Kartam and Kartam, 2001, Li *et al.*, 2013, Zayed *et al.*, 2008, Fang *et al.*, 2004, Ibrahim, 2011, Wang and Chou, 2003) |
| | Incorrect use | Wrong type and quantity; Problems of using special and new materials; Inadequacy of the equipment maintenance or overloading operations of the construction equipment | | (Tah *et al.*, 1993, Casey, 1979, Hastak and Shaked, 2000, Kangari, 1995, Tang *et al.*, 2007, Kartam and Kartam, 2001, Li *et al.*, 2013, Zayed *et al.*, 2008, Fang *et al.*, 2004, Ibrahim, 2011, Wang and Chou, 2003) |
| | Incorrect operation | Instability of the construction equipment and unsafe operation | | (Tah *et al.*, 1993, Casey, 1979, Hastak and Shaked, 2000, Kangari, 1995, Tang *et al.*, 2007, Kartam and Kartam, 2001, Li *et al.*, 2013, Zayed *et al.*, 2008, Fang *et al.*, 2004, Ibrahim, 2011, Wang and Chou, 2003) |

| | | | | |
|---|---|---|---|---|
| | Disqualification | Defective material or equipment; | | (Tah *et al.*, 1993, Casey, 1979, Hastak and Shaked, 2000, Kangari, 1995, Tang *et al.*, 2007, Kartam and Kartam, 2001, Li *et al.*, 2013, Zayed *et al.*, 2008, Fang *et al.*, 2004, Ibrahim, 2011, Wang and Chou, 2003) |
| Physical | Unforeseeable adverse ground/underground conditions | Unexpected underground utilities; Archaeological finds; Unforeseen soil conditions; Poor terrain condition; underground river or water | | (Zayed *et al.*, 2008, El-Sayegh and Mansour, 2015, Ibrahim, 2011, Casey, 1979, Wang and Chou, 2003) |
| Organisational | Personnel | *1) Negative work attitude:*<br>Tight project schedule; Project funding problems; Contractors’ poor management ability; Lack of readily available utilities on site; Poor competency of labour; Labour dispute and strike; Poor coordination;<br>*2) Wages:*<br>Contractors’ difficulty in reimbursement;<br>*3) Insurance:*<br>Employees do not have safety insurance;<br>*4) Collaboration:*<br>Lack of good communication | | (Rezakhani, 2012, van Well-Stam *et al.*, 2004, El-Sayegh and Mansour, 2015, Tang *et al.*, 2007, Ibrahim, 2011, Kartam and Kartam, 2001, Mustafa and Al-Bahar, 1991, Kangari, 1995, Andi, 2006, Zou *et al.*, 2007, Bing *et al.*, 2005) |
| | Organisational regulations | Lack of quality plan | | (van Well-Stam *et al.*, 2004) |
| | Organisational management | *1) Lack of capability:*<br>Failure to take projects in the area partially or fully into account; Inaccuracy and incompleteness in the estimate; Incompleteness or carelessness in the drawing up of contractual documents;<br>*2) Late response:*<br>Late ordering materials;<br>*3) Lack of necessary manpower at a certain point:*<br>Problems in setting up and organizing project organization; withdrawal of key individuals; modifications in project staffing;<br>*4) Lack of clarity on:*<br>Requirements; project limits; | | (van Well-Stam *et al.*, 2004, Choi and Mahadevan, 2008, Rezakhani, 2012) |

| | | | | |
|---|---|---|---|---|
| | | *5) Lack of project procedures:* Modifications procedure, POR, planning, estimates; completion and acceptance procedures; Administrative organization (AO) procedures; tender offer plan or procedure; award procedure; *6) Difficulties in commercial management:* Failure to enter into agreements with the parties concerned; Insufficient link between subprojects | | |
| | Organisational changes | *Modifications in the programme of requirements (POR):* Lack of clarity on basic principles; changes in project definition | | (Rezakhani, 2012, van Well-Stam *et al.*, 2004) |
| Natural | Natural (Acts of God) | Flood; Earthquake; Fire; Wind damage; Lighting; Collapse and Landslide; Storm; Epidemic diseases; Heavy snow; Extreme high temperature; Volcanic cloud; Thaumasite sulphate attack | 1) Insurance Carried by Owner 2) Contractual Clauses for delay and Payments for Incurred Damages 3) Contingency Plan 4) monitoring behaviour of structures in risk zones 5) mitigation strategy for erosion of base due to intensive rain | (Bing *et al.*, 2005, Mustafa and Al-Bahar, 1991, Fang *et al.*, 2004, Casey, 1979, El-Sayegh and Mansour, 2015, Al‐Bahar and Crandall, 1990, Kangari, 1995, Wang *et al.*, 2004, Zayed *et al.*, 2008, Andi, 2006, Ibrahim, 2011, Kartam and Kartam, 2001, Wang and Chou, 2003, Choi and Mahadevan, 2008, Structural-Safety, 2015i, Structural-Safety, 1997h, Structural-Safety, 1999d, Structural-Safety, 1999f, Atkins, 2006, Therrien, 2011, Transport for London, 2013, WSP, 2013, Thames-Water, 2013, Gardenbridgetrust, 2014) |
| Financial | Financial difficulty or failure | *1) Funding issues:* Variations by the client; Unavailability of sufficient cash flow; improper measurement and pricing of Bill of Quantities (BOQ); ill planned schedule and client's delay in payment; Financing difficulties because of tax or capital movement restrictions; Delay in payments; | 1) Escalation Clause 2) Price Contingency in the Bid 3) Project Financing by a Reputable Owner 4) Owner Purchase of Equipment & Material | (Casey, 1979, Mustafa and Al-Bahar, 1991, Al‐Bahar and Crandall, 1990, Kangari, 1995, Tang *et al.*, 2007, Tah *et al.*, 1993, Therrien, 2011, Bing *et al.*, 2005, Rezakhani, 2012, El-Sayegh and Mansour, 2015, Fang *et al.*, 2004, Andi, 2006, Ibrahim, 2011, Kartam |

