FEDERAL UNIVERSITY OF PERNAMBUCO
CENTER OF INFORMATICS
COMPUTER SCIENCE GRADUATE PROGRAM

JEFFERSON FERREIRA BARBOSA

A Model for Dealing with Epistemic Uncertainties in Agile Software Project Management

Recife

2023

JEFFERSON FERREIRA BARBOSA

A Model for Dealing with Epistemic Uncertainties in Agile Software Project Management

A Ph.D. presented to the Center of Informatics of Federal University of Pernambuco in partial fulfillment of the requirements for the degree of Philosophy Doctor in Computer Science.

Concentration Area: Software Engineering

Advisor: Hermano Perrelli de Moura

Co-Advisor: Marcelo Luiz Monteiro Marinho

Recife

2023

Catalogação na fonte
Bibliotecária Nataly Soares Leite Moro, CRB4-1722

B238m Barbosa, Jefferson Ferreira
A model for dealing with epistemic uncertainties in agile software project management / Jefferson Ferreira Barbosa – 2023.
174 f.: il., fig., tab.

Orientador: Hermano Perrelli de Moura.
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2023.
Inclui referências e apêndices.

1. Engenharia de software e Linguagens de programação. 2. Gerenciamento de projetos de software. 3. Incerteza epistêmica. 4. Quantificação da incerteza. 5. Gestão ágil. I. Moura, Hermano Perrelli de (orientador). II. Título

005.1 CDD (23. ed.) UFPE - CCEN 2023 – 121

Jefferson Ferreira Barbosa

"A Model for Dealing with Epistemic Uncertainties in Agile Software Project Management"

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação. Área de Concentração: Engenharia de Software e Linguagens de Programação.

Aprovado em: 03/02/2023

____

Orientador: Prof. Dr. Hermano Perrelli de Moura

BANCA EXAMINADORA

____

Profa. Dra. Simone Cristiane dos Santos Lima
Centro de Informática / UFPE

____

Prof. Dr. Adiel Teixeira de Almeida Filho
Centro de Informática / UFPE

____

Profa. Dra. Sheila dos Santos Reinehr
Programa de Pós-Graduação em Informática / PUC/PR

____

Prof. Dr. Ivaldir Honório de Farias Junior
Universidade de Pernambuco / Campus Garanhus

____

Prof. Dr. Marcos Felipe Falcão Sobral
Departamento de Administração / UFRPE

A list of acknowledgements always starts with the injustice of never representing all people who, during our journey, have been able to contribute directly or indirectly to our evolution.

I would like to thank the Espírito Santo SPS team for all the help and time dedicated to using and evaluating the Euler model, especially to professional colleagues Sandro Tonini da Silva and Tiago Zamperini.

I would also like to extend this thanks to the FABWORK startup team for their help and support with excellent comments and critical suggestions for improvements that made the Euler model make sense not only for public companies but also for private companies.

I would like to thank the day-to-day companions of SPS Paraíba: Marcio Augusto Araújo de Barros, Augusto César Benvenuto de Almeida, Noberto Ohara do Santos, Ewerton Pierre Miranda Lemos, José Júnior Andrade, Victor Augusto Rocco Ribeiro, José Wanderlúcio Lira, Vladimir Michel Bacurau Magalhães who took part in the Pandora project of the MPPB.

Some people appear in our lives in strange ways and shape our lives in ways that leave an eternal mark. I thank the prosecutor Dr. Octavio Celso Gondim Paulo Neto, for the courage and vision to believe in my dream and not only to believe in it but to live this dream.

I would also like to thank the professors Hermano Perrelli de Moura and Marcelo Luiz Monteiro Marinho for believing in me and sharing their knowledge and extremely helpful insights throughout these years.

And finally, I would like to thank my children Marianna Alencar da Costa Barbosa, Gabriel Alencar da Costa Barbosa and Manuela Alencar da Costa Barbosa, and my faithful wife Sammara de Fátima Alencar da Costa Barbosa for their sacrifices over the years. Without their support, none of this would have been possible.

ACKNOWLEDGEMENTS

A list of acknowledgements always starts with the injustice of never representing all people who, during our journey, have been able to contribute directly or indirectly to our evolution.

I would like to thank the Espírito Santo SPS team for all the help and time dedicated to using and evaluating the Euler model, especially to professional colleagues Sandro Tonini da Silva and Tiago Zamperini.

I would also like to extend this thanks to the FABWORK startup team for their help and support with excellent comments and critical suggestions for improvements that made the Euler model make sense not only for public companies but also for private companies.

I would like to thank the day-to-day companions of SPS Paraíba: Marcio Augusto Araújo de Barros, Augusto César Benvenuto de Almeida, Noberto Ohara do Santos, Ewerton Pierre Miranda Lemos, José Júnior Andrade, Victor Augusto Rocco Ribeiro, José Wanderlúcio Lira, Vladimir Michel Bacurau Magalhães who took part in the Pandora project of the MPPB.

Some people appear in our lives in strange ways and shape our lives in ways that leave an eternal mark. I thank the prosecutor Dr. Octavio Celso Gondim Paulo Neto, for the courage and vision to believe in my dream and not only to believe in it but to live this dream.

I would also like to thank the professors Hermano Perrelli de Moura and Marcelo Luiz Monteiro Marinho for believing in me and sharing their knowledge and extremely helpful insights throughout these years.

And finally, I would like to thank my children Marianna Alencar da Costa Barbosa, Gabriel Alencar da Costa Barbosa and Manuela Alencar da Costa Barbosa, and my faithful wife Sammara de Fátima Alencar da Costa Barbosa for their sacrifices over the years. Without their support, none of this would have been possible.

ABSTRACT

Since agile methods began to be used for software development, project managers have been looking for ways to improve these projects. Agility coexists with uncertainty, as one of the agile project's principles is the possibility of rapid change. Uncertainty quantification allows comparative purposes and evaluating alternative approaches to real-world problems in managing uncertainty. Some recent studies show that current approaches to managing uncertainty organize known project information but give little or no indication of the unknown information or uncertainties associated with the project. These uncertainty management approaches do not consider the quantitative aspect of uncertainty management beyond the interrelationships between sources of uncertainty in software projects. This thesis aims to build a model to deal with epistemic uncertainties based on quantification approaches. In addition, it seeks to identify interdependent relationships between sources of uncertainty in the agile management of software projects. The method used in this work was action research conducted to investigate the quantification of epistemic uncertainty in the real context of software engineering design. The author also conducted a quasi-systematic literature review to support action research searching for approaches to quantify epistemic uncertainty in software projects. The application of the method illustrates the benefit of applying uncertainty quantification approaches to reduce and prioritize epistemic uncertainties in software projects. Applying the above research methods resulted in the construction of a model to quantify epistemic uncertainty in software projects. This thesis also discusses the impact of a proposed model for dealing with epistemic uncertainty management. These results share findings that can help agile software teams improve their uncertainty handling.

Keywords: software project management; epistemic uncertainty; uncertainty quantification; agile management.

RESUMO

Desde que os métodos ágeis começaram a ser utilizados para o desenvolvimento de software, os gerentes de projeto têm buscado maneiras de melhorar esses tipos de projetos. A agilidade coexiste com a incerteza, pois um dos princípios do projeto ágil é a possibilidade de mudanças rápidas. A quantificação da incerteza permite a comparação e avaliação de abordagens alternativas para problemas do mundo real no gerenciamento da incertezas. Alguns estudos recentes mostram que as abordagens atuais para gerenciar a incerteza organizam as informações conhecidas do projeto, mas dão pouca ou nenhuma atenção as informações desconhecidas ou incertezas associadas ao projeto. Essas abordagens de gerenciamento de incerteza não consideram o aspecto quantitativo do gerenciamento de incertezas além das inter-relações entre fontes de incerteza em projetos de software. Esta tese visa construir um modelo para lidar com incertezas epistêmicas com base em abordagens quantitativas. Além disso, busca identificar relacionamentos interdependentes entre fontes de incerteza na gestão ágil de projetos de software. O método utilizado neste trabalho foi a pesquisa-ação conduzida para investigar a quantificação da incerteza epistêmica no contexto real de projetos de engenharia de software. O autor também realizou uma revisão quase sistemática da literatura para apoiar a pesquisa-ação em busca de abordagens para quantificar a incerteza epistêmica em projetos de software. A aplicação do método ilustra o benefício de aplicar abordagens de quantificação de incertezas para reduzir e priorizar incertezas epistêmicas em projetos de software. A aplicação dos métodos de pesquisa acima resultaram na construção de um modelo para quantificar a incerteza epistêmica em projetos de software. Esta tese também discute o impacto do modelo proposto ao lidar com a gestão da incerteza epistêmica. Esses resultados compartilham descobertas que podem ajudar as equipes de software ágil a melhorar seu gerenciamento de incertezas.

Palavras-chave: gerenciamento de projetos de software; incerteza epistêmica; quantificação da incerteza; gestão ágil.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

This chapter places the topic of the thesis in context and presents the primary motivations for undertaking it. The following sections describe the objectives and procedure used to develop them and their contributions.

1.1 CONTEXTUALIZATION

The idea that defines the boundaries between modern and past times is the notion of risk control. The notion that the future is more than the sudden change of the will of the gods and the change in the perception that men and women are passive towards nature (BERNSTEIN; BERNSTEIN, 1998).

Seminal works such as Boehm (1991) have already highlighted that identifying and dealing with risks as quickly as possible during the project reduces the cost of risk prevention, helps avoid disasters in software development, and helps managers understand and manage uncertainties in software projects. Moreover, even today, with digital transformation, dealing with risks is still of the utmost importance for organizational performance (CHOUAIBI et al., 2022).

The interest in managing uncertainty arose from the evolution of thinking about project management (MOURA, 2011). Software projects can be characterized as projects that involve a high level of uncertainty, which is related to the level of innovation of these projects. Although risk and uncertainty management has attracted much attention in recent years in academia and by project management practitioners, there is still considerable development potential in this field (MARINHO; SAMPAIO; MOURA, 2014).

Recent trends in project management highlight the need to address the issue of project uncertainty (MARINHO; SAMPAIO; MOURA, 2017)(HE; HUSSAINI, 2023). In this context, uncertainty management becomes essential for risk management. It provides strategies for a manager more efficiently to transform the unknown into the known as a way to succeed in project management (RAMASESH; BROWNING, 2014; MARINHO; SAMPAIO; MOURA, 2017).

Software development teams cannot ignore the management of uncertainties because this can cause increased costs associated with software development, can affect the quality of the software product delivered, and can lead to delays and dissatisfaction with the software produced by the customer. Uncertainty affects software projects in many ways. A project manager with the ability to manage uncertainty can determine the success or failure of software projects in different types of companies. (DöNMEZ; GROTE, 2018).

Reinforcing the importance of managing uncertainties, Padalkar e Gopinath (2016), in their analysis of six decades of project management research, present three main research areas and practices that they call deterministic, exploratory, and non-deterministic. In the prevailing deterministic view, projects were measured by performance, focusing on the "iron triangle" of cost, time, and quality. Efficiency was achieved by constructing and implementing an optimized schedule of project activities, which were assumed to have fixed and deterministic attributes. In the exploratory view, the search was to explain the project phenomenon. Lastly, in the non-deterministic view, the emphasis was on the study of complexity and uncertainty in projects and it was argued there was a need for a mixture of empirical and conceptual approaches.

Padalkar e Gopinath (2016) also report dissatisfaction with the results of the first two areas, which led to the need to rethink project management, thus directing interest to areas considered non-deterministic, such as managing complexity and uncertainties in projects.

One of the significant challenges always faced by project managers is the need to make decisions about the future. However, these decisions are made in the present, thereby making this situation inherently uncertain. Applying uncertainty management can be a determining factor for success in software projects (MARINHO; SAMPAIO; MOURA, 2017).

For Ward e Chapman (2003), conventional project management techniques work best with well-defined artifacts and a relatively stable environment, but where the construction of these artifacts needs to be more fluid and uncertain, what is necessary is a broad perspective associated with managing uncertainty. Critics of the conventional form of project management argue that there has been an excessive focus on the execution and delivery of project artifacts such as the project management plan or the risk management plan.

1.2 MOTIVATION AND RESEARCH QUESTIONS

Ever since agile methods began to be used for software development, project managers have been looking for ways to improve how these projects are conducted. Agility coexists with uncertainty, as one of the principles of agile projects is that there is the possibility of making rapid changes. This is also identified in the interactive nature of agile projects, and allows for the possibility of changing course when necessary (DINGSØYR et al., 2012).

The current trend of using agility in software development indicates the need to manage uncertainty by taking advantage of the inspection cycles of software development and adapting to changes, in addition to the perception that uncertainties can generate opportunities through the mitigation of threats and the rise of opportunities (FONTANA et al., 2015)(DöNMEZ; GROTE, 2018).

Chapman e Ward (2011) show that tools such as the Risk Matrix present very simplistic definitions of risk, which are limited to possible adverse events, and are measured by the product of the well-known equation of probability versus impact, resulting in revealing the degree of risk exposure. Thus, the assessment of risk is limited to identifying some sources of risk, thereby ignoring other uncertainties identified by researchers, such as ambiguity, variability, and systemic uncertainty. These tools focus on specific sources of a low level of uncertainty, whereas applying more complex models can produce better results.

According to Khodakarami, Fenton e Neil (2007), not measuring the impacts of the interdependence between the existing relationships of the different sources of uncertainty can lead to many sources of uncertainty not being identified, and possibly have adverse impacts on the project.

In order to measure and analyze uncertainty properly, practitioners must consider that projects are unique experiences. Thus, the type of uncertainty found in the endeavor is called epistemic uncertainty (i.e., it is related to a lack of knowledge about the project) rather than random uncertainty (i.e., it is related to randomness). To illustrate, lack of knowledge can be related to many factors, such as to an inadequate understanding of the process or phenomenon and to an inaccurate assessment of the particular characteristic of a particular project (BASU, 2016).

Chapman e Ward (2011) recognize that many sources of uncertainty are not independent of each other and that their interaction can lead to considerable impacts on projects. Therefore, a successful approach to managing uncertainty needs to recognize the various sources of uncertainty existing in projects and to model the interdependence relationships between the variables and their control/response relationships within the project (MARINHO; SAMPAIO; MOURA, 2013; MARINHO; SAMPAIO; MOURA, 2014).