| | | | | |
|---|---|---|---|---|
| | | <u>*2) Other issues:*</u><br>Financial default of subcontractor; Constructor’s difficulty in reimbursement; Incomplete or inaccurate cost estimate; Delay in payments | 5) Providing Performance Bond and Prequalification of Suppliers<br>6) Forward Contracts for Hedging Exchange Rate Changes | and Kartam, 2001, Zayed *et al.*, 2008, Wang *et al.*, 2004, Zou *et al.*, 2007, Castles and Parish, 2011) |
| | Insurance | Bad credit of the insurance companies or bank; Inadequacy insurance; Difficulty in relevant insurance compensation; Delay in payments | | (Bing *et al.*, 2005, El-Sayegh and Mansour, 2015, Kangari, 1995, Fang *et al.*, 2004, Andi, 2006, Ibrahim, 2011, Kartam and Kartam, 2001, Tang *et al.*, 2007, Li *et al.*, 2013) |
| | Cost increase | <u>*1) Because of planning or pre-construction:*</u><br>Quotation errors in tendering or construction time prediction errors made by contractors;<br><u>*2) Because of design:*</u><br>Design variation; Delay in documentation; Inadequate program scheduling; Delay in payments; Delay in documentation;<br><u>*3) Because of construction:*</u><br>Tight project schedule; Contractors’ poor management ability; Inadequate site information (soil test and survey report); Inadequate program scheduling; Delay in payments; Delay in documentation; Default by Sub-Contractors and Suppliers | | (Bing *et al.*, 2005, Rezakhani, 2012, El-Sayegh and Mansour, 2015, Kangari, 1995, Fang *et al.*, 2004, Andi, 2006, Ibrahim, 2011, Kartam and Kartam, 2001, Zou *et al.*, 2007, Al - Bahar and Crandall, 1990) |
| Time | Delays because of client | Project funding problems; Postponement of project; Variations by the client; Excessive procedures of government approvals | | (Ibrahim, 2011, Zou *et al.*, 2007, Therrien, 2011) |
| | Delays because of design | Design Variations; Tight project schedule; Excessive procedures of government approvals | | (Zou *et al.*, 2007, Therrien, 2011) |
| | Delays because of construction | Inadequate program scheduling; Contractor’s difficulty in reimbursement; Tight project schedule; Contractors’ poor site management; Excessive procedures of government approvals | | (Zou *et al.*, 2007, Therrien, 2011) |
| | Delays because of supplier | Suppliers’ incompetency to delivery materials on time | | (Zou *et al.*, 2007, Therrien, 2011) |

| | | | | |
|---|---|---|---|---|
| Quality | Design product quality | Project funding problems; Variations by the client; Tight project schedule; Design Variations | | (Zou *et al.*, 2007, Kangari, 1995, Kartam and Kartam, 2001) |
| | Construction product quality | Project funding problems; Tight project schedule; Contractors' poor management; Unavailability of sufficient amount of skilled labour; Poor competency of labour; Design Variations or changes; Low management competency of subcontractors; Inadequate site information; Bad quality of materials; Bad quality of workmanship; Improper quality control; Bad quality of materials (e.g. steel components) | | (Wang *et al.*, 2004, Hastak and Shaked, 2000, Kangari, 1995, Zou *et al.*, 2007, Fang *et al.*, 2004, Tang *et al.*, 2007, Structural-Safety, 2015o, Structural-Safety, 2015n) |
| Contractual and Legal | Delays | Delayed contractual dispute resolution; Delayed payment on contract and extras; Unfairness in tendering; Third-party poor liability; Conflict in contract documents and laws; Problems in dispute settlement due to country's laws; Enforceability of contracts; Intellectual property protection problems; Errors of omission of the bill of quantities; Errors of the unit price or total price of the project; Indeterminate or defective terms of the contract | | (Li *et al.*, 2013, Wang *et al.*, 2004, El-Sayegh and Mansour, 2015, Hastak and Shaked, 2000, Tang *et al.*, 2007, Fang *et al.*, 2004, Casey, 1979, Mustafa and Al-Bahar, 1991, Zayed *et al.*, 2008, Andi, 2006, Kartam and Kartam, 2001, Tah *et al.*, 1993) |
| | Changes or variation | Unfairness in tendering; Owners' breach of contracts and disputes with contractors; Change order negotiation; Poorly tailored contract forms; Third-party poor liability; Conflict in contract documents and laws; Problems in dispute settlement due to country's laws; Enforceability of contracts; Intellectual property protection problems; Errors of omission of the bill of quantities; Errors of the unit price or total price of the project; Indeterminate or defective terms of the contract | | (Rezakhani, 2012, Tah *et al.*, 1993, Li *et al.*, 2013, Wang *et al.*, 2004, El-Sayegh and Mansour, 2015, Hastak and Shaked, 2000, Tang *et al.*, 2007, Casey, 1979, Mustafa and Al-Bahar, 1991, Kangari, 1995, Fang *et al.*, 2004, Castles and Parish, 2011) |
| | Failure | Insolvency of contractor or owner; Problems in dispute settlement due to country's laws; | | (Tah *et al.*, 1993, Wang *et al.*, 2004, Hastak and Shaked, 2000, |

| | | | | |
|---|---|---|---|---|
| | | Enforceability of contracts; Intellectual property protection problems | | Casey, 1979, Mustafa and Al-Bahar, 1991) |

## A.3 Reference in Appendix A


Al-Bahar, J. and Crandall, K. (1990), "Systematic Risk Management Approach for Construction Projects", *Journal of Construction Engineering and Management,* Vol. 116 No. 3, pp. 533-46.

Andi (2006), "The importance and allocation of risks in Indonesian construction projects", *Construction Management and Economics,* Vol. 24 No. 1, pp. 69-80.

Atkins (2006), *Walton Bridge Flood Risk Assessment,* Atkins, Surrey, UK.

Bing, L., Akintoye, A., Edwards, P.J. and Hardcastle, C. (2005), "The allocation of risk in PPP/PFI construction projects in the UK", *International Journal of Project Management,* Vol. 23 No. 1, pp. 25-35.

Casey, J.J. (1979), "Identification and Nature of Risks in Construction Project: A Contractor's Perspective", in *Proceedings of ASCE Construction Risks and Liability Sharing Conference in Scottsdale, Arizona*, American Society of Civil Engineers, New York, NY, pp. 17-23.

Castles, C. and Parish, D. (2011), *The Risks and Financing of the Mersey Gateway Bridge,* North West Transport Roundtable and Chapaign for Better Transport, London, UK.

Choi, H.-H. and Mahadevan, S. (2008), "Construction project risk assessment using existing database and project-specific information", *Journal of Construction Engineering and Management,* Vol. 134 No. 11, pp. 894-903.

El-Sayegh, S. and Mansour, M. (2015), "Risk Assessment and Allocation in Highway Construction Projects in the UAE", *Journal of Management in Engineering,* Vol. 31 No. 6, pp. 04015004.

Fang, D.P., Li, M.G., Fong, P.S.W. and Shen, L.Y. (2004), "Risks in Chinese construction market - Contractors' perspective", *Journal of Construction Engineering and Management,* Vol. 130 No. 6, pp. 853-61.

Gardenbridgetrust (2014), *Garden Bridge Planning Application Environmental Statement Volume 5: Technical appendices,* London, UK.

Hastak, M. and Shaked, A. (2000), "ICRAM-1: Model for international construction risk assessment", *Journal of Management in Engineering,* Vol. 16 No. 1, pp. 59-69.

HSE (2015), *The Construction (Design and Management) Regulations,* Health and Safety Executive (HSE), London, UK.

Ibrahim, M. (2011), "Risk matrix for factors affecting time delay in road construction projects: owners' perspective", *Engineering, Construction and Architectural Management,* Vol. 18 No. 6, pp. 609-17.

Kangari, R. (1995), "Risk Management Perceptions and Trends of U.S. Construction", *Journal of Construction Engineering and Management*, Vol. 121 No. 4, pp. 422-29.

Kartam, N.A. and Kartam, S.A. (2001), "Risk and its management in the Kuwaiti construction industry: a contractors' perspective", *International Journal of Project Management*, Vol. 19 No. 6, pp. 325-35.

Li, Q.F., Zhang, P. and Fu, Y.C. (2013), "Risk identification for the construction phases of the large bridge based on WBS-RBS", *Research Journal of Applied Sciences, Engineering and Technology*, Vol. 6 No. 9, pp. 1523-30.

Mehdizadeh, R., Breysse, D., Taillandier, F. and Niandou, H. (2013), "Dynamic and multi perspective risk management in construction with a special view to temporary structures", *Civil Engineering and Environmental Systems*, Vol. 30 No. 2, pp. 115-29.

Mulholland, B. and Christian, J. (1999), "Risk Assessment in Construction Schedules", *Journal of Construction Engineering and Management*, Vol. 125 No. 1, pp. 8-15.

Mustafa, M.A. and Al-Bahar, J.F. (1991), "Project risk assessment using the analytic hierarchy process", *Engineering Management, IEEE Transactions on*, Vol. 38 No. 1, pp. 46-52.

Rezakhani, P. (2012), "Classifying Key Risk Factors In Construction Projects", *Buletinul Institutului Politehnic din lasi. Sectia Constructii, Arhitectura*, Vol. 2, pp. 27-38.

SEStran (2008), *Cramond Bridge Feasibility Study*, South East of Scotland Transport Partnership, SEStran, Edinburgh, UK.

Structural-Safety (1974), "Death from wall collapse", available at: http://www.structural-safety.org/publications/view-report/?report=3111 (accessed 25 December 2015).

Structural-Safety (1980), "Long term risks for ground anchors", available at: http://www.structural-safety.org/publications/view-report/?report=3083 (accessed 25 December 2015).

Structural-Safety (1997a), "Bridge assessment and strengthening", available at: http://www.structural-safety.org/publications/view-report/?report=2155 (accessed 23 December 2015).

Structural-Safety (1997b), "Bridge strikes", available at: http://www.structural-safety.org/publications/view-report/?report=2156 (accessed 23 December 2015).

Structural-Safety (1997c), "Cranes", available at: http://www.structural-safety.org/publications/view-report/?report=2168 (accessed 26 December 2015).

Structural-Safety (1997d), "Design against fatigue in steel structures", available at: http://www.structural-safety.org/publications/view-report/?report=2162 (accessed 26 December 2015).

Structural-Safety (1997e), "Hidden tension members", available at: http://www.structural-safety.org/publications/view-report/?report=2163 (accessed 26 December 2015).

Structural-Safety (1997f), "Pin connections in bridges and building structures", available at: http://www.structural-safety.org/publications/view-report/?report=2159 (accessed 26 December 2015).

Structural-Safety (1997g), "Post-tensioned concrete bridges", available at: http://www.structural-safety.org/publications/view-report/?report=2154 (accessed 23 December 2015).

Structural-Safety (1997h), "Scour and flood damage to bridges", available at: http://www.structural-safety.org/publications/view-report/?report=2157 (accessed 31 December 2015).

Structural-Safety (1999a), "Bridge assessment", available at: http://www.structural-safety.org/publications/view-report/?report=2173 (accessed 21 December 2015).

Structural-Safety (1999b), "Bridge strikes", available at: http://www.structural-safety.org/publications/view-report/?report=2177 (accessed 21 December 2015).

Structural-Safety (1999c), "Resistance to disproportionate collapse", available at: http://www.structural-safety.org/publications/view-report/?report=2171 (accessed 21 December 2015).

Structural-Safety (1999d), "Scour", available at: http://www.structural-safety.org/publications/view-report/?report=2182 (accessed 21 December 2015).

Structural-Safety (1999e), "Steel connector failures and forged certificates", available at: http://www.structural-safety.org/publications/view-report/?report=3212 (accessed 26 December 2015).

Structural-Safety (1999f), "Thaumasite sulphate attack in structural concrete", available at: http://www.structural-safety.org/publications/view-report/?report=2185 (accessed 21 December 2015).

Structural-Safety (1999g), "Unsafe lamp posts", available at: http://www.structural-safety.org/publications/view-report/?report=3110 (accessed 19 December 2015).