In addition, authors like Chenarani e Druzhinin (2017) also highlight that due to the significant impact of uncertainty on project goals, approaches associated with quantification and monitoring can be very useful and informative for stakeholders and project managers. However, little attention has been given to the uncertainty management(TAIPALUS; SEPPäNEN;

PIRHONEN, 2020).

Epistemic uncertainty uses expert judgment, which can be deemed as the expertise provided by any group or person with specialized education, knowledge, skill, experience, or training, to assist in managing uncertainty. However, there is an inherent difficulty in analyzing expertise, which concerns the way of measuring the accuracy of the information of an expert's perception. The credibility of the information can vary greatly depending on the expert's familiarity with the theme of a project and thus one of the biggest challenges is to add some measure or weight to the information from different experts in accordance with the expert's level of expertise (CRUZ; TROFFAES; SAHLIN, 2022).

The above text reinforces the need to search both for alternative approaches to managing uncertainties in software projects and for quantitative forms of measuring uncertainty that can consider the impact of existing interdependence relationships between sources of uncertainties in the project. Examples of such approaches found in the literature are those based on probabilistic methods such as Bayesian Networks, and Evidence Theory (DANTAS et al., 2021).

Hence, the specific problem for this research area is the need for quantitative approaches to managing epistemic uncertainty rather than relying on qualitative approaches. Also, we need more approaches that consider the interdependent relationships between the various sources of uncertainties in software projects. So, we understand that developing a model focusing on techniques that quantify uncertainty that consider the relationships of interdependence between the sources of uncertainties in software projects can be of great value for the community of project managers.

Therefore, based on the context presented above, the main research question investigated by this study is *RQ - How do agile teams quantify epistemic uncertainties in software projects?*. To help answer the fundamental research question, we present a further four related questions:

- RQ1.1 What are the approaches to quantifying epistemic uncertainty in software projects that are already to be found in the literature?
- RQ1.2 How do we measure the causal relationship between sources of epistemic uncertainty in software projects?
- RQ1.3 How do agile teams approach quantifying epistemic uncertainty in software projects?

- RQ1.4 What are the impacts of a model based on quantitative approaches that would help manage epistemic uncertainty in software projects?

The above questions summarize the concerns that motivate undertaking this research.

1.3 PRIMARY AND SPECIFIC OBJECTIVES

The primary objective of this research is to construct a model to deal with epistemic uncertainties based on quantification approaches. Besides this, it seeks to identify interdependence relationships among sources of uncertainty in the agile project management of software development.

In order to seek to achieve this main objective, the following specific objectives were outlined:

- To collect existing evidence in the literature regarding approaches to quantifying epistemic uncertainty in software projects;
- To investigate how agile teams deal with the quantification of epistemic uncertainty in software projects;
- To construct a model for quantifying epistemic uncertainty in software projects.
- To perform an assessment session to collect opinions from project management practitioners on the built model to identify points for improvement that will help answer the research questions more thoroughly.

1.4 SUMMARY OF THE THESIS

In this chapter, the general ideas have been presented and placed in context,and what has motivated the thesis has been set out as have its objectives. The overall structure of the thesis consists of this Introduction, called Chapter 1, and this is followed by:

- Chapter 2: discusses the state-of-the-art approach to uncertainty and its characteristics. This chapter also discusses the methods and techniques used to manage and quantify epistemic uncertainty in software project management;
- Chapter 3: presents the research method that guides this thesis;

- Chapter 4: presents a quasi-Systematic Literature Review with a view to finding approaches for quantifying epistemic uncertainty in software projects;
- Chapter 5: presents Action Research applied in the industry and the actions that led to developing the first version of the model proposed in this thesis
- Chapter 6: presents the process used to apply the model developed with risk and uncertainty management practitioners in real projects in the industry and the method used to evaluate the results of applying the first version of the model;
- Chapter 7: presents a model to deal with the management of epistemic uncertainty in software projects;
- Chapter 8: concludes this thesis, summarizes the initial findings, discusses and proposes next steps, possible future enhancements, current limitations, and future research areas;
- Appendix A: presents a practical guide for applying the Euler model;
- Appendix B: presents the semi-structured interview protocol used in Action Research;
- Appendix C: presents the protocol for the proof of concept research;
- Appendix D: presents the Python language scripts used to assist applying the model.

2 THEORETICAL BACKGROUND

This chapter aims to present the basic concepts of this research. It will present concepts related to risk, uncertainty, epistemic uncertainty, uncertainty management, Bayesian Networks, and Dempster-Shafer Evidence Theory.

2.1 RISKS

The word "risk" derives from Old Italian *risicare*, which means "to dare". Risk refers to future events and changes in thinking, changes in opinions, actions or places (BERNSTEIN; BERNSTEIN, 1998).

A risk definition focused on project management comes from the Project Management Institute (PMI) in its standard, The Standard for Risk Management in Portfolios, Programs and Projects (PMI, 2019a). The PMI states that risk is an uncertain event that, if it does occur, has a positive or negative effect on the results of a project. This definition includes two critical dimensions of risk: uncertainty and effect. Other widely used terms for these dimensions are probability and impact, respectively, for uncertainty and effect.

Boehm (1991) introduces two fundamental concepts of risk: risk exposure which is a measure of the potential impact of risk on project objectives at any point in the life of a project, program or portfolio and risk factor (PMI, 2019a). Boehm also relates these two concepts through the formula in (2.1):

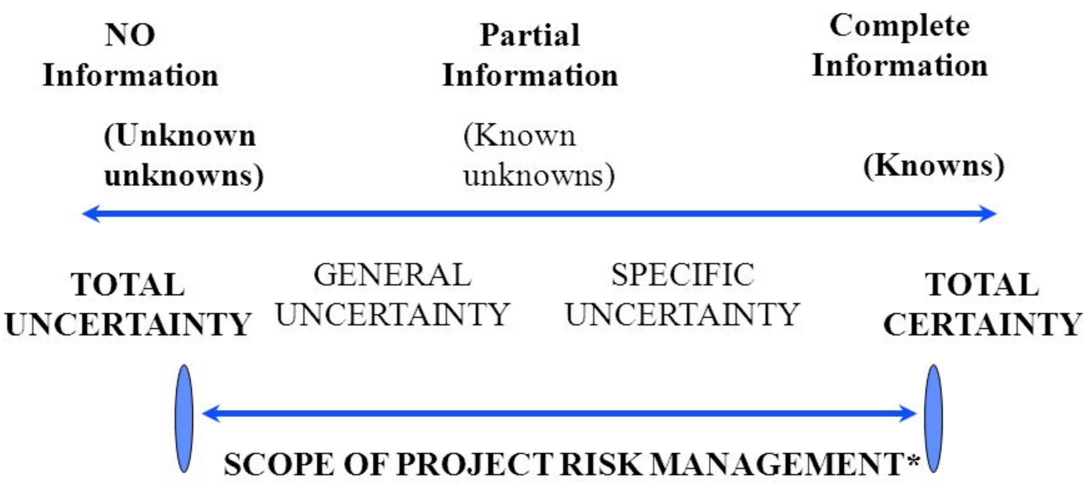
$$RE = P(UO) * L(UO), \tag{2.1}$$

where RE - Risk Exposure, P(UO) - the probability of an unsatisfactory outcome means the probability of an unexpected outcome and L(UO) - loss to the parties affected means the loss suffered by stakeholders.

Lister (1997) tells us that risk is any variable in the project that can or cannot be controlled and the value of which can contribute to the failure or success of the project. The author also mentions that a project may be threatened by political risks, communication risks, timing, legislation and technical risks.

The concept of uncertainties in software project management is commonly associated with the concept of risk. It can be seen in Figure 1 that there is a region associated with a lack of information (no information) before risk management (partial information) that is called uncertainty management. Thus, there is space for traditional project management tools and risk management techniques when there is complete information. The challenge lies in the spectrum where there is a lack of knowledge and traditional risk management tools are ineffective.

Figure 1 – Spectrum of Uncertainty and Risk Management.



Source: Wideman (1992), Loch, DeMeyer e Pich (2006)

It is essential to highlight that this thesis is in line with Marinho (2015) since it understands that uncertainty is something greater than risk and that risk is part of a process for managing uncertainty.

2.2 RISK AND UNCERTAINTY

A natural confusion exists between the definitions of risk and uncertainty. Understanding the concepts of risk and uncertainty is essential for understanding the approaches to managing uncertainty found in the literature.

The seminal work of (Knight Frank, 1921) asserts that risk refers to events that are statistically capable of being modeled, with known probability distributions and based on past events and reliable data. It is possible or feasible to have probabilities associated with the outcomes of such an uncertain events.

These types of uncertain events are characteristic of projects with a high degree of innovation. Innovation projects impose a high degree of uncertainty, considering that in addition to seeking to identify probabilities associated with the event, there is the fact that it is impossible to recognize which variables can influence project management.

2.3 UNCERTAINTY

The term uncertainty can be widely applied as the "lack of certainty", which means there is an absence of information. Therefore uncertainty covers not only probabilistic factors or indefinite results but also ambiguous factors or lack of clarity about a given phenomenon (HOWELL; WINDAHL; SEIDEL, 2010; MARINHO; SAMPAIO; MOURA, 2017).

To Perminova, Gustafsson e Wikström (2008), "Uncertainty is an event or a situation which was not expected to happen, regardless of whether it could have been possible to consider it in advance". In the context of projects, according to Marinho (2015) "Uncertainty in projects is the resulting phenomenon of limitations in seeing signs that may affect project success. Thus, it is something that cannot obtain an occurrence probability, even if subjective. This difficulty may be generated by a lack of experience, sufficient information, perceptive ability, or even the member's mindset in the project. At this point, the organizational culture can have a strong influence".

In contrast, Thomé et al. (2016) consider uncertainty as something that cannot be perceived, something that cannot be observed (i.e., unknown-unknown). Therefore something for which it cannot be prepared, much less estimated or quantified.

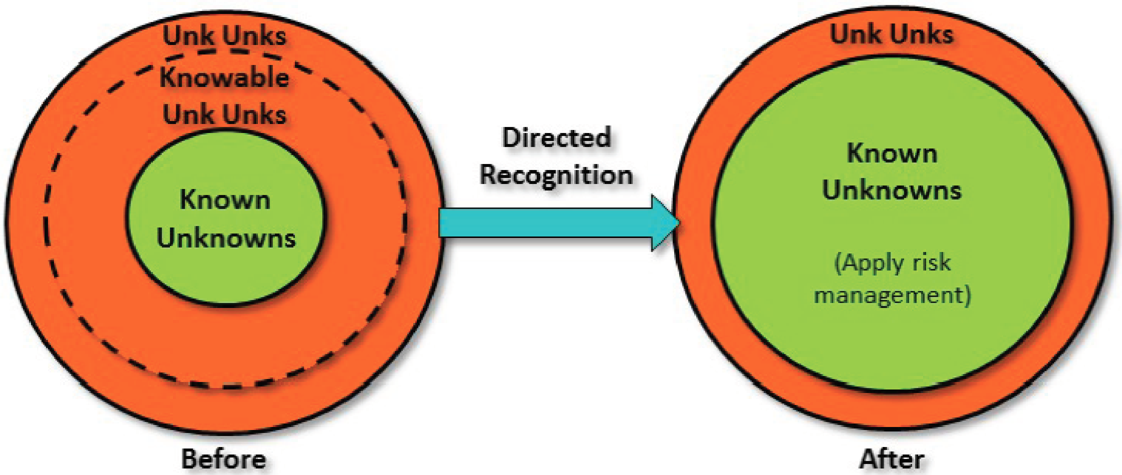
According to JCGM et al. (2008) uncertainty is "the quality or state that involves imperfect and/or unknown information. It applies to predictions of future events, estimates, physical measurements or unknown properties of a system".

Browning e Ramasesh (2015) based on Wideman (1992) also agree that the level of uncertainty is related to the amount of information that the project managers know and that can be classified into:

- knowns: when there is sufficient information for uncertainty to become a risk and we can apply the known techniques for risk management in projects;
- known unknowns: when the information is partially known, and there is a probability that the uncertain event occurs, and the likely impact on the objectives of the project can be assessed;

- unknown unknowns (unk-unks): when there is complete ignorance of what will happen and thus it is impossible for any type of verification to be made on the impacts of the occurrence of uncertainty. This last level can also be called unpredictable and highly complex.

Figure 2 – Converting knowable unk-unks into known unknowns.



Source: Browning e Ramasesh (2015)

Figure 2 illustrates the need to transition between a level of uncertainty (lack of information) and the level of knowledge in which the known techniques for risk management can be applied. As a result, for Marinho et al. (2015b), project managers should actively seek new information and make adjustments to project activities as new data emerge, applying new solutions to problems as they arise as a way of dealing with uncertainty.

Although the purpose is to reach the level of applying the known risk management techniques, a growing body of research suggests that dealing with uncertainty at the project level requires an approach that differs from the approach related to risk management that focuses on linear planning, forecasting, and control (MEYER; LOCH; PICH, 2002; LENFLE; LOCH, 2010; MARINHO; SAMPAIO; MOURA, 2017).

As stated above, we can conclude from the definitions about uncertainty that there is a common factor: the little information or lack of information associated with the definitions in the literature. Even the definition that deals with unforeseen events shows that it is not expected that when there is little or no information about the possibilities of something occurring, it will happen. It can be concluded that it is possible to deal with uncertainty by searching for information. Therefore,, searching for information is a possible approach to dealing with uncertainty.

Hence, this thesis takes the position that uncertainty can be anticipated and managed by using appropriate approaches. Uncertainty can be anticipated and quantified in order to bring a greater degree of precision to the decision-making process.

2.3.1 A Classification of Uncertainty

To understand the central theme of this thesis completely, it is interesting to identify and discuss the types of uncertainties present in the literature. The types of uncertainties found in the literature are random, agnostic, and epistemic.

2.3.1.1 Random Uncertainty

The word random derives from Latin *alea*[1], which means a game of chance with dice. Thus, it can be seen that uncertainty is at the heart of this type of phenomenon.

Random uncertainty is associated with the probabilistic variability or randomness of the elements associated with a phenomenon, such as the act of making a decision based on data entry. This type of uncertainty is considered irreducible since there will always be variability in the elements involved in the phenomenon.