Structural-Safety (1999h), "Unsafe street lighting columns", available at: http://www.structural-safety.org/publications/view-report/?report=3094 (accessed 19 December 2015).

Structural-Safety (2000), "Dynamic behaviour of bridges under pedestrian loading", available at: http://www.structural-safety.org/publications/view-report/?report=2387 (accessed 19 December 2015).

Structural-Safety (2004), "The collapse of the Nicoll Highway", available at: http://www.structural-safety.org/publications/view-report/?report=1481 (accessed 19 December 2015).

Structural-Safety (2006a), "Counterweight failure – Plank Lane Bridge", available at: http://www.structural-safety.org/publications/view-report/?report=3010 (accessed 19 December 2015).

Structural-Safety (2006b), "Issues for CDM co-ordinators", available at: http://www.structural-safety.org/publications/view-report/?report=3112 (accessed 19 December 2015).

Structural-Safety (2007), "Culvert roof slab failure", available at: http://www.structural-safety.org/publications/view-report/?report=3093 (accessed 28 December 2015).

Structural-Safety (2008a), "Bridge failure report and CDM", available at: http://www.structural-safety.org/publications/view-report/?report=3145 (accessed 28 December 2015).

Structural-Safety (2008b), "Partial collapse of the 'de la Concorde' overpass bridge: Laval Canada", available at: http://www.structural-safety.org/publications/view-report/?report=1883 (accessed 28 December 2015).

Structural-Safety (2008c), "Temporary bridge jacking", available at: http://www.structural-safety.org/publications/view-report/?report=3135 (accessed 28 December 2015).

Structural-Safety (2009a), "Network Rail bridge collapse", available at: http://www.structural-safety.org/publications/view-report/?report=1462 (accessed 28 December 2015).

Structural-Safety (2009b), "River Crane bridge failure - news", available at: http://www.structural-safety.org/publications/view-report/?report=1464 (accessed 28 December 2015).

Structural-Safety (2013), "Adjacent excavations in Australia", available at: http://www.structural-safety.org/publications/view-report/?report=3601 (accessed 28 December 2015).

Structural-Safety (2015a), "Adequacy of termination connectors for tensile bars", available at: http://www.structural-safety.org/publications/view-report/?report=3214 (accessed 17 December 2015).

Structural-Safety (2015b), "Asymmetric bridge design and construction – a near miss", available at: http://www.structural-safety.org/publications/view-report/?report=3218 (accessed 17 December 2015).

Structural-Safety (2015c), "Brick bridge collapse", available at: http://www.structural-safety.org/publications/view-report/?report=3132 (accessed 20 December 2015).

Structural-Safety (2015d), "Bridge component failures", available at: http://www.structural-safety.org/publications/view-report/?report=3134 (accessed 20 December 2015).

Structural-Safety (2015e), "Concern about bridge hangers", available at: http://www.structural-safety.org/publications/view-report/?report=3193 (accessed 17 December 2015).

Structural-Safety (2015f), "Concrete half joints", available at: http://www.structural-safety.org/publications/view-report/?report=3084 (accessed 20 December 2015).

Structural-Safety (2015g), "Dangerous arches", available at: http://www.structural-safety.org/publications/view-report/?report=3143 (accessed 23 December 2015).

Structural-Safety (2015h), "Documentation for imported large diameter steel pins", available at: http://www.structural-safety.org/publications/view-report/?report=1456 (accessed 16 December 2015).

Structural-Safety (2015i), "Engineering disaster mitigation", available at: http://www.structural-safety.org/publications/view-report/?report=3100 (accessed 20 December 2015).

Structural-Safety (2015j), "Fall of bridge deck support due to bolt over-tightening", available at: http://www.structural-safety.org/publications/view-report/?report=3187 (accessed 17 December 2015).

Structural-Safety (2015k), "Falsework support to a bridge – a near miss", available at: http://www.structural-safety.org/publications/view-report/?report=1444 (accessed 16 December 2015).

Structural-Safety (2015l), "Foundation base plates for tower cranes", available at: http://www.structural-safety.org/publications/view-report/?report=3215 (accessed 16 December 2015).

Structural-Safety (2015m), "Girder Collapse", available at: http://www.structural-safety.org/publications/view-report/?report=2995 (accessed 25 December 2015).

Structural-Safety (2015n), "GRP panels as permanent shuttering", available at: http://www.structural-safety.org/publications/view-report/?report=1440 (accessed 16 December 2015).

Structural-Safety (2015o), "Quality of some imported steel components", available at: http://www.structural-safety.org/publications/view-report/?report=3217 (accessed 17 December 2015).

Structural-Safety (2015p), "Scaffold protection", available at: http://www.structural-safety.org/publications/view-report/?report=3071 (accessed 20 December 2015).

Structural-Safety (2015q), "Scaffolds and traffic protection", available at: http://www.structural-safety.org/publications/view-report/?report=3018 (accessed 23 December 2015).

Structural-Safety (2015r), "Stud failure", available at: http://www.structural-safety.org/publications/view-report/?report=1447 (accessed 16 December 2015).

Tah, J.H., Thorpe, A. and McCaffer, R. (1993), "Contractor project risks contingency allocation using linguistic approximation", *Computing systems in engineering*, Vol. 4 No. 2, pp. 281-93.

Tang, W., Qiang, M., Duffield, C., Young, D. and Lu, Y. (2007), "Risk Management in the Chinese Construction Industry", *Journal of Construction Engineering and Management*, Vol. 133 No. 12, pp. 944-56.

Thames-Water (2013), *Thames Tideway Tunnel Blackfriars Bridge Foreshore Navigational Issues and Preliminary Risk Assessment*, London, UK.

Therrien, J.-C. (2011), *Pre-feasibility Study Concerning the Replacement of the Existing Champlain Bridge*, Montreal, Canada.

Transport for London (2013), *Nine Elms - Pimlico bridge Feasibility study summary report*, Transport for London, London, UK.

TranSystems (2012), *Washington Street Pedestrian Bridge Feasibility Study*, TranSystems Corporation, Naperville, Illinois.