Random uncertainty can be characterized by known probability distributions (KIUREGHIAN; DITLEVSEN, 2009) and categorized as aleatory if the modeler does not foresee the possibility of reducing them (JAKEMAN; ELDRED; XIU, 2010).

Random uncertainty is measured or characterized through relative frequencies: the number of times a participating event occurs from N repeated experiments. For example, with the dice roll, there is complete knowledge of the potential states that dice may be in (integer values 1-6), even so, there is a degree of variability in the possible occurrences of these values at each roll of the dice. This type of uncertainty is objective. It is a characteristic of the real world (DAMNJANOVIC; REINSCHMIDT, 2020).

[1] https://www.latinitium.com/latin-dictionaries?t=lsn1775

2.3.1.2 Agnostic Uncertainty

Agnostic uncertainty stems from the competitive nature of projects. As an example of this kind of uncertainty, we can mention that the profit of a project is at risk not only due to the uncertainty associated with quantities, prices, performance, but also due to the competitive nature of the project participants who are competing for customers, or the best members of the project team or the best suppliers. Agnostic uncertainty is about our degree of confidence, knowledge, and assumptions on which we build our judgments (KIUREGHIAN; DITLEVSEN, 2009).

2.3.1.3 Epistemic Uncertainty

The word epistemic derives from the Greek *epirsglg* (episteme), which means knowledge. Thus, epistemic uncertainty is caused by a lack of knowledge, information, or data (KIUREGHIAN; DITLEVSEN, 2009).

Epistemic uncertainty refers to the lack of information (lack of knowledge) about the modeled phenomenon or the lack of information about the elements that make up a phenomenon. For example, we can mention the degree of confidence (or lack of confidence) we have about the occurrence of a specific event. Therefore Jakeman, Eldred e Xiu (2010) conclude that this kind of uncertainty can be reducible by introducing additional information.

Again as mentioned at the beginning of this chapter, the focus of this work is on epistemic uncertainty for two reasons: because it is the type of uncertainty associated with the lack of information, among other types of uncertainties found in agile projects, and because it is the reducible type of uncertainty, precisely through the search for more information about the project.

2.3.2 Sources of Uncertainty

This subsection presents the studies on sources of uncertainty found in the literature.

One of the approaches to deal with uncertainty is to look for new sources of uncertainty (MARINHO; SAMPAIO; MOURA, 2013), especially when discussing epistemic uncertainty, the main characteristic of which is lack of knowledge. Therefore, knowledge about the sources of uncertainty is paramount for managing uncertainty.

To Macheridis e Nilsson (2004), there are two primary sources of uncertainty that can be quantified in the structure of a project: the number of activities and the total duration of the project.

McLain (2009) added a further three additional sources of uncertainty that can be quantified: (1) complex information dependencies among activities, (2) lack of knowledge about the duration of activities, and (3) unfamiliar project work.

With a more comprehensive view on sources of uncertainties, Marinho, Sampaio e Moura (2017), Marinho et al. (2015c), Marinho, Sampaio e Moura (2013) affirm that the sources of uncertainty are:

- *Technological uncertainty*: this dimension is associated with uncertainty about the knowledge and application of technology that the project makes use of;
- *Market uncertainty:* this dimension refers to the degree of novelty of the product, service, or outcome of the project;
- *Environmental uncertainty*: this dimension is directly determined by the complexity and dynamic environment in which organizations are inserted;
- *Socio-human uncertainty*: this dimension is determined by the relationships among the staff of an organization. These relationships have cognitive factors and are intrinsically related to how people perceive, learn, think, and assimilate the information within the organization.

The management of uncertainties in projects is not limited to the perception of threats, opportunities, and implications. Uncertainty management must identify the sources of uncertainties in projects and understand how to deal with them in the best possible way (MARINHO et al., 2015c). For this happen, it is necessary to model relationships between the sources of uncertainties and their impacts. Projects are unique experiences as defined by the Project Management Institute (PMI), and according to Ward e Chapman (2003), the uncertainty associated with projects is epistemic i.e., related to a lack of complete knowledge rather than random, or related to randomness.

2.4 AGILE SOFTWARE PROJECT MANAGEMENT

Agile management is a dynamic and iterative approach to software project management that emphasizes flexibility, collaboration, and continuous improvement. It emerged as a response to the traditional waterfall model, which often faced challenges adapting to changing requirements and delivering value to customers promptly. Agile methodologies provide a set of principles and practices that enable project teams to respond effectively to evolving customer needs and market conditions, resulting in higher customer satisfaction, better product quality, and increased project success rates (DYBÅ; DINGSØYR, 2008).

Agile management provides a dynamic and collaborative framework for software project management. It enables teams to adapt to changing requirements, fosters collaboration and empowerment, promotes continuous improvement and learning, and prioritizes customer satisfaction and value delivery. These qualities make agile methodologies an essential and valuable approach for managing software projects in today's fast-paced and evolving business landscape.

At the core of agile management are four key values outlined in the Agile Manifesto: individuals and interactions over processes and tools, working software over comprehensive documentation, customer collaboration over contract negotiation, and responding to change over following a plan Beck et al. (2001). These principles emphasize the importance of prioritizing people and their interactions, delivering a working product incrementally, involving customers throughout the development process, and embracing change as a natural and valuable aspect of software projects. Agile methodologies such as Scrum, Kanban, and Extreme Programming (XP) operationalize these principles through specific practices and frameworks.

One of the primary reasons agile management is important for project management is its inherent flexibility and adaptability. Traditional project management methodologies typically involve detailed upfront planning, which can lead to rigidity when faced with changing requirements or unexpected challenges. Agile embraces change and allows teams to continuously reprioritize and refine their work based on customer feedback and emerging insights. By embracing iterative development and frequent feedback loops, agile methodologies enable teams to adapt quickly and deliver value incrementally, mitigating risks associated with unforeseen changes and improving overall project success rates. However, because it is easy to adapt and quickly respond to changes, it is also prone to uncertainties arising precisely from the above characteristics that define agile management.

Agile management promotes a culture of continuous improvement and learning. Agile methodologies encourage teams to reflect on their work regularly, identify areas for improvement, and implement changes to enhance their processes, tools, and collaboration. This iterative approach allows teams to evolve and adapt, continuously refining their practices to optimize productivity, quality, and customer satisfaction. By embracing the "inspect and adapt mindset," agile teams can learn from their experiences and make data-driven decisions to drive better project outcomes (SERRADOR; PINTO, 2015).

2.5 APPROACHES TO QUANTIFYING UNCERTAINTY

The following subsections will present the ways of quantifying uncertainty that were found in the literature, and will focus on quantifying epistemic uncertainties, i.e., those associated with the lack of information.

Uncertainty Quantification (UQ) can be described as the quantitative characterization and propagation of uncertain inputs through a model until it produces a measure or response value in experimental and computational models (SHAH; HOSDER; WINTER, 2015).

Therefore, quantifying uncertainties can be of paramount importance for comparative purposes and thus, evaluating alternative solutions to real-world problems in project management. Measures or quantification of each source of uncertainty enable between-project comparisons of the resulting values. However, many sources of uncertainty are not quantified, e.g., sources from inside and outside the organization where the project is running, such as potential changes in the economy, project leadership, customer churn, and government regulations (MCLAIN, 2009).

According to Elmaghraby (2005), project management is a probe to uncertainty, whether to estimate resources, estimate deadlines, or costs, so the question is not to do with recognizing the existence of uncertainty but how to measure it.

According to Oberkampf et al. (1998), several methods have been used to estimate uncertainty, either by identifying possible sources of variability, uncertainty, and errors in computer simulations (random uncertainty) or through multiple sources of uncertainty coming from different experts' opinions (epistemic uncertainty). The application of these methods results in formulating mathematical structures that can adequately represent uncertainty (e.g., Dempster Shafer Theory (SHAFER, 1976)).

It is essential to highlight that quantification here refers to the numerical measurement of an uncertain factor to obtain greater precision in comparing actions to be taken or to foster more accurate decision-making to aid project management.

2.5.1 Dempster-Shafer Evidence Theory

The seminal work on the Dempster-Shafer Theory (DST) or Evidence Theory (ET) can be found in Shafer (1976), and Dempster (2008).

DST is a mathematical theory that allows the mathematical representation of uncertainty or lack of knowledge in complex domains. This theory is widely used when it is desired to assign uncertainty measures to a set of disjoint hypotheses (KOVACH, 2012).

In Evidence Theory, uncertainty is separated into belief and plausibility. While traditional Probability Theory uses only the probability of an event to analyze uncertainty, DST uses belief and plausibility to prevent bounds on probability. In Evidence Theory, uncertainty, belief, and plausibility represent a range of potential values for a given parameter without assumptions about the probability of the underlying data.

2.5.1.1 Belief and Plausibility Functions

A critical component of the DST is defining a set of mutually exclusive elements and hypotheses known as the frame of discernment θ. From the frame of discernment, another set is defined. This other set is called the Power Set, 2^{θ}. The last set represents all the possible combinations of the hypotheses from the frame of discernment.

In DST, the expert elicitation process specifies the attribution of a value in the interval [0, 1] by the expert, associating the expert's belief in the occurrence of a particular hypothesis. The relevance of the evidence for each of the elements (possible events that can occur) is represented by a function called the basic probability assignment (BPA) function or mass function or Body of Evidence (BOE) that can be seen in (2.2) (2.3).

$$m : 2^{\theta} \rightarrow [0, 1] \tag{2.2}$$

$$\sum_{A \subset \theta} m(A) = 1 \tag{2.3}$$

Belief Functions

Based on the evidence available to experts, the total belief in a given event A can be represented by the belief function. The total belief in A can be obtained from (6.2).

$$Bel(A) = \sum_{B \subset A} m(B) \tag{2.4}$$

Plausibility Function

Another critical concept to understand DST is the plausibility function. The plausibility function (Pl(A)), as can be seen in (6.1), can be interpreted as the risk of occurrence of event A. The function Pl(A) presents the degree to which A is plausible (admissible), given the evidence known by experts, i.e., it is the maximum belief attributed to an element or a set of elements based on existing evidence. Another way of describing plausibility is to realize that it indicates the degree to which we do not believe it ~A negation.

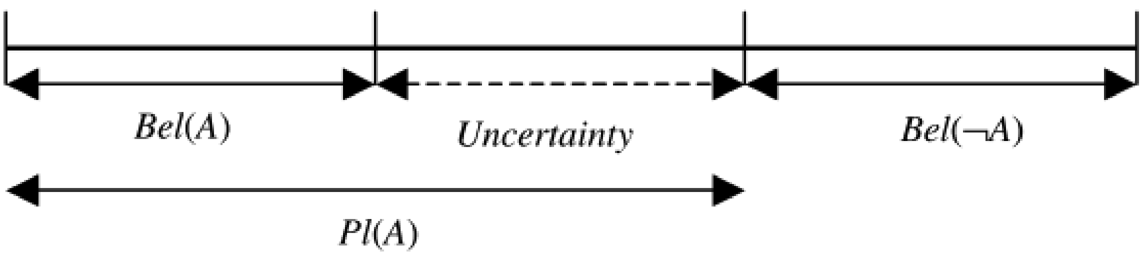
$$Pl(A) = 1 - Bel(\sim A) \;\; or \;\; Pl(A) = \sum_{A \cap B \neq \theta} m(B) \tag{2.5}$$

In addition to calculating the belief function and the plausibility function, another exciting feature of the DST is that it is possible to calculate the degree of ignorance (I). This calculation is made regarding the precise probability value due to the difference with belief plausibility, as seen in (2.6). With DST, the more the plausibility of a hypothesis increases, the lessbecomes the belief in the opposite hypothesis (FILHO; SOUZA; SCHILLING, 2007).

This statement can be viewed graphically in Figure 3.

$$I(A) = Pl(A) - Bel(A) \tag{2.6}$$

Figure 3 – Belief and Plausibility Domain



Source: Knight Frank (1921)

Figure 3 shows that Bel (~A) is the denial of the belief in the occurrence of A. Pl(A) reflects the maximum probability limit that can be distributed among the elements of A (upper limit function). It can also be seen graphically that the difference between Pl (A) and Bel (A) is defined as the expert's degree of ignorance in the occurrence of A (Uncertainty). Bel (A) and Pl (A) express the degree of uncertainty in the occurrence of A. When the ignorance to proposition A is decreased, the length of the interval is diminished.

2.5.1.2 Dempster's Rule of Combination

The main advantage of Dempster-Shafer Theory (DST) comes from Dempster's Rule of Combination (DRC), which allows two belief functions or independent bodies of evidence (BOE) to be merged (ZADEH, 1984). The use of this rule will allow the opinion of two or more experts to be combined, thus improving the outcome of the uncertainty assessment when using DST. Assuming there are two Bodies of Evidence m_1 and m_2 , the mathematical formulae for DRC can be seen in (2.7)(2.8)

$$m(A) = \sum_{B_1 \cap B_2 = A} m_1(B_1)m_2(B_2)/K \tag{2.7}$$

wherein

$$K = 1 - \sum_{B_1 \cap B_2 = \emptyset} m_1(B_1)m_2(B_2) \tag{2.8}$$

Despite the advantage associated with combining two BOEs, there is an issue associated with the DST when the BOE conflicts. Equation (2.8), represents the level of conflict between the two BOEs. Thus, k = 0 means a complete absence of conflict between the experts, whereas k = 1 means they contradict each other. DST has problems when BOE is in total conflict. The use of DRC in this type of situation can have counterintuitive results.

To solve this kind of problem, the paper by Silva e Almeida-Filho (2018) presents a solution based on a PROMETHEE-based approach applied within a framework for conflict analysis in Evidence Theory that integrates three conflict measures.

Another paper that presents a solution to the conflict indicated is Schubert (2008). It presents a solution for conflict management in Dempster-Shafer Theory using the degree of falsity.

2.5.2 Bayesian Networks

A Bayesian Network (BN) can be seen as a model that uses Graph Theory, Markov conditions, and probability distribution to model a phenomenon. If the modeling is done correctly, it is possible to make inferences about the phenomenon. However, this type of modeling has its limitations when there are many variables or characteristics to be modeled. Other names for BN are Bayesian Belief Networks (BBN), Probabilistic Networks (PN), or Causal Networks (CN) (NEAPOLITAN et al., 2004).

Bayesian Networks were introduced in 1986 by Pearl (1986) and are defined as Directed Acyclic Graphs (DAG) in which the nodes represent propositions or variables. The arcs indicate dependency relationships between propositions, and conditional probabilities quantify the strength of these dependencies.