Well-Stam, D. van, Lindenaar, F. and Kinderen, S. (2004), *Project Risk Management: An essential tool for managing and controlling projects*, Kogan Page Publishers, London, UK.

Wang, M.T. and Chou, H.Y. (2003), "Risk allocation and risk handling of highway projects in Taiwan", *Journal of Management in Engineering*, Vol. 19 No. 2, pp. 60-68.

Wang, S.Q., Dulaimi, M.F. and Aguria, M.Y. (2004), "Risk management framework for construction projects in developing countries", *Construction Management and Economics*, Vol. 22 No. 3, pp. 237-52.

WSP (2013), *London Bridge Station Redevelopment Design for Future Climate Final Report*, WSP UK limited, London, UK.

Zayed, T., Amer, M. and Pan, J. (2008), "Assessing risk and uncertainty inherent in Chinese highway projects using AHP", *International Journal of Project Management*, Vol. 26 No. 4, pp. 408-19.

Zou, P.X.W., Zhang, G. and Wang, J. (2007), "Understanding the key risks in construction projects in China", *International Journal of Project Management*, Vol. 25 No. 6, pp. 601-14.

## Appendix B. Pre-defined Risk-related Lexicon

Table 12. Pre-defined risk-related lexicon

| No. | Keywords | Values |
|---|---|---|
| 1 | inflation | deflation, hyperinflation, price, inflationary, devaluation, recession, stagflation |
| 2 | currency | banknote, monetary |
| 3 | tax | taxation, surtax, taxpayer, tariff, taxable, revenue |
| 4 | restriction | limitation, requirement, limit, constraint, prohibition, regulation |
| 5 | demand | need, requirement, expectation |
| 6 | government | administration, governmental, regime, authority, legislation, judiciary, policy |
| 7 | power | strength, authority |
| 8 | stability | stabilization, robustness |
| 9 | nation | country, world |
| 10 | relation | relating, relative, connection |
| 11 | regulation | regulatory, regulating, regulated, legislation, regulator, guideline, provision, regulate, directive |
| 12 | variation | variability, variant, difference, divergence, alternation |
| 13 | bribe | kickback, blackmail, corruptly, defraud |
| 14 | corruption | cronyism, bribery, nepotism, fraud, malfeasance, abuse, corrupt, malpractice, mismanagement, lawlessness |
| 15 | pressure | stress, overpressure |
| 16 | foreign | overseas, foreigner, diplomatic, abroad |
| 17 | local | regional, locally, community, municipal, provincial |
| 18 | morality | moral, ethics, ethical, selfishness |
| 19 | criminal | crime, felon |
| 20 | cultural | culture, multicultural, religious, social, sociocultural, linguistic, socio, intercultural |
| 21 | fraud | fraudulent, bribery, malfeasance, embezzlement, corruption, extortion, malpractice, scam |
| 22 | competition | contest, challenge, competitor |
| 23 | temporary | temporarily, interim, makeshift |
| 24 | permanent | permanently, semipermanent |
| 25 | risk | likelihood, vulnerability, incidence |
| 26 | structure | substructure, structural, building, component, bridge, road, tunnel, drainage, framework |
| 27 | building | edifice, courthouse, construction, tower, structure |
| 28 | bridge | footbridge, drawbridge, culvert, roadbridge, causeway |
| 29 | collapse | disintegration, collapsing, collapsed, destruction, failure, damage |
| 30 | damage | damaging, damaged, devastation, disruption, harm |
| 31 | failure | failing, collapse |
| 32 | durability | reliability, toughness, robustness |
| 33 | fall | falling, drop |

| | | |
|---|---|---|
| 34 | electric | electrics, electrical, elec |
| 35 | fire | gunfire, afire, wildfire, ablaze, burning |
| 36 | disease | infection, poliomyelitis, lymphangitis, lytico, beriberi, malaria, pellagra, myocarditis |
| 37 | design | designing, designer, layout, architecture, redesign |
| 38 | deficiency | insufficiency |
| 39 | error | mistake, inconsistency, incorrect, erroneous, inaccuracy, flaw |
| 40 | mistake | error, flaw |
| 41 | change | alter, alteration |
| 42 | rework | reworking |
| 43 | delay | delayed, delaying, pause, postponement, shutdown, postpone, interruption |
| 44 | poor | inadequate, weak, substandard |
| 45 | increase | increased, increasing, boost, improve |
| 46 | cost | costing, expense, expenditure, price |
| 47 | river | valley, estuary, creek, lake, rivulet, gorge, headwater, riverbank |
| 48 | site | area |
| 49 | adverse | unfavourable, unfavourable, deleterious, harmful |
| 50 | weather | thunderstorm, thundery, cloudiness, showery, wind, fog, forecast |
| 51 | improper | inappropriate, unethical, unlawful, excessive, inadvertent, unprofessional, inadequate |
| 52 | construction | constructing, reconstruction, rebuilding, erection, building, refurbishment, redevelopment |
| 53 | deviation | variance, divergence, deviating |
| 54 | planning | plan, development, planner, planned, budgeting |
| 55 | method | technique, methodology, process, procedure, approach |
| 56 | scheme | plan |
| 57 | quality | durability |
| 58 | traffic | congestion, rail, pedestrians, train |
| 59 | defect | flaw, deficiency |
| 60 | environment | ecosystem, environmental, climate, ecological, atmosphere, sustainability |
| 61 | resource | environment, data, information |
| 62 | equipment | machinery, consumable, vehicle, device, facility, appliance, hardware |
| 63 | operation | deployment, operating, operational |
| 64 | breakdown | disintegration, deterioration, disruption, collapse, breakup |
| 65 | incorrect | erroneous, inaccurate, misleading, inconsistent, incorrectly, unreliable, imprecise |
| 66 | use | usage, utilize, using, employ, reuse, utilizing, utilise, used |
| 67 | ground | surface, hillslide, terrain, soil, aboveground |
| 68 | underground | subterranean, tunnel |
| 69 | personnel | staff, officer, technician |
| 70 | insurance | insurer, reinsurance, underwriter, policyholder, mortgage |