For this research, we will always adopt the nomenclature of Bayesian Networks and its acronym BN.

An example of BN can be seen in Figure 4. This figure shows how the relationship between Age (A) and Sex (S) with Education (E) is modeled.

The representation of a Conditional Probability Table which numbers the occurrence probabilities of Age (A) within the possible domain (young, adult, old) and Sex (S) (M and F) is shown beneath the diagram.

The objective of this Bayesian Network is to verify what causal relationships exist between nodes and what are the probabilities of Travel (T) occurring from previous nodes, including Occupation (O) and Residence (R).

As indicated in Figure 4, it is common for some Bayesian Network models to recommend the use of a letter representing the node in question for network readability.

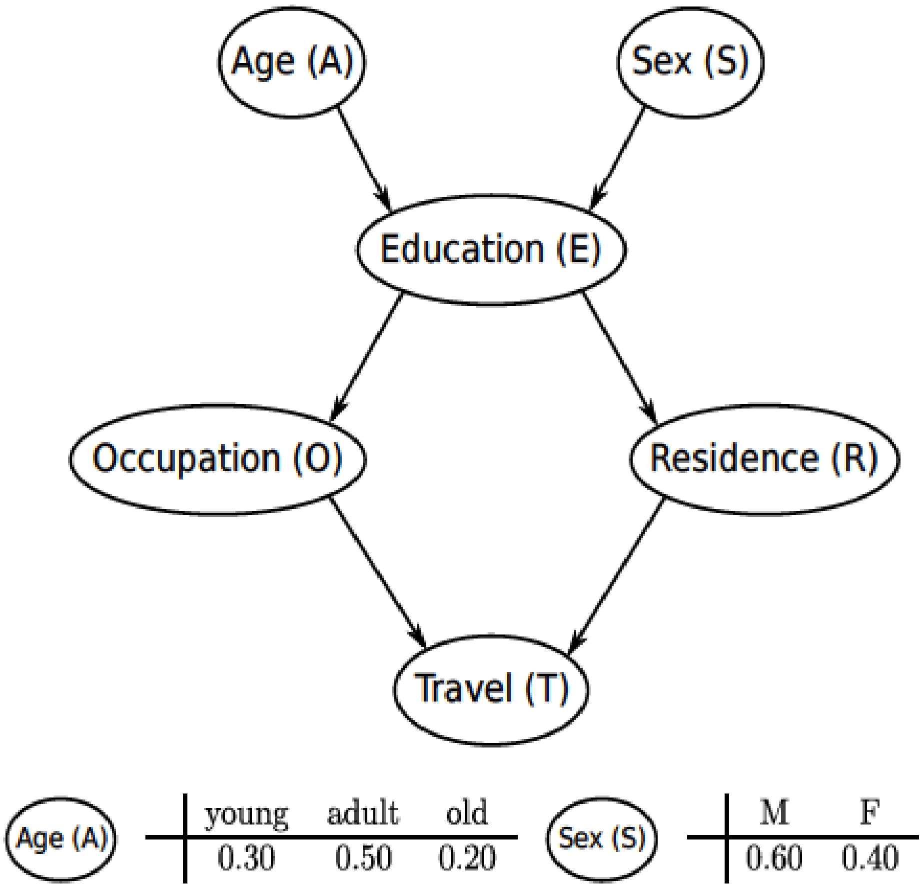
Expert knowledge plays a significant role in the construction of Bayesian Networks and the definition of relative probabilities. Thus, the attribution of probabilities using experts' knowledge is called expert elicitation (KNOCHENHAUER et al., 2013).

To understand the importance of Bayesian Networks, it is essential to know their origin, associated with Bayes' Theorem.

Bayes' Theorem can be expressed mathematically by the formula (2.9) below:

$$P(\mathbf{H}|\mathbf{E}) = P(\mathbf{H})\frac{P(\mathbf{E}|\mathbf{H})}{P(\mathbf{E})}, \tag{2.9}$$

Figure 4 – Discrete Bayesian Network



Source: Scutari e Denis (2014)

Where, P(**H**) is the probability of the hypothesis being true (before any evidence), P(**E**|**H**) is the probability of knowing the evidence if the hypothesis is true, P(**E**) is the probability of knowing the evidence and P(**H** | **E**) is the probability that the hypothesis is true given evidence of the event that occurred.

Evidence can be defined as the actual observation of some variable states or events. From these variables observed, it is possible to infer the state or occurrence of other variables. Examples of evidence may be information coming from databases, running a simulation, a project event, or an expert's opinion on a particular subject (MKRTCHYAN; PODOFILLINI; DANG, 2015).

Our interest in Bayes' Theorem comes from the fact that it is one of the mathematical ways of characterizing learning with experience, i.e., the modification of the initial attitude towards the evidence and hypotheses after having the information (piece of evidence) that a given hypothesis is materialized (PAULINO et al., 2018).

The creation in the mid-1980s of Bayesian Networks as a way of reasoning under uncertainty in order to represent probabilistic inference and knowledge was one of the essential factors for the emergence of expert systems (PEARL, 2014).

According to Scutari e Denis (2014), BN can be represented in three different types: a Multinominal Bayesian Network, a Gaussian Bayesian Network, and Hybrid Bayesian Networks.

A Multinominal Bayesian Network aims to model discrete data, such as modeling a Bayesian network that has 'sex' as one of its nodes or components (possible values male (M) and Female (F), education (possible values - high (h) or university (u), or a person's job occupation (possible values self-employed (self) or employee (emp).

As for Gaussian Bayesian Networks, their purpose is to model continuous data under the normal multivariate assumption. For example, we can think of a BN the purpose of which is to model the behavior of relationships between the production of vegetative mass and the mass of harvested grains, which is called a crop (for a crop analysis problem).

The last type, a Hybrid Bayesian Network, aims to model problems the variables of which belong to multinominal or normal multivariate families. For this type of network, we will use a mix of discrete and continuous variables.

The process of building a Bayesian Network goes through the following steps (RUSSELL; NORVIG, 2003):

1. Choose a set of relevant variables Xi that describes the domain;

2. Choose an order for the variables;

3. Select a variable Xi and add a node in the network;

4. Determine the parent nodes pa(Xi), among the nodes that are already in the network so that the Markov Condition is satisfied;

5. Determine the conditional probability table for Xi;

By definition, the Markov condition supposes the joint probability distribution of random variables in a set of nodes V in a DAG $G = (V, E)$. So we say that (G, P) satisfies the Markov condition if each variable $X \in V$, X is conditionally independent of the non-descendant nodes given its parents (NEAPOLITAN et al., 2004).

Bayesian network topology aims to create graphs that can represent the causal relationships between events. These graphs, called Directed Acyclic Graphs (DAGs), have main characters that are non-cyclic. Sometimes some confusion arises from trying to understand the difference between a DAG and Bayesian Networks. A DAG expresses only the conditional independence structure of a BN. A DAG must be understood as being the visual representation of the BN.

In order to make quantitative statements about the behavior of variables in a BN, we need to specify their joint probability distribution.

A Joint Distribution or Joint Probability Distribution is a set of random variables that provides the probability of combining the events on any of the variables. If the variables are independent, the joint distribution simplifies into the product of the marginal distributions (SCUTARI; DENIS, 2014).

Despite the advantages of Bayesian networks, there is a concern associated with the bias that experts can introduce. Whenever an expert is used as a source of information, their beliefs and experiences bias how they view the problem and what information they provide to help solve it. These biases take the form of either intentional or unintentional biases. Intentional biases result from the expert's willful decision to bias their assessment results.

2.6 RELATED STUDIES

This section presents the related studies published in the literature that technically is related to our thesis. This section engages on a brief discussion about papers that propose a different way to solve the same problem or the same way to solve different problems.

The work developed by Marinho, Sampaio e Moura (2017), resulting from applying several research methods such as Systematic Literature Review and Action Research, resulted in developing an Approach for Uncertainty Management in Software Projects.

Marinho's approach begins with a characterization of the type of project (traditional and agile), followed by the identification of sources of uncertainty and detection of early signs, passing through the application of Sensemaking techniques, ending with the integration with traditional risk management techniques and response plan.

Altogether, the approach contains five methods and 18 practices for reducing uncertainties. Marinho's Approach to Uncertainty Management mainly describes strategies that allow team members to formalize and manage uncertainty in software projects explicitly.

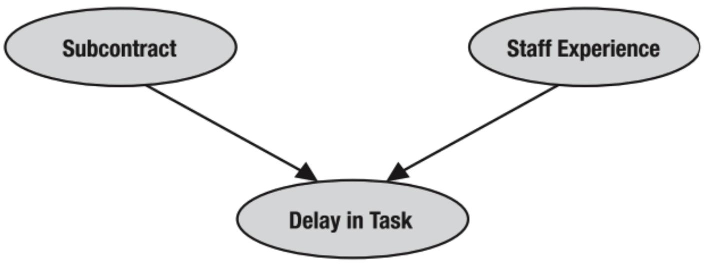
The focus of the study by Khodakarami, Fenton e Neil (2007) is on identifying causal relationships between the sources of uncertainty existing in projects, specifically in managing project schedules. For these authors, the basic inputs of schedule management such as time, cost, and resources of activities are non-deterministic and are affected by several sources of uncertainty. When this author refers to the term non-deterministic,this means that a procedure or a function will not always produce the same output. The approach presented by the author aims to model the relationships of causality and uncertainties by using Bayesian Networks. The strategy adopted by the author was to empower the known technique of the Critical Path Method and also to provide exploratory analysis to elicit, represent and manage different sources of uncertainty in project planning.

To support their choice of non-traditional approaches, the author highlights an existing weakness in the Monte Carlo Simulation (MCS) that does not deal with statistical dependence between the duration of activities in the project network. Furthermore, the MCS assumes project risks are independent events and does not identify the dependent relationship between sources of uncertainty.

These authors argues that an appropriate approach to managing uncertainty must provide means of identifying various sources of uncertainty, understanding their origins, and then managing them in ways that deal with the unwanted implications of uncertain occurrences.

Figure 5 shows the Bayesian Network with nodes, arcs and probability table. As we can see, this Bayesian Network comprises three nodes, the first two nodes, Subcontract and Staff Experience, related to Delay in Task. Next to each node, the Conditional Probability Table is shown.

Figure 5 – Bayesian Network with nodes, arcs and probability table



On Time	0.95
Late	0.05

High	0.7
Low	0.3

Subcontract		On Time		Late	
Staff Experience		High	Low	High	Low
Delay	Yes	0.95	0.7	0.7	0.01
	No	0.05	0.3	0.3	0.99

Source: Khodakarami, Fenton e Neil (2007)

For Bayesian Networks with discrete nodes (discrete as opposed to continuous, something that is separate, distinct, individual, or countable), the probabilistic dependence, i.e., the cause and effect relation, is often represented via a table called a Conditional Probability Table (CPT).

In the example, there are two possibilities for the Subcontract event On Time or Late. The project team estimated the values based on the probabilities of the occurrence of one or another event. Below the Delay in Task node, we can see a table of probabilities representing the junction of the probabilities of the nodes above Delay in Task. This table contains the probability of occurrence of Delay in Tasks taking into account the conditional probability of the occurrence of the above nodes.

For Khodakarami, Fenton e Neil (2007), the key benefits of Bayesian Networks are:

- They explicitly allow uncertainties to be quantified and modeling to be made of the existing interrelationships between the variables (nodes) that make up the Bayesian Network;
- They allow visualization of cause and effect between the nodes that make up the Bayesian Network;
- They make it possible to review past beliefs in light of new data;
- It is possible to build predictions from incomplete data (epistemic uncertainty) by combining subjective and objective data;
- They allow decisions to be made with the aid of visual and editable diagrams.

Finally, these authors conclude that a classic project management problem is how to deal with project uncertainty and that more advanced techniques are needed to capture different aspects of project uncertainty.

Another important study for the comprehension of our work is Santos e Cardoso-Junior (2018). This study deals with planning execution deadlines in R&D projects, given the high level of innovation inherent to projects of this category in Higher Education institutions. As a way to solve this problem, Dempster-Shafer's Theory of Evidence will be applied because of its ability to allow uncertainties based on the knowledge of experts to be dealt with. The study is applied to the knowledge of two experts, and then the data coming from the two experts is aggregated Thus, it is possible to estimate the most significant uncertainty interval and conclude which activity should be monitored or prioritized to mitigate its impact on the objectives of the project.

In DST, experts must assign mass to the element (probability of occurrence of a given event) in the interval [0, 1] based on their experiences and available evidence (pieces of

Figure 6 – Expert 1 - Basic Probability Assignment (BPA)

Expert 1	
Project R&D – Goals 1 and 2 (expected to be executed in 18 months)	
Goals (Evidence)	BPA(m1)
M1	0.5
M2	0.3
(M1, M2)	0.2

Source: Santos e Cardoso-Junior (2018)

information), and quantify their belief in a given hypothesis within the possibilities of events possible in the planning of project activities. In this way, using DST, it is possible to aggregate the knowledge of experts and calculate the relevance of the evidence for each activity within the project.

As we can see in Figure 6, using their knowledge and evidence of similar or previous activities, it was possible to carry out the elicitation of Expert 1. The expert assigned a mass of 0.5 to Target 1 (M1) and 0.3 to Target 2 (M2), with a mass of 0.2 for M1 and M2 (belief of occurrence of simultaneous events).

Using the mathematical formulae provided for the DST, it is possible to calculate the Belief (Bel) and Plausibility (Pl) of the occurrence of a given event.

Figure 7 – Result of expert 1 elicitation and calculation of belief and plausibility

Expert 1			
Project R&D - Goals 1 and 2 (expected to be executed in 18 months)			
Goals (Evidences)	BPA (m1)	Belief (Bel)	Plausibility (Pl)
M1	0,5	0,5	0,7
M2	0,3	0,3	0,5
(M1,M2)	0,2	1	1

Source: Santos e Cardoso-Junior (2018)

This process is repeated for expert 2, after which the experts' opinions can be aggregated, thereby giving more strength to the belief and plausibility of the occurrence of certain events (reduction of uncertainty).

From the results presented in Figure 7, it can be inferred, based on expert 1's elicitation, that target M1 is the one that needs more follow-up and monitoring. This result represents the minimum and maximum limit of the epistemic uncertainty involved and presents greater chances of impacting project delays.

2.7 CLOSING REMARKS

This chapter presented some concepts related to the theoretical basis that supports this thesis, including definitions of risk and uncertainty, epistemic uncertainty, and approaches for quantifying uncertainty. This chapter ends with a related work section where evidence found in the literature is discussed to support the construction of a new model for managing epistemic uncertainty.

3 RESEARCH METHOD

This chapter presents the research method for this thesis. The scientific method is necessary to make the research results more reliable and this includes assuring they can be reproduced independently by other researchers. Initially, we present a methodological framework used in this research, followed by a research roadmap for this thesis, and finally, we include a section about threats to the validity of a research study.

3.1 CONTEXT

There is no science without the use of scientific methods. Thus, the method is the set of systematic and rational activities that, with greater security and economy, is most likely to reach the expected objective. In other words, it will ensure that real and robust knowledge is obtained, an audit trail of the path followed is fully described, errors are detected and scientists are assisted in their decision-making (MARCONI; LAKATOS, 2017.).

One of the widely used approaches to seek to achieve and experience this goal is called qualitative research. Qualitative research is a nonlinear process. It is a set of interconnected methodological practices which is based mainly on seeking to interpret and understand how people see, approach, or experience the world around them and how they make sense out of that experience or of experimenting with specific phenomena (RAVITCH; CARL, 2015).

For Merriam e Tisdell (2015), qualitative research is an umbrella term covering a range of interpretative techniques that seek to describe, decode, translate, and create meaning for certain phenomena in the social world. They assert that qualitative researchers are interested in understanding the meaning constructed by people: how people make sense of their world and their world experiences.

The following section defines and classifies the qualitative research methodology used in this thesis. It sets out the definitions of Action Research, its process, and its phases for investigating how agile teams approach quantifying epistemic uncertainty in software development projects. We aim to answer the research questions presented in chap:introduction.

3.2 METHODOLOGICAL FRAMEWORK

The framework and fundamental structures of all research studies need to be well defined. Reviewing previous thinking and studies found in the literature can help illuminate our context and format the problem statement of the problem and the purpose of the investigation (MERRIAM; TISDELL, 2015). The framework for this research is set out in Table 1.

Table 1 – Methodological Framework

Philosophical Roots	Constructivist
Method of Approach	Inductive
Method of the Procedure	Exploratory Review of the Literature/ quasi-Systematic Literature Review / Action Research / Proof of Concept Research
Nature of the Variable	Qualitative / Quantitative

The philosophical roots are called Constructivist as it is assumed that reality is socially constructed. There is no single observable reality. This kind of research has multiple realities, or interpretations, of a single event. For this research, an adequate philosophical position is a Constructivist thought. This argues that knowledge must be based on the inductive as a primary mode of analysis and on comparative methods.

The stance of the inductive method strives to derive meaning from the data. The findings of the method for this approach are derived from the data in a qualitative study. They are shown in the form of themes, categories, typologies, concepts, and even theories about a particular aspect of practice (MERRIAM; TISDELL, 2015).

Data collection is a vital process for qualitative research. By means of this process, the researcher will access data of the phenomenon under study and start searching for the answer to the research question. The data collected can be numerical, and contain quantitative or "direct quotations from people about their experiences, opinions, feelings, and knowledge" that feature qualitative data or even of entire passages extracted from various types of documents (PAUL, 2015).

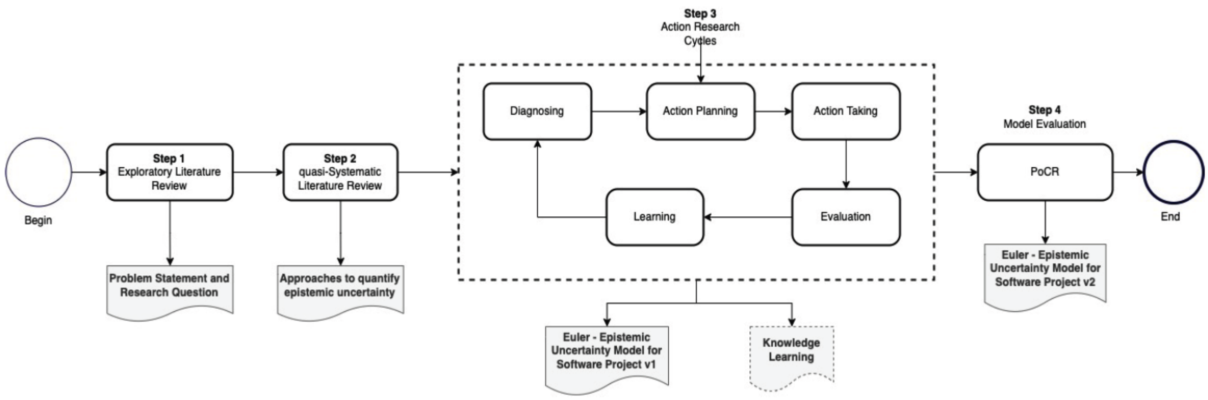
Each approach to collection should have different techniques for analyzing the data. These techniques are based on statistics in the first case or based on thematic analysis in the second case. Interviews, observations, and documents can be used to collect data. This thesis uses qualitative data, from interviews and observations and quantitative data from the project management systems used by the project on the focus of the research.

3.3 DESIGN OF RESEARCH

When starting a research effort, we need some idea of what we want to do and a plan that will assist us to reach our research goals. This plan or design of the research is the logical plan for getting from here to where the initial set of questions to be answered is, and there is some set of conclusions (answers) about these questions (YIN, 2013).

As we can see in Figure 8, this research consisted of four main steps. In the following sections, more details of these steps will be presented.

Figure 8 – Research Steps



Source: Author

3.3.1 Step 1 - Exploratory Review of the Literature

An exploratory review of the literature was carried out in the first step, which consisted of a broad review of concepts associated with uncertainty and epistemic uncertainty. The main output of this step was the definition of the problem statement, and of the research questions that would guide the proposed investigation.

Exploratory research is a scoping study guided by a requirement to identify the relevant literature associated with a given topic, regardless of a specific research method, so as to determine the value of undertaking a full systematic review, and to identify gaps in the existing literature that future studies could usefully fill (ARKSEY; O'MALLEY, 2005).

According to Arksey e O'Malley (2005), this kind of research can follow a framework comprised of the following stages for conducting a scoping study:

- Stage 1. Identifying the research question
- Stage 2. Identifying relevant studies
- Stage 3. Selecting studies
- Stage 4. Charting the data
- Stage 5. Collating, summarizing, and reporting the results

Following these stages, an exploratory review of the literature was conducted using search engines and databases such as Google, Elsevier ScienceDirect, Scopus, IEEEXplore, and ACM Digital Library to identify relevant studies. Two search blocks were performed as the accessible text search block conducted on the above engines and databases with the search criteria for peer-reviewed journals.

The first search block was represented by the search strings associated with "uncertainty", "software project management", "project management", "software project", and "quantitative techniques".

The final result of this step indicated there was a gap in the Software Engineering field, namely, the lack of approaches available to take into account the impact of existing dependency and interdependence relationships between the various sources of uncertainties in project management.

We were always concerned with ensuring that the process was documented in sufficient detail to enable the study to be replicated by other researchers.

3.3.2 Step 2 - Quasi-Systematic Literature Review

After completing the first step, the need was felt for a more extensive search in the literature for approaches to quantifying epistemic uncertainty in projects. For this task the meaning of the word approaches must be understood in a broader way (techniques, strategies, tools, theories, and a wide range of ways of dealing with quantifying uncertainty that can be found in the literature.). The main result of this step was a list of approaches to quantifying uncertainty taken from the current literature by using Evidence-Based Software Engineering processes.

According to Travassos et al. (2008), a quasi-Systematic Literature Review (qSLR) is also called a secondary study, the purpose of which is to collect evidence from the literature extensively and systemically to answer a research question. Compared to the traditional Systematic Literature Review (KITCHENHAM; CHARTERS, 2007), this method has the same formalism and rigor. However, it differs as no baseline criterion is available, which makes it hard to perform a comparative meta-analysis since both the SLR and qSLR focus on identifying and aggregating empirical evidence.

Despite the impossibility of performing a meta-analysis, Travassos et al. (2008) emphasize that qSLR is a valuable method for establishing an initial knowledge base for comparing primary studies and can produce valuable knowledge.

For Magdaleno, Werner e Araujo (2012), this type of secondary study can also be defined as an exploratory study designed to characterize the research area, with no baseline being established for comparing the results obtained.

The qSLR protocol and the results are presented respectively in APPENDIX C and Chapter 4 of this thesis.

3.3.3 Step 3 - Action Research

Step three was when Action Research was conducted to investigate how agile teams approach the quantification of epistemic uncertainty in software development projects in a real project scenario. This step resulted in constructing the first version of the model for this.

Action Research (AR) processes use the protocol described in Staron (2020) derived from his experience of conducting different Action Research studies in Software Engineering. In each cycle, we have the **Diagnosing** phase, which deals with exploring the research problem and certifies that it can be feasible. The **Action Planning** phase deals with the planning of the actions to be performed during the Action Research. In this phase, we will identify what will be done, when it will be done, how it will be done, and how the research actions will be carried out to solve the problem identified in the previous cycle. The next phase, **Action Taking**, deals with conducting Action Research. In this phase, there will be an object of study of Action Research to evaluate the consequences and effects of this intervention. The next phase deals with the **Evaluation** of impacts, reducing bias in the researcher's observations, and the effects of the intervention on the research object. Finally, we have the **Learning** phase which is considered one of the most important in the cycle, as it helps to raise the organization's skills in dealing with the problem identified. In this phase, methods are used to increase the learning process in the organization.

An overview of the Action Research process can be seen in Figure 9.

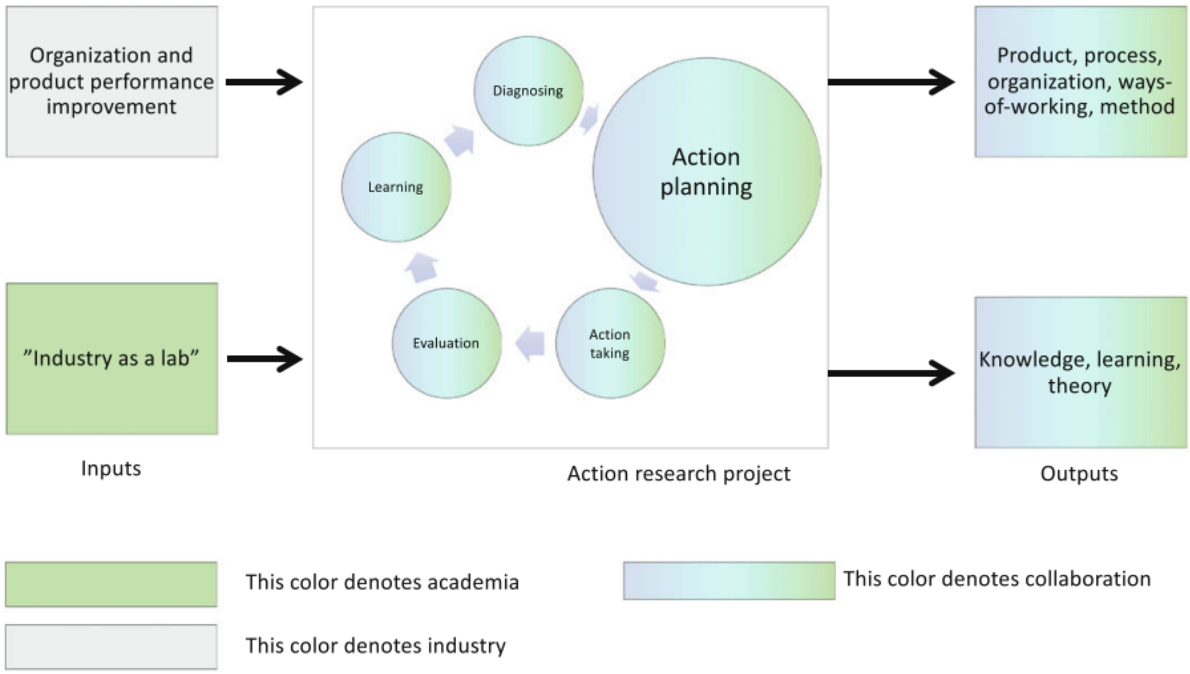


Figure 9 – Action Research in Software Engineering (Staron, 2020)

The diagnosis phase was used as a form of data collection, for observation, semi-structured interviews, and data collection made directly from the project management system used by the institution where we ran the Action Research. The search for more than one source of information in the project proved to be necessary as the research which evolved within the company also proved beneficial because it is a way to combat threats to the validity of the research. This was accomplished by triangulating data sources.

For the thematic analysis, CAQDAS software (Computer-assisted Qualitative Data Analysis Software) was used, specifically Atlas.it[1] because it has a very smooth learning curve and allows limited use without purchasing an expensive license. We conducted semi-structured interviews in both face-to-face and online formats using online meeting software.

APPENDIX B shows the protocol and a more general structure of the interviews.

Thematic Analysis is a method to identify, analyze, and report patterns (themes) from data collected (BRAUN; CLARKE, 2006). Thematic Analysis was chosen because it is a valuable method used in participatory research, where participants act as collaborators during research

[1] https://atlasti.com/

and analysis (ROBSON; MCCARTAN, 2016). Another advantage is that it is an analysis method accessible to researchers with little or no experience in qualitative research, which is undoubtedly the reality of a doctoral student writing his thesis. In addition to being a relatively easy and quick method to use, it is also appropriate for the planning made for a thesis that must be carried out in a specific period of time. However, the flexibility of the method means that there is a wide range of possibilities which makes it difficult for a researcher to decide what data to focus on.

The analysis helped us meet the objective of the diagnostic phase, which is to explore the problem and certify that a solution can be found for the problem identified and thereafter to proceed to the action planning phase of our study.

The Action Research results are presented in Chapter 5 of this thesis.

3.3.4 Step 4 - PoCR

Finally, the last step deals with evaluating the model developed in the previous step, with specialists, the output being the production of a second version of the proposed model.

Kendig (2016) defines 'proof of concept' as research carried out in its early stages on new or cutting-edge technologies. It also highlights that 'proof of concept' can be used for search results that can be scalable or search results that have the potential to be extended.