| | | |
|---|---|---|
| 71 | negative | negatively, unfavourable, bias, negativity, critical, biased |
| 72 | attitude | stance, mindset, memeanor, mentality, tendencies, behaviour, cynicism, ambivalence, indifference |
| 73 | wage | salary, income, pension, tariff, compensation |
| 74 | modification | alteration, modifying, revision, configuration, change, redesign, modified |
| 75 | requirement | provision, required, criteria, restriction, limitation, mandatory, regulation, obligation |
| 76 | lack | lacking, paucity, dearth, inadequacy, unavailability, scarcity, lacked, inadequate, ineffectiveness, insufficient |
| 77 | agreement | treaty, pact, negotiation, memorandum, deal, contract, compromise |
| 78 | communication | interaction, communicating, communicative, connectivity, interworking, integration, communicational, coordination |
| 79 | limit | restriction, allowable, threshold, limitation |
| 80 | subproject | project |
| 81 | manpower | logistical, workforce, resource |
| 82 | inaccuracy | imprecision, inconsistency, unreliability, inaccurate |
| 83 | estimate | estimation, estimating, calculate, estimated, calculation, approximate, calculated, approximation |
| 84 | contractual | obligation, confidentiality, contract, enforceability, licensing, contractually, renegotiation |
| 85 | document | documentation, memo, memoranda, letter |
| 86 | flood | flooding, inundation, floodwater |
| 87 | flooding | flood, inundation, floodwater, erosion |
| 88 | earthquake | manshock, aftershock, mudslide, flood, quake, disaster |
| 89 | wind | etesian, crosswind, fohn, rain |
| 90 | lightning | thunder, thunderbolt, flame |
| 91 | landslide | mudslide, mudflow, earthflow, rockfall, landslip |
| 92 | snow | snowfall, snowdrift, sleet, ice, snowpack, thaw |
| 93 | temperature | humidity, precipitation, thermal, moisture, celsius |
| 94 | storm | hurricane, cyclone, typhoon, thunderstorm, windstorm, hailstorm, tornado |
| 95 | volcanic | volcano, lava, rhyolitic, ignimbrite, magmatic |
| 96 | cloud | stratocumulus, cumulonimbus, fractus, cumuliform, cirrostratus, cirrocumulus, mist, altocumulus |
| 97 | attack | assault, counterattack, raid, ambush, attacking, airstrike, invasion |
| 98 | funding | fund, financing, subsidy, grant, budget, funding |
| 99 | financial | economic, banking, investment, monetary, liquidity, finance, financing |
| 100 | difficulty | difficultly, problem, trouble, disadvantage |
| 101 | credit | mortgage, loan, saving, payment, billing, netbank |
| 102 | time | period, decade, year, timeframe, weekend |
| 103 | supplier | manufacturer, provider, distributor, importer, wholesaler, reseller, exporter |
| 104 | contract | deal, lease, loan, agreement, contractual, licence, warranty |

| | | |
|---|---|---|
| 105 | contractual | obligations, contract, licensing, contractually |
| 106 | law | jurisprudence, statute, legislation, legal, constitution, constitutional, judicial, policy |
| 107 | legal | judicial, constitutional, litigation, statutory, law, political, governmental, regulatory |

# Appendix C. INTERVIEWS WITH INDUSTRY EXPERTS

## C.1 Invitation letter

Dear [Name of participants],

Hope you are well and had a nice summer. My name is Yang Zou, a PhD student at University of Liverpool. I would like to invite you to participate in a research study. The format will be interview between you and me, which will last about 1 hour.

According to my research plan, so far I have completed most of my primary research, i.e. 1) I developed a risk case retrieval system by using Natural Language Processing and a prototype was developed with Python programming language (details could be found in the email below), 2) I developed a theory of establishing a collaborative environment for project risk management through establishing a linked relationship between risk information and Building information model, 3) to implement and validate our proposed theory, a tool prototype was developed based on a 4D BIM tool called Navisworks. I am now doing a small case study by using an existing 4D BIM model and our proposed tool to test and validate our theory.

The purpose of this interview is: 1) to discuss my PhD project and review the current process, and 2) to take an interview (between us) to validate our proposed theory and tool.

I have attached the participant information sheet and consent form. Before you decide whether to participate, it is important that you will take time to read the information in those documents carefully. We would like to stress that you do not have to accept this invitation and should only agree to take part if you want to.

If you would like to accept this invitation, could you please sign the participant information sheet and consent form, and back the signed documents to me by email before our interview. Meanwhile, may I suggest we find a date and time for this interview?

Thank you for reading this. Look forward to hearing from you soon.

Best regards,
Yang Zou

## C.2 Participant information sheet

1. **Title of Study**

   BIM and Knowledge Based Risk Management System

2. **Invitation**

   ***You are being invited to participate in a research study. Before you decide whether to participate, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and feel free to ask us if you would like more information or if there is anything that you do not understand. Please also feel free to discuss this with your friends or colleagues if you wish. We would like to stress that you do not have to accept this invitation and should only agree to take part if you want to.***

   ***Thank you for reading this.***

3. **What is the purpose of the study?**

   The purpose of the whole PhD research project is to develop a methodology to support collaborative project risk management using advanced information technology (e.g. BIM, database).

4. **Why have I been chosen to take part?**

   We would like to invite experts whose specialty are in BIM and/or project risk management to participate in the interviews to comments and discuss the PhD research.

5. **Do I have to take part?**

   The participation is totally voluntary and participants are free to withdraw at anytime without explanation and without incurring a disadvantage.

6. **What will happen if I take part?**

   Each interview will last about 1 hour and consist of two parts: 1) general discussion about the PhD research and the student's case study; 2) a semi-structured question guided interview between the student researcher and the participant.

7. **Expenses and / or payments**

   Potential expenses will not be covered by the researchers and there is no remuneration.

8. **Are there any risks in taking part?**

   None

9. **Are there any benefits in taking part?**

   The participants' comments and suggestions will be valuable for improving this PhD research as well as import guidelines for future research.