According to the Oxford English Dictionary, 'proof of concept' is a term associated with evidence collected, usually from an experiment or pilot project, thus demonstrating that the proposed idea or concept is feasible.

The Proof of Concept Research (PoCR) assesses whether a prototype, model, method, or tool has achieved the desired results. By using this assessment, it is also desirable to identify improvement points that can take prototype development to a new level beyond the research level (ELLIOTT, 2021).

As part of the PoCR application process, it was decided to conduct a set of semi-structured interviews to evaluate the first version of the model to quantify the epistemic uncertainty.

In this phase, three proofs of concept were carried out using three teams from three companies where agile management techniques were being applied to develop software projects. These proofs of concept consisted of the practical use of the first version of the Euler model (alpha). By holding meetings with project teams, it was possible to explain the concepts associated with risks and uncertainties and apply the existing processes in the first version of the model. Suggestions for changing the model emerged, such as adjustments to the nomenclatures and new processes resulting from this application. These suggestions make applying the model more fluid and the terms used in the model clearer for project management practitioners, especially concerning the process of managing uncertainty in software projects. The PoCR protocol and results are presented in Chapter 6 of this thesis.

Regarding the type of interview, choosing a semi-structured interview with open questions, even following a pre-defined protocol, allows the researcher to make adjustments to the questions as the interview develops. For the development of the interviews, several aspects were taken into account. These include preparing consent documentation to be signed by respondents, the researcher's preparation as he/she needs to rehearse how to conduct the interview, and preparing technical material so as to record the interview and subsequently to compile an audio file for analysis (ROBSON; MCCARTAN, 2016).

3.4 CLOSING REMARKS

This chapter presented the basic concepts of philosophical roots, the research design and research methods used in this thesis. Thereafter, we showed the method used in this research.

4 EPISTEMIC UNCERTAINTY QUANTIFICATION: QUASI-SYSTEMATIC LITERATURE REVIEW

This chapter presents a quasi-Systematic Literature Review developed to find approaches to quantifying epistemic uncertainty in software projects. It will give us knowledge of what approaches and scenarios are being used to reduce epistemic uncertainties and we shall also gain broad knowledge of studies related to quantifying epistemic uncertainty that are found in the literature.

4.1 QUASI-SYSTEMATIC REVIEW PROCESS

Epistemic uncertainty has been widely studied, and many approaches have been proposed to deal with this type of uncertainty (HELTON; OBERKAMPF, 2004). However, in practice, few of these methods have been used in industry (WHEELER, 2010). Some of these approaches consider evidence-based qualitative criteria. Other approaches are based on numerical methods to perform analyses and calculations with probability distributions that describe the uncertainties in a given scenario. A quasi-Systematic Literature Review (qSLR) was conducted on the protocol set out by Travassos et al. (2008) and Matalonga, Rodrigues e Travassos (2017).

4.1.1 Search Environment

Since 2000, research related to Uncertainty Quantification (UQ) has been highly valued and strongly supported by academia and industry and has become an important research direction (ZHANG; YIN; WANG, 2020). The time criterion for choosing articles to include in this qSLR (2000 to 2020) considered when interest in research on Uncertainty Quantification started.

4.1.2 Research Questions

For Keele et al. (2007), specifying the research questions is essential for the entire development of any systematic review. Similarly, what strategies will be chosen to identify primary studies, data extraction, and analysis process must be included.

Therefore, based on the context presented above, the research questions investigated by this qSLR are:

- qSLR-RQ1 - What are the approaches to quantifying epistemic uncertainty in software projects that are already in the literature?
- qSLR-RQ2 - In which context of project management have these approaches to quantifying epistemic uncertainty been applied?
- qSLR-RQ3 - How can these approaches to quantifying epistemic uncertainty be applied to agile projects?

These research questions align with the primary and specific objectives of this thesis.

4.1.3 Search Strategy

The search strategy executed by this quasi-systematic literature review used seven search engines (academic databases) for automatic search and snowballing research purposes.

The selection criteria for the search engines were based on selecting the most relevant ones identified by Dyba, Dingsoyr e Hanssen (2007) and on the opinion of fellow researchers from our research group at the Federal University of Pernambuco.

The selection features some of the most relevant and popular research engines for Computer Science. Despite the relevance of each search engine, it is clear that some search engines have limitations in how they perform searches. These limitations can threaten the validity of the research. It is expected that the application of quality criteria can mitigate these risks.

The search engine called the Google Scholar database was not selected because we understand that it is an indexer of other databases and that its content would present a larger number of duplicated papers than other engines.

A manual search was conducted in a journal related to the research topic. This journal is the International Journal for Uncertainty Quantification[1].

Finally, the snowball method was applied backward and forward to identify relevant literature among the reviewed articles (WOHLIN, 2014).

Table 2 shows the search engines used by this thesis.

The strategy PICO - Population, Intervention, Comparison, and Outcome - was used to structure the string elements, followed by the most relevant terms and synonyms for this search to construct the search string.

[1] http://uncertainty-quantification.com/

Table 2 – Research Engines

ID	Name	Link
1	IEEE Xplore	http://ieeexplore.ieee.org
2	ACM Digital Library	http://dl.acm.org
3	Scopus	http://www.scopus.com
4	Wiley Online	http://onlinelibrary.wiley.com
5	Web of Science	http://webofknowledge.com
6	Engineering Village	https://www.engineeringvillage .com
7	Elsevier ScienceDirect	http://www.sciencedirect.com

For this research, the comparison criterion, the 'C' of PICO, is represented by an empty set, thus characterizing what Travassos et al. (2008) understood as being a quasi-Systematic Literature Review. This strategy helped assemble the search string shown in Table 3.

Table 3 – Key terms following PICO

PICO	Relevant Terms	Synonyms
Population	Project management	"project management" OR "software project management"
Intervention	Approaches	“model” OR “metric” OR “guideline” OR “check- list” OR “template” OR "approach” OR “strategy” OR “method” OR “methodology” OR "tool” OR “technique” OR “heuristics” or "theory"
Comparison	-	-
Outcome	epistemic uncertainty quantification	

A note regarding our understanding and use of the word 'approach' throughout this thesis: by 'approach', we intend to identify uncertainty quantification techniques, strategies, tools, theories, and a wide range of ways of dealing with uncertainty quantification found in the literature.

4.1.4 Paper Selection

Inclusion and exclusion criteria were created to improve the quality of collecting articles. This study has one inclusion criterion and nine exclusion criteria.

Tables 4 and 5 shows the inclusion and exclusion criteria for this research.

Table 4 – Inclusion Criteria

Acronym	Criteria
IC-01	Primary studies that include epistemic uncertainty quantification approaches in the context of agile project management helping to answer the research questions qSLR-RQ1, qSLR-RQ2, and qSLR-RQ3.

Table 5 – Exclusion Criteria

Acronym	Criteria
EC-01	Studies that are not included in the inclusion criteria.
EC-02	Studies that are not accessible for downloading from search engines via open and institutional access from our university.
EC-03	Duplicated studies, i.e., the same study from different search engines. Just one incidence will be considered.
EC-04	Replicated studies, i.e., studies that are very similar in content and authorship. The less detailed study will be discarded.
EC-05	Studies that are not in the English language.
EC-06	Secondary studies.
EC-07	Tertiary studies.
EC-08	Studies in the format of Posters, Chapter Books, and Technical Reports
EC-09	Studies, including the first author of this document.

4.2 DATA EXTRACTION AND QUALITY ASSESSMENT

One of the essential steps of extensive research in the literature is the textual extraction of data (quotes) from primary studies to answer the research questions.

For this research, the extraction criteria defined by Cruzes e Dybå (2011) were followed based on the findings aggregation according to their selected study context.

In order to register the references of articles and update the evolution of the qSLR, a tool called Jabref[2] version 5.3 was used. Jabref is an open-source citation and reference management software that uses BibTex and BibLaTex formats. The original version of this software was released in 2003, which shows that it is a mature software program. It can be used to create

2 https://www.jabref.org/

groups to organize the articles. Groups were created for the inclusion and exclusion criterion, so when an article was included or excluded, it was also moved to the respective group. When the article was deleted, comments on why it was deleted were also added to the Comments tab for future qSLR review.

Table 6 – Total of articles in each stage of the review.

Stage	Number of articles
Total articles found	204
Articles selected by the inclusion criteria	41
Manual Research	1
Articles kept after full reading	7
Articles included after snowballing	2
Total papers selected	10

Table 6 presents the total of articles in each stage of the review. As can be seen at the end of the process, 10 articles were selected to extract abd synthesize data with a focus on answering the research questions.

In the synthesis phase, 17 quotes were extracted from the articles. 13 of these quotes were responses to qSLR-RQ1 (What are the approaches to quantify epistemic uncertainty in software projects that are already found in the literature?). Four areas of project management were found that answer the question qSLR-RQ2 (In which context of project management have these approaches to quantifying epistemic uncertainty been applied?) and another four quotes answer qSLR-RQ3 (How can these approaches to quantify epistemic uncertainty be applied to agile projects?).

4.3 DISCUSSION

This section presents the results of the research questions defined above. All evidence used to answer the research questions was duly referenced by means of quotes from the ten studies that present evidence collected in the literature. As can be seen, Table 7 presents the results found after qSLR synthesis process. This table presents the nine Uncertainty Approaches (UA) for quantifying the epistemic uncertainty applied in projects found in the literature.

The first column presents a coding for each approach for quantifying epistemic uncertainty

Table 7 – Approaches for Epistemic Uncertainty Quantification

ID	Approaches	In Studies
UA-001	The Dempster-Shafer Theory of Evidence	(SWILER; PAEZ; MAYES, 2009);(TEGELTIJA; OEHMEN; KOZIN, 2017)
UA-002	Coherent upper and lower probability	(TEGELTIJA; OEHMEN; KOZIN, 2017)
UA-003	The NUSAP Scheme - Number, Unit, Spread, Assessment and Pedigree	(TEGELTIJA; OEHMEN; KOZIN, 2017)
UA-004	Interval Analysis	(SWILER; PAEZ; MAYES, 2009)
UA-005	Bayes Linear Method	(REVIE; BEDFORD; WALLS, 2011)
UA-006	Bayesian Belief Networks	(KHODAKARAMI; FENTON; NEIL, 2007); (MENDES et al., 2018); (MANZANO et al., 2018); (FREIRE et al., 2018) (RADU, 2019)
UA-007	C-FASEE - Consistent Fuzzy Analogy-based Software Effort Estimation	(EZGHARI; ZAHI, 2018)
UA-008	Sensitivity Analysis	(REVIE; BEDFORD; WALLS, 2011)
UA-009	Possibility Theory	(HAO et al., 2020)

found in the literature. This coding will help better reference the approach within the text of this thesis. Each approach was given a code based on the acronym UA (for the Uncertainty Approach) and a sequential number. The second column cites the most common name found in the literature. Some approaches can have several names, like the case of Bayesian Networks which can be found as Bayesian Network or Bayesian Belief Networks. The third column indicates the studies where references to these approaches were found. Only studies that underwent the entire qSLR process were used.

One of the findings that emerged from the extraction process was a classification of the approaches used to quantify epistemic uncertainty (lack of knowledge). Swiler, Paez e Mayes (2009) presents a classification of approaches to quantifying epistemic uncertainty. For these authors, there are ways to represent epistemic uncertainty, including probability theory, fuzzy sets, possibility theory, and imprecise probability. They also highlight the challenge of choosing an appropriate approach to represent epistemic uncertainty.

Tegeltija, Oehmen e Kozin (2017) highlight the merit of probabilistic approaches that deal with uncertainties of a stochastic nature (random uncertainty). They also shed light on the need for frameworks that use approaches beyond probabilistic ones regarding uncertainties associated with a lack of knowledge (epistemic). They present three classifications for these types of approaches:

1. Imprecise probability methods allow reasoning of natural language statements, such as 'A is more likely to happen than B';

2. Semi-quantitative methods tackle the problem of communicating complex uncertainty analysis to decision-makers;

3. The post-probabilistic methods decrease the "predict and act" reflection and introduce the "monitor and adopt" model.

It is essential to draw attention to the number of studies that emerged in our research that referred to approaches called the Dempster-Shafer Theory of Evidence and Bayesian Networks associated with quantifying epistemic uncertainty in projects.

That frequency indicated greater attention from the academic community to mainly these two approaches. This demonstrates the mature and well-established use of these two approaches for quantifying epistemic uncertainty in projects. Thus, it is clear that building a model from them would be an evident course of action to follow.

Since nine approaches emerged for quantifying epistemic uncertainty in the literature, it can be concluded that there is an answer to the research question qSLR-RQ1. In addition, it could be perceived that each approach found has a specific field of action related to the problems found and evidenced in this thesis.

It was identified that each approach to UQ contributes to the objectives of this thesis in a specific way. Bayesian Networks can solve the problem of the need to identify and quantify interrelationships between sources of uncertainty. However, Bayesian networks have a bias problem in creating Conditional Probability Tables (CPT). This problem can be mitigated using the Theory of Evidence, which can be used to aggregate experts' opinions in forming the CPT of Bayesian Networks, thus combating the effect of bias coming from individual experts.

4.3.1 qSLR-RQ1 - What are the approaches for quantifying epistemic uncertainty in software projects that are already in the literature?

Bayesian Belief Networks

There are approaches such as Bayesian Belief Networks (UA-006) which, in addition to being concerned with quantifying epistemic uncertainty, also consider the relationships between the various sources of uncertainty identified in projects.

"The most well-established approach to handling uncertainty in these circumstances is the Bayesian approach. Where complex causal relationships are involved, the Bayesian approach is extended by using Bayesian networks (KHODAKARAMI; FENTON; NEIL, 2007, p. 2)."

"The key benefits of BNs that make them highly suitable for the project planning domain are that they explicitly quantify uncertainty and model the causal relation between variables (KHODAKARAMI; FENTON; NEIL, 2007, p. 4)."

Dempster-Shafer Theory

There are approaches such as the Dempster-Shafer Theory of Evidence (UA-001) concerned with the bias introduced by experts quantifying uncertainty in a given domain. To resolve this issue, a mathematical way of aggregating the opinions of these experts has been proposed so that the result of this approach can represent not only the opinion of an expert but the opinion of a group of experts. This aggregation has some limitations, but suggestions for resolving these limitations are being proposed in the literature (SILVA; ALMEIDA-FILHO, 2018) (SILVA; De Almeida-Filho, 2016) (SCHUBERT, 2011)

"This paper has presented a basic overview of three methods that are often used to quantify and propagate epistemic uncertainty in uncertainty analyses.(...) We outlined and demonstrated three methods used in propagating epistemic uncertainties: interval analysis, Dempster-Shafer evidence theory, and second-order probability (SWILER; PAEZ; MAYES, 2009, p. 15)."