10. **What if I am unhappy or if there is a problem?**

    If you are unhappy, or if there is a problem, please feel free to let us know by contacting either Yang Zou (mob: 0742 130 3532) or Steve Jones (Tel: 0151 794 5228) and we will try to help. If you remain unhappy or have a complaint which you feel you cannot come to us with then you should contact the Research Governance Officer at ethics@liv.ac.uk. When contacting the Research Governance Officer, please provide details of the name or description of the study (so that it can be identified), the researcher(s) involved, and the details of the complaint you wish to make.

11. **Will my participation be kept confidential?**

    The student will use voice recorder to record the whole interview so that the student and supervisors can revisit and check the data. The data will be stored on the safe university server. All the data obtained will be finally approved by the experts and supervisors before use. Only the approved data and participants' names will be publicly available in the PI's final thesis. The interview recordings will be destroyed after the thesis have been approved.

12. **What will happen to the results of the study?**

    The participants' comments and suggestions for the PhD research will be included in the final PhD thesis, which means the data will be publicly available.

13. **What will happen if I want to stop taking part?**

    Participants can withdraw at anytime, without explanation. Results up to the period of withdrawal may be used, if you are happy for this to be done. Otherwise you may request that they are destroyed and no further use is made of them. If any of the results are anonymised the results may only be withdrawn prior to anonymisation. However, data cannot be withdrawn after the thesis is published.

14. **Who can I contact if I have further questions?**

    Please contact the PI if you have further questions:

    **Principal Investigator: Yang Zou**
    **Address: School of Engineering, University of Liverpool, Liverpool L69 3GH**
    **Email: yang.zou@liverpool.ac.uk**
    **Mob: 07421303532**

**C.3 Questionnaire**

Q1: What is your profession and how long have you worked in this area?

Q2: Do you think the observed knowledge gaps in this PhD study are correct? If yes, why do the observed gaps exist in the industry from your perspective? If no, could you please help point out what existing documents have covered them?

Q3: Do you think that the proposed theory and tool prototype have the potential to address the observed problem? And why?

Q4: Do you have any comments and suggestions about the proposed theory and tool, and for future research in the area of BIM-based risk management?

### C.4 Expert evaluations

*Alastair Soane* *17 November 2016*

- **Answer to Q1**: I am a civil and structural engineer and my career spans 50 years. For the last ten I have been Director of Structural-Safety which is an organisation dedicated to learning from failures and dissemination the information to others.

- **Answer to Q2**: You have correctly identified gaps in our knowledge about how safety related matters are identified and how engineers can learn them. Firstly, failures have to be identified and information about them published but this does not always take place because of a reluctance by individuals and organisations to admit to events which might reflect badly on them. Secondly the information must be retrieved and lessons that can be learned should be used to improve the design and construction of new projects.

- **Answer to Q3**: The technique that you have developed for data-mining existing reports to extract information is unique in the field of structural engineering and will prove to be of real value to practicing engineers. The application of this to BIM is a starting point for highlighting safety issues to designers. As discussed if BIM models can be interrogated for potential safety issues such as a lack of stability, particularly when changes are made, this will benefit designers and builders. I have concerns that some computer modelling is so complex that users do not recognise potential problems and your system has the potential to address this issue.

- **Answer to Q4**: You demonstrated the tool for a single model and I suggest that if the data base were shared between different models then the learning process would be quicker and better. Safety related information gained from the first model would benefit the second model and so on. The growth of the data base and algorithms to aid design within a BIM environment could result in safer and quicker design processes. Ultimately this knowledge, gained from the experience of engineers in developed countries could be transferred to those in developing counties. If so it would be a major step because at present there are very many failures and a high death toll from structural collapses in some developing areas.

**Participant biography**

Alastair Soane

BSc PhD CEng FICE FIStructE

Director of Structural-Safety the combined group encompassing CROSS (Confidential reporting on structural safety) and SCOSS (Standing committee on structural safety) and sponsored by the Institution of Structural Engineers, the Institution of Civil Engineers and the Health and Safety Executive. Formerly CEO of Bingham Cotterell consulting engineers with extensive experience on UK and International projects. Former member of Building Regulations Advisory Committee for England and Chair of the Structures Working Party, member of the Advisory Group on Temporary Structures. Lewis Kent award winner from IStructE in 2012, author of numerous publications and visiting professor of Civil Engineering at the University of Liverpool. Past president of the Smeatonian Society of Civil Engineers.

*Gordon Crick* *23 November 2016*

- **Answer to Q1**: Health 4 Safety Inspector employed by Health and Safety Executive (HSE) since 1992; Chair of BIM 4 Health and Safety Working Group since Dec 2015.

- **Answer to Q2**: Yes I agree there is a knowledge gap: 1) current practice in the industry tries to start to carry risk information in BIM models, but it's very rudimentary and each business has their own way of doing it; 2) there is a wealth of experience among designers and architects, but there is no theory that applies this to data enriched models in BIM.

- **Answer to Q3**: I think the approach is powerful in showing how a Risk Breakdown Structure can be used to structure the product of risk information down to the tasks and activities. The linkage to BIM models using Navisworks is very creative, and the first time I have seen this done in its way. The search routines are very interesting and I see great potential in developing the risk library future.

- **Answer to Q4**: The link to the BIM model requires information about mitigation and the development of ideas what has mitigating can be recorded and enhanced using the BIM model is very interesting. The RBS presented and its analysis in worthy of future research, and in particular how the risk information is visualised and used in BIM. Can risk information be automatically presented to support expert judgements on risk required doing the design phase?

**Participant biography**

Gordon Crick

Gordon has been HM Inspector of Health and Safety with the HSE since 1992. Prior to this time he had his own construction company for 8 years and so has a vast amount of experience from both perspectives. Over the last 5 years Gordon has been specialising in leadership, and the operational aspects of CDM and the question of co-ordination and competence.

*Benedict Wallbank* *1 December 2016*

- **Answer to Q1**: I am now the BIM Strategy Manager at Viewpoint Construction Software as well as a Chartered Architect with over 35 years' industry experience.

- **Answer to Q2**: Yes, I agree the observed gaps are the real problems to be solved. In the industry some contractors are using the digital models for safety planning in weekly meetings. But I think the existing technologies can go a long way to overcome them.