Possibility Theory

Another approach to quantifying the epistemic uncertainty that emerged from the qSLR was Possibility Theory (UA-009). Possibility Theory uses two types of measures: possibility and necessity. In practice, each of these terms will represent functions that define the limit above the probability distributions (possibility) and below the probability distributions (necessity).

"Possibility theory is based on the specification of a pair (X,), where (i) X is the set of possible values for x and (ii) is a function defined such that 0 X x) 1 for x X and sup (x):

x X = 1. The function can be seen as a measure of the amount of confidence assigned to each element of x and is referred to as the possibility distribution function for x"(HAO et al., 2020)."

"Rationally, epistemic uncertainty can be described using possibility distribution because a possibility distribution defines a family of probability distributions (HAO et al., 2020)."

Interval Analysis

Interval Analysis (UA-004) is an approach focused on using limits on rounding errors and measurement errors. This technique can guarantee more reliable results by adopting the option of presenting a value, not as a single number but representing each value as a range of possibilities.

"The simplest way to propagate epistemic uncertainty is by interval analysis. In interval analysis, it is assumed that nothing is known about the uncertain input variables except that they lie within certain intervals [6,8]. That is, there is no particular structure on the possible values for the epistemic uncertain variables except that they lie within bounds (SWILER; PAEZ; MAYES, 2009, p. 3)."

Bayes Linear Method

In Bayes Linear Method (BLM)(UA-005), a node represents an uncertain variable, and an arc represents the potential of this node to influence the decision-maker's belief.

BLMs can also be represented by using acyclic graphs such as Bayesian Networks. However, the difference is that the links between the arcs (relationships) between the nodes of a BLM do not necessarily represent causal relationships as with Bayesian Networks.

"The Bayes linear method has features that make it appealing for our problem. First, the Bayes linear method uses influence diagrams to qualitatively structure the problem. Those regularly using these visual representations believe that it is an effective tool for communication between decision-makers and analysts that engenders credibility in the model. Second, if elicitation to structure and quantify the Bayes linear method is carried out transparently, it allows future decision-makers to understand the justification for previous decisions. Third, the Bayes linear method is a sound theoretical approach for combining multiple sources of information in a consistent and mathematically robust way (REVIE; BEDFORD; WALLS, 2011, p. 663)."

This method for quantifying uncertainty is not the same Bayesian Networks method as seen in the quote below, which highlights differences in how belief functions quantify uncertainty, thus making the two approaches distinct.

"The main difference between Bayes linear and conventional probabilistic Bayesian meth-

ods, such as Bayesian belief networks, revolves around how beliefs about uncertain quantities are specified (REVIE; BEDFORD; WALLS, 2011, p. 663)."

"we use the term Bayes linear network (BLN) as a descriptor of the graph representing the model. In a BLN, a node represents an uncertain variable while an arc represents the potential for a source node to influence the decision-maker's belief about the destination node. An arc does not necessarily represent a causal relationship (REVIE; BEDFORD; WALLS, 2011, p. 664)"

The NUSAP Scheme

The emergence of this method for quantifying epistemic uncertainty also suggested a classification for these types of methods. Semi-quantitative methods are methods that, according to Tegeltija, Oehmen e Kozin (2017) are methods considered quantitative that are combined with additional qualitative information as a way to improve the uncertainty analysis. These types of methods more thoroughly tackle the problem of communicating the result of uncertainty analysis to decision-makers. However, some methods, such as the NUSAP Scheme (UA-003), are also used to communicate uncertainty analysis to the lay public.

"From the various semi-quantitative representations that are developed in different fields, we here present the NUSAP Scheme (TEGELTIJA; OEHMEN; KOZIN, 2017, p. 172)."

"The NUSAP Scheme can again be seen as an extension of established probabilistic modeling of uncertainty. It adds qualitative information to the uncertainty and risk analysis in a structured manner, informing the modeling, analysis, and decision-making process by making issues such as data origin, quality, and key assumptions transparent. The acronym "NUSAP" stands for Number, Unit, Spread, Assessment, and Pedigree, the five elements that constitute an information set regarding uncertainty in the method (TEGELTIJA; OEHMEN; KOZIN, 2017, p. 172)."

What stands out in this method is the concern to communicate the result of the uncertainty analysis. This concern is an item that will be observed throughout the construction of the model which is a product of this thesis.

Coherent upper and lower probability

This method is characterized as Imprecise probability. Approaches in this classification, including the Coherent upper and lower probability (UA-002), allow decision-makers to review coherent and plausible ranges of probabilities. This type of approach considers that the expert only has part of the information for his analysis, so the probabilities cannot be precisely known. This approach suggests creating probability measures as accurately as possible, as far as the available information (data) allows.

"In coherent upper and lower probability, the major novelty is the idea to drop a central assumption of Bayesian theory, which states that uncertainty should always be measured by a single (additive) probability measure. There is a large number of arguments which support the concept of coherent upper and lower probability and why it is needed. Given that it does not require unjustified assumptions which is the case with traditional approaches (TEGELTIJA; OEHMEN; KOZIN, 2017, p. 171)"

C-FASEE - Consistent Fuzzy Analogy-based Software Effort Estimation

This model (UA-007) uses Fuzzy Logic Theory to estimate effort in software projects in an analogous way. This means that similar projects (analogs) are observed to estimate similar efforts in the planned project. For this reason, this model is classified as analogy-based. This model has two main characteristics: it presents consistent criteria in representing attributes by fuzzy sets and assumes that similar projects make similar efforts, thus introducing a relationship to measure the extent to which these results are reliable.

"the advantages of C-FASEE representation of uncertainty against the prediction interval approach is that our approach results from an estimation model. Hence, the uncertainty in the estimation can be reduced by optimizing the parameters of the estimation model. Also quantifying the uncertainty using fuzzy sets associated with terms like low, medium or high is more interpretable and less ambiguous (EZGHARI; ZAHI, 2018, p.544). "

4.3.2 qSLR-RQ2 - In which areas of project management have these approaches to quantifying epistemic uncertainty been applied?

It can be seen from Table 8 that the application of approaches for quantifying epistemic uncertainty in the areas of project management, in response to qSLR-RQ2, was in project quality management, as well as in scope management. One of the oldest studies found in our research applied UQ to project scheduling management. Evidence of epistemic uncertainty quantification applied to software production effort planning emerged in two studies.

4.3.3 qSLR-RQ3 - How can these approaches to quantifying epistemic uncertainty be applied to agile projects?

This section presents the evidence following our research in response to qSLR-RQ3. Nine approaches were found and applied in the context of general management. One of these ap-

Table 8 – Approaches by PM Areas

Approach ID	Project Management Areas	In Studies
UA-006	Quality Management	(MANZANO et al., 2018)
UA-006	Scope Management	(FREIRE et al., 2018)
UA-006/UA-007	Software Effort Estimation	(RADU, 2019); (EZGHARI; ZAHI, 2018)
UA-006	Project Scheduling	(KHODAKARAMI; FENTON; NEIL, 2007)

proaches, Bayesian Networks, stands out concerning application in contexts of agile methodologies. We found the application of BN to improve teamwork, process improvement, and productivity prediction of the teams.

"The company providing the use case develops distributed systems for telecommunication networks using a release-based development process based on agile and lean principles (MANZANO et al., 2018, p. 3)."

Within the agile development process with the application of lean principles, the purpose of using at least one of the approaches to Epistemic Uncertainty Quantification evidenced in the literature was to apply Bayesian Networks to improve teamwork. This process of improving teamwork was obtained by seeking to improve the way of evaluating the quality of the work produced by the team. For this purpose, a Bayesian Network was created to quantify the relationships between the key factors that impact the quality of teamwork.

"we propose a Bayesian networks-based model to assess the teamwork quality of agile teams. To build the model, we used the factors presented in Table 2 as a foundation. However, since there are many different agile methods, it is not possible to build a model for all cases due to the differences between these methods and the context of the projects (FREIRE et al., 2018, p. 123)."

However, even for this application of Bayesian Networks to agile processes, lean principles suggest the challenge related to the number of different agile methods, thus making the construction of the model complex.

"We decided to use Bayesian networks due its capability to handle uncertainty [17], formally represent knowledge and the easiness to model and quantify the relationships between the key factors that have impact on the teamwork quality (FREIRE et al., 2018, p. 123)."

Another study proposes the application of Bayesian Networks in evaluating the effort applied to agile software development. This work explicitly speaks of the agile process XP -

Extreme methodology[3]. This model looks for process improvement, and productivity prediction of the teams.

"we propose a BN-based model to assess the effort prediction in the case of developing software with agile methodologies. Researchers have applied such a model in software engineering for process improvement, productivity prediction of the teams that use XP - Extreme Programming, risk evaluation and product quality prediction (RADU, 2019, p. 240)."

4.4 CLOSING REMARKS

This chapter presented the results of the quasi-Systematic Literature Review carried out with the aim of finding approaches for quantifying epistemic uncertainty in projects.

[3] http://wiki.c2.com/?ExtremeProgramming

5 A MODEL FOR QUANTIFYING EPISTEMIC UNCERTAINTY IN SOFTWARE PROJECTS: AN ACTION RESEARCH INVESTIGATION

This chapter presents how action research was used to investigate the impact of epistemic uncertainty on members of a real software development team in a public sector context to understand how they deal with quantifying epistemic uncertainty.

5.1 DIAGNOSIS

Meetings with practitioners and researchers (our Action Research team) were held with academics and professional practitioners to discuss actions to address the diagnosis phase in Action Research problem.

The methods used in this action research phase were Data collection in the project system, Participatory Observation, and Interviews. The use of three methods in Diagnosing can be observed as a form of triangulation to combat threats to the validity of this research. Participatory Observation allows the researcher to be part of the team and to observe the team simultaneously and it is very popular in surveys of types of action research (STARON, 2020).

5.1.1 Context of the Company

This Action Research was conducted with an agile software development team at the State Prosecution Service (SPS), based in the state of Paraiba, located in the Northeast of Brazil. The SPS consists of independent public prosecutors at the state level, which means that the SPS operates independently (financially and administratively) from the three branches of government (the Executive, the Legislative, and the Judiciary).

According to Article 127 of the Federal Constitution of Brazil, Caput: The Public Prosecutor's Office is a permanent institution, essential to the jurisdictional function of the state, responsible for defending the legal order, the democratic regime, and indispensable social and individual interests.

Each state of the federation has a State Public Prosecutor's Office that has the same function as the federal (SPS) but with scope in their respective States. Each state SPS is made up of several members with training and expertise in law. This number is most often equivalent to the number of judges in the respective Courts of Justice of each State. Currently, the Paraíba SPS has 215 prosecutors. Likewise, there is a set of people who comprise what is called the Auxiliary Staff of the institution. There are currently auxiliary staff among public servants, transferred from other bodies and occupants of commissioned positions (these are temporary posts and postholders have not passed an entry exam to the public service).

Each Public Prosecutor's Office has administrative and financial independence and therefore has departments that assist in its administrative tasks, such as the Administrative Directorate, and the Financial Directorate.

The author of this thesis is a public servant and is assigned to the Directorate of Information Technology of the Public Prosecutor's Office of the State of Paraíba.

The main task attributed to the Information Technology Board is to develop systems to assist SPS prosecutors by providing technological solutions, in addition to providing computational resources such as computers and other peripherals. The Directorate of Information Technology currently has 30 auxiliary staff (between public tenders and commissioned). Of these 30 auxiliary staff for Information Technology, seven are developers. The System Development area operates within the SPS by developing and maintaining support systems for prosecutors both in the end area of the SPS (virtual system of processes and investigative systems such as the Pandora) and in the middle area (administrative systems such as the financial one).

5.1.2 Context of the Project

The Pandora software is a web-based application (in NodeJS [1]) that began to be developed in 2016 to automate a recurring demand for information such as updated addresses of people investigated in operations by members (the name given to prosecutors and attorneys) of the SPS. Access to the most up-to-date address is paramount to delivering documents such as subpoenas to appear in court (in procedural law, this refers to listening to witnesses or parties to legal proceedings).

This demand for addresses, which used to be met manually by consulting the most diverse databases, started to be automated by developing software that allowed each member and employee of SPS to access information without asking staff from the Department of Information Technology.

[1] https://nodejs.org/en/

As time went by, the need to add more features to the software arose, which also increased the complexity of managing the Pandora software development project. In 2019, the Pandora project underwent a complete restructuring which meant that it could no longer be the investigation and intelligence software of the SPS of Paraiba and it became a system that would serve the paraiban judiciary, public prosecutors, and the Secretariat for Security in the State of Paraiba.

Currently, the Pandora system can be defined as an analytical and consultative software developed by the SPS IT team to assist the prosecutors of the courts to make their work more effective and efficient. Secondary sources are used to obtain information such as telephone numbers, addresses, and vehicle number plates and information related to family relationships and risk typologies (this type of risk typology is associated with the risk of a person committing fraud or crime of a specific nature). The software uses data mining techniques and artificial intelligence to produce high precision and accurate reports.

The user's direct action of the software helps through the speed and economy in the search for information that helps the prosecutor, magistrate, or intelligence agent's performance since countless searches can be done directly in the Pandora software without the need for intermediaries.

Table 9 presents the composition of the Pandora project team. All of the participants are active members of the Pandora project.

Table 9 – Team IDs and description of roles

Team ID	Role	Experience(years)
P1	Scrum Master	7
P2	Developer	3
P3	Developer	3
P4	Database Administrator (DBA)	3
P5	Infrastructure Engineer	4

The software requirements are approved by a P.O (Product Owner) represented by the coordinator of the investigation department and prosecutor. The coordinator or PO did not participate directly in the Action Research due to the lack of an agenda.

From the table presented, it can be seen that the participants of the Pandora project have at least three years of experience in the area of systems development and Software Engineering. All have a degree in Computer Science or related fields. Two of them, P1 and P4, have a postgraduate degree in Project Management and experience in Software Engineering prior to the Pandora project. P5, an Infrastructure Engineer, graduated in Computer Networks.