- **Answer to Q3**: Yes. Firstly, the risk case retrieval system can help retrieve valuable information from previous problems and this makes sense. The approach of linking risks with BIM definitely works. Using open API and Navisworks to validate this research is fine. You could also use multiple APIs to develop a system to support different software.

- **Answer to Q4**: In the real world, a potential problem from my observation of using APIs is timing intensive because you have to re-do your APIs when any one of those tools goes to the next version. In the long term, I suggest that future research in this area should embrace Common Data Environment (CDE) and open BIM standards, e.g. IFC, and BIM Collaboration Format (BCF). Globally from an Information Management perspective, in the future you may consider extending your work to link site photos, task document, etc., within the CDE for regular risk checking.

**Participant biography**

Benedict Wallbank

Benedict is now the BIM Strategy Manager at Viewpoint Construction Software, a member of buildingSMART UK Technical Committee, and a Chartered Architect and BIM expert with over 35 years' industry experience.

*David Philp* *19 December 2016*

- **Answer to Q1**: With over 23 years' industry experience, I am now working as the Global BIM and Information Management (IM) Consultancy Director at AECOM. Chair of the Scottish BIM Delivery Group, a Professor at the Glasgow Caledonian University, and a Fellow of the CIOB, ICE and RICS.

- **Answer to Q2**: Yes, there is a theoretical gap. Many years ago I saw several US guys tried to do Monte Carlo simulations to the BIM models. However, in fact the application of risk management in the BIM environment is very limited at this moment. Most of the current applications include, for example, using clash detection or rule checkers for detecting geometrical risks. In terms of a wider risk management, there is still a need of human intervention, and the integration of technology (e.g. 4D simulation) and human knowledge and experience. Risk information is generated during the dynamic process and should be well recorded and managed throughout the project lifecycle.

- **Answer to Q3**: Yes, I agree that the proposed approach is helpful to the research questions. The linkage between BIM and risks does it work for the new projects and future research also needs to think about how to improve the risk management for existing buildings and bridges. The risk retrieval system you developed is very key in my opinion. For example, each contractor has their own risk databases and by using such a semantic search engine you can quickly find out a similar case and compare the two "products".

- **Answer to Q4**: 1) Managing risks within the CDE environment. The CDE is a key piece because CDE contains so many things, e.g. geometrical objects, geo-spatial, and non-graphical data, and in my view is beyond the IFC. 2) Strengthening risk assessment for the project operation stages, particularly for those existing buildings or bridges that have no BIMs. An existing method is to take the advantage of the 3D Laser Scanning and obtain the Point Cloud data to establish the BIMs for these structures. 3)

Improving the risk knowledge transferring and management from one project to the others. 4) Use of other information technologies, e.g. standard rule based checking, to improve the automatic detection of risks.

**Participant biography**

David Philp

MSc, BSc, FICE, FRICS, FCIOB, FCInstES, FGBC

With over 23 years' industry experience, David is now working as the Global BIM and Information Management (IM) Consultancy Director at AECOM. He is also the Chair of the Scottish BIM Delivery Group, a Professor at the Glasgow Caledonian University, and a Fellow of the CIOB, ICE and RIC.

*Martin Simpson* *27 February 2017*

- **Answer to Q1**: I am a Chartered Civil and Structural Engineer with 20 years' industry experience on Structural Design and Construction. I specialise in BIM and stadium design.

- **Answer to Q2**: Yes the identified knowledge gaps from an academic perspective are correct. BIM is currently not well aligned with all risk management processes. It is possible to build some traditional tools (e.g. Risk Breakdown Matrix used in this study) into BIM. For example the Arup approach is based around the traditional Risk Management Spreadsheet and embeds risk warning triangles into REVIT which are linked back to the data in the spreadsheet. These "flags" sit alongside the components in 3D and 2D space and are used to communicate to users that there is a risk that needs to be managed. It is a communication and visualisation of risk that then needs to be managed in a traditional manner. It is also important to note that this tool is not in widespread use in the wider industry. What is currently missing is the ability to link more advanced tools (e.g. Root Cause Analysis) into BIM object structure. IFC does not support the CDM analysis and the risk sharing and communication is still not clear in the CDE environment. Therefore a more holistic and eventually automated way of avoiding risk is not possible.

- **Answer to Q3**: Yes this piece of research is a good thing because it articulates the problem space. Although it is not accurate to say this research by itself can fully solve the two knowledge gaps, the concepts, approaches, and tool prototypes are valuable to the observed problems. Without further work on embedding risk into the object structure of the IFC, linkage is the most efficient way to communicate risk information management. From a long-term perspective, I think this research gives other researchers a holistic framework and ideas to do some further investigations. More work is required on the role that systems and system to system interaction, for example some systems interaction in a minor way during normal operation that could have critical interactions in emergency. For example emergency lighting cabling sits on a cable tray. In a seismic event the emergency lighting needs to be triggered, but if the cabling tray fails because it is not designed for seismic, then the emergency lighting will also fail. This critical scenario inter-dependency is not adequately catered for in current tools and approaches to BIM.

- **Answer to Q4**: 1) This research proposed the linkage relations between risks and Levels of Contents (LOC), Systems, and Objects of BIM. However, accidents in the real world often have consequential effects, e.g. failure of an important bridge may have impacts on the people and operation of some important social facilities, infrastructure system and a whole city. Further research is required in this scenario based system linkage, both within the project and it's linkage to wider city, environment etc. 2) Further work is required to expand the linkage and relations for IFC and the deployment of such thinking in the CDE environment. 3) Semantic searching engine developed by this research can work, however it is only as good of the body of work that the search engine can access. The reports from SCOSS are edited, abridged and verbose accounts to aid learning from past accidents. A significant practical challenge is how to manage risks through project to project learning so that previous knowledge can be transferred and learned effectively in new projects. It is recommended that both theoretical and practical investigations are needed in the future.

**Participant biography**

Martin Simpson

Martin is now a Reader in Digital Structural Design at the University of Liverpool and his speciality is in BIM and stadium design. Before joining Liverpool, he was the Associate Director of Arup and has been a Chartered Structural Engineer with more than 20 years' industry experience.