The author of this thesis is also the project manager or Scrum Master (P1) from an agile perspective.

The development process follows agile methodological practices such as Scrum and Kanban (BRECHNER, 2015)(BECK et al., 2001). The most common agile practices are the development through sprints that occur every fortnight, meetings of revision of the sprint to identify points of improvement, and the definition of a backlog of features to be developed during a sprint. These tasks are monitored by means of a Kanban implemented with the Trello[2] tool and monitored by all project members.

The Pandora system currently indexes 23 databases. This database consists of open sources and other data sources acquired via agreements. In Paraíba, Pandora is used by five government agencies, and other SPSs have requested its source code so that they can instal it for installation in their infrastructure. As to the dimension of the software, Pandora currently has 839 files and 85398 Lines of Code - LoC (data collected from the April 2021 report).

5.1.3 Description of the Problem

The project described here is not the first software development project undertaken by the SPS team. Other projects have been developed using traditional project management methodology. However, they were unsuccessful from the point of view of achieving the objectives of the project, of complying with the schedule and the costs involved in project deliveries. One of the characteristics present in the projects carried out by this team is that generally, not all software requirements are known a priori, thus making it necessary to use a methodology more suitable for this type of project.

After conducting some research, one of the project team members and the project manager identified the need for a more focused approach to agile project management. Thus, we adopted the use of practices such as sprints, the use of a Kanban, retrospective Sprint meetings, and other practices that could lead to greater probabilities of success for the projects in which the team was involved.

A meeting held among study participants identified that the State Prosecution Service does not have a formal process for identifying uncertainty applicable to its software development

[2] https://trello.com/

projects. This activity has been performed in an exploratory manner by the project manager and a few team members, which does not offer security for the continuous development of the department's activities. Thus, an investigation was proposed to explore how the agile team manages the uncertainty in software projects since the team's software development activities are paramount for the smooth running of the systems that the team makes available.

Despite the efforts of the entire team, it has failed to identify sources of uncertainties such as those reported in Marinho, Sampaio e Moura (2017). Among these sources of unidentified uncertainties, we can highlight technology and social uncertainties which, if not properly managed, can impact the success of the project (MARINHO et al., 2015a).

There is a need to understand how to measure uncertainty to better manage the uncertainty associated with the software project (MARINHO; SAMPAIO; MOURA, 2014). The team needs to quantify the amount of uncertainty in the project, and it also needs to find ways to take into account the impact of existing dependency and interdependence relationships between the various sources of uncertainties in the project.

Furthermore, to better contextualize the impacts of not considering quantification strategies for uncertainty management, authors like Chenarani e Druzhinin (2017) report the significant impact of uncertainty on project goals, quantification, and that monitoring can be very useful and informative for project managers. Also Hester (2012) declares that it is difficult to assess the accuracy of the expert's input with regard to managing the quantification of uncertainty. Jakeman, Eldred e Xiu (2010) add that little attention has been given to quantifying epistemic uncertainty.

Thus, the project manager proposed conducting action research to investigate how managing epistemic uncertainty and taking into account its quantative aspects, can contribute to the project's success. The SPS coordinator accepted the proposal. The entire project team also agreed to participate in the action research.

The Diagnosis process of this Action Research had three data collection methods: observation, semi-structured interviews, and data collection directly from the project management software used by the agency. The observation process allowed us to monitor and record the team's software development process in a more focused way. Data collection through the Redmine3 project management system allowed us to identify uncertainties previously registered in Redmine. The semi-structured interviews contributed to the diagnosis by making use of Thematic Analysis, where identified patterns and themes were observed.

Table 10 presents the quotes taken from the transcripts of the semi-structured interviews and the themes that arise from the Thematic Analysis process.

Table 10 – Patterns identified in the Thematic Analysis

Participant	Citation	Theme
P2	"I think it is limited, nowadays, our role in managing risk and uncertainty, even for a matter of change of scope. Sometimes within the sprint **there are several demands, and we end up leaving aside** some things that should be finished."	Constant changes
P2	"I think I could have as there are these daily meetings in Scrum; maybe they have something like that to let everyone know what is going on."	Lack of Transparency in the process
P4	"There is this point of concern, as the process is not formal. It is more of an individual thing."	Absence of a formal process
P5	"Answering the question whether there is a formal uncertainty management process? When it comes to infrastructure, I say no. Despite improvements always being provided to reduce the risks that exist for the infrastructure as a whole."	Absence of a formal process

The first column presents a coding used to identify the participants in the semi-structured interview process. The second column presents the quote translated into English. Finally, the third column presents the theme from the thematic analysis process. The absence of participant one (P1) is justified because he is the researcher who applied the semi-structured interview protocol and the author of this thesis.

The protocol of the questions asked in the semi-structured interviews can be found in APPENDIX B.

The author respected ethical considerations before conducting the interviews. The participants signed a consent form, and a brief explanation was given about what actions (interventions) would be expected from each team member.

Thus, it is concluded that the team needs to understand how to measure epistemic uncertainty to better manage the uncertainties associated with the project. The team needs to quantify the amount of uncertainty in the project. They must find or develop a new approach to quantifying uncertainty in software projects.

5.1.4 Research Subject

According to the scenario presented above, the investigative object of this Action Research is, in a real project scenario, how agile teams approach the quantification of epistemic uncertainty in software development projects.

5.2 ACTION PLANNING

In this section, the action research planning phase will be described. In this phase, we need to specify what actions will be taken, when they will be taken, and how they will be measured in the organization. Furthermore, to better evaluate the result of the action, it is recommended that the status before and after the execution of a particular action to assess the impact of the action be recorded. Always with the principle of maximizing the impact of our activities, thereby demonstrating the value of the actions developed by action research for the organization.

In addition to the literature review, semi-structured interviews were planned and conducted with the Pandora project team in collaboration with the coordination unit of the project. The main objective is to understand a little more about the point of view of project participants, how projects are managed in the institution, how the (agile) software development process happens, and especially how the systems development team perceives the management of project uncertainty.

During this research, five sprints were followed in the software development cycle of the Pandora project. These five sprints are listed in Table 11.

Table 11 – Project Sprints and Duration

Cycle	Date Interval	Sprints Covered
1	September 1st to September 15th, 2020	Sprint 1
2	September 28th to October 13th, 2020	Sprint 2
3	October 19th to November 2nd, 2020	Sprint 3
4	November 9th to November 23rd, 2020	Sprint 4
5	November 30th to December 14th, 2020	Sprint 5

Sprints are agile Scrum process events with a fixed duration of one month or less that aim to turn ideas into value. A new Sprint starts immediately after completing the previous Sprint

(SCHWABER; SUTHERLAND, 2011).

This Action Research consisted of three cycles. As shown in Table 3, the first cycle followed the sprint from September 1st to September 15th, 2020, totaling 15 days and was called Sprint 1 and so on.

The focus of the action research was to establish approaches to manage and quantify epistemic uncertainty in a software project and to generate a model to manage them. Therefore, based on the context presented above, the research question that this thesis investigates is *How do agile teams approach quantifying epistemic uncertainty in software projects?*. As expected, results were used to create a software model that project managers can adopt to quantify and monitor epistemic uncertainty.

This action research used semi-structured interviews, follow-up meetings, retrospective meetings, and follow-up activities as an operational aspect. The thematic analysis techniques were applied to the results of the semi-structured interviews as a way to identify patterns (themes) that could help answer the research question. Online tools were used to support the action research, such as Wiki software (Media Wiki), in order to centralize the knowledge produced during action research. Examples include audio transcription tools within the structure of Google Docs and Atlas.ti[3] to visualize relationships and to identify themes from the texts of the transcripts.

Data were collected and measured within the sprints and by using the Project Management System (PMS). The PMS is used by the team to record the demands and activities of the project as well as to record the uncertain events that occurred throughout the project.

5.3 ACTION TAKING

The presentation of the results uses the format suggested by Staron (2020) for Action Research studies in Software Engineering when he adapts the format of Runeson e Höst (2009) and Host, Rainer e Regnell (2012) for case studies. This type of presentation is focused on results rather than the type of formatting focused on cycles.

As previously reported, the literature search was carried out in a quasi-systematic way to investigate which approaches can be found in the literature to quantify epistemic uncertainty in research project software. The result of this research supports the actions developed and reported on in this section.

[3] https://atlasti.com/

According to Staron (2020), an action can be any intervention made in the context of a study in action research. Therefore, it can be as complex as a change in the formula of a study, or as simple as an analysis and presentation of the results to the members of the research team and the company team. Simply collecting data from the project management system can be considered an action.

It can be understood that there are types of actions that represent changes in the process, actions that represent data analysis outside the primary process (in a test environment, for example) for later comparison of the results produced between the test environment and the production environment (actual data). Also, some actions can represent inputs so significant that they can cause changes in the structure of the organization as a whole.

Action is carried out based on data collected and the context of the action research. Another feature is that it is done in collaboration with the various parties involved in the research. They also are aligned with the problem diagnosed in the diagnostic phase of action research. Lastly, it is also based on the literature review previously conducted.

Each action reported below was designed with Staron's recommendation in mind that action requires preparation, execution, and data collection.

This highlights the importance of clearly defining a way of measuring each action so we can be sure of the effects achieved at the end of its execution. The instrument that is the main element of the action also stands out. This instrument can be defined as the element causing the expected effect. The instrument can be understood as a new set of codes (when the action is to improve the testing code) or an Excel spreadsheet when an action is exemplified to improve the way of reporting quantitative reports.

The actions reported in this section have information about in which cycle the action was carried out. However, we believe a better reporting form would be oriented towards actions, not action research cycles.

5.3.1 Action 1 - Research approaches for quantifying project uncertainties

As reported in the description of the problem section and identified by applying the Thematic Analysis technique in the transcript of the semi-structured interviews present in Table 10, there is no formal process or model for managing or quantifying epistemic uncertainty in SPS projects. The absence of this type of approach led to problems such as delays in the delivery of project activities and losses associated with the need to re-plan actions and activities of the

Pandora project.

By conducting semi-structured interviews, the project team could be asked if they had a notion of uncertainty. For conceptual alignment, a presentation of themes associated with uncertainty and risks present in the literature was made. Participants were asked whether, now with an understanding of the concept of uncertainty, they could recall the occurrence of uncertainties that occurred during the Pandora project. The following actions will present the identification of uncertainties that occurred at past moments of the project.

This action was associated with reviewing the strategies for identifying and understanding the uncertainties that occurred in the project. The conclusion drawn from analyzing the questionnaire given to the project team is that they did not take any approach to manage or quantify the Pandora project's uncertainties.

Another action taken was to carry out a quasi-systematic literature review to seek approaches present in the literature that can be adapted and adopted by the project team to quantify and relate sources of uncertainty. At the end of this action, it was concluded that what must be done is to adapt or build an approach to quantify epistemic uncertainty in software projects.

5.3.2 Action 2 - Identify sources of uncertainty

After identifying the approaches present in the literature for quantifying epistemic uncertainty, the need arose to identify the sources of uncertainty present in the project. To help the process, we searched the literature for sources of uncertainty that could support the process of identifying uncertainties.

According to Marinho, Sampaio e Moura (2017), Marinho, Sampaio e Moura (2013), Marinho, Sampaio e Moura (2014), Marinho (2015), sources of uncertainty are Technological, Market, Environmental, Social, Human and Organizational issues. The sources of uncertainty could guide the project team members' search for and identification of uncertainties.

At the end of this action, the team can count on four sources of uncertainty that could be used in the following action to identify the uncertainties in the Pandora project.

5.3.3 Action 3 - Review the team's experiences of project uncertainties

Based on the sources of uncertainty identified in the literature and which could support the identification and recording of uncertainties in the Pandora project, an action was initiated to identify the uncertainties that occurred or may occur in the Pandora project.

The analysis of the transcript of the semi-structured interviews with the project team generated six uncertainties from Technological, Human, and Environmental sources.

The analysis of the project management system used by the team to manage the activities of the Pandora project also contributed to identifying uncertainties. From this system, three uncertainties can be collected from the sources of technological and environmental uncertainty. Observing the project's sprints contributed to five more uncertainties from the Technological, Environmental, and Socio-Human sources. All these uncertainties mapped in the Pandora project are set out in Table 12.

As an output of this action, Table 4 presents the uncertainties identified in the Pandora project. The first column presents a numerical sequence to identify the uncertainties identified and registered in the project. The second column presents the instrument used to collect and identify uncertainty. This instrument can be the data collection method, such as the transcription of semi-structured interviews or the method of observing a sprint. Another form that we call an instrument was the collection of previously recorded uncertainties or records of uncertain events that meet the definition of uncertainty adopted in this thesis which was carried out directly in the project management system used by the team to record the demands and activities of the project.

The third column presents the classification of the source of uncertainty among the classifications described and adopted for this thesis. The fourth column comprises the uncertainty identified in this action preceded by an uncertainty identification code (The letter U refers to the word uncertainty). This codification has two purposes: to facilitate the reference to uncertainties when this thesis was being written and the reference to the uncertainty in creating Bayesian Networks that will be presented in the following actions.

It can be observed that all uncertainties came from sources of uncertainty identified in the literature. However, the lack of uncertainties associated with the source of uncertainty called Market is justified because the project is carried out in a public company that, by its nature, does not sell or commercialize any type of product type.

At the end of this action, the team had a catalog of 14 uncertainties identified with the

Table 12 – Sources of Uncertainty

Cod.	Instrument	Uncertainty Source	Uncertainty
1	Semi-structured interviews	Technological	U1 - Lack of expertise in the technology/programming language used in the project.
2	Semi-structured interviews	Technological	U2 - Lack of formal process for managing uncertainty.
3	Semi-structured interviews	Socio-Human	U3 - Insufficient team to meet the growing demands of the project.
4	Semi-structured interviews	Socio-Human	U4 – Turn-over of project team members.
5	Semi-structured interviews	Environment	U5 - Lack of Transparency in the process.
6	Semi-structured interviews	Environment	U6 – Structural damage to the civil infrastructure of the data center.
7	Sprint observation	Environment	U7 -Lack of expertise in the process of adopting new container technology (Docker).
8	Sprint observation	Environment	U8 – Constant requests to change the design.
9	Sprint observation	Environment	U9 – Damage to the civil structure (building) of the data center.
10	Sprint observation	Socio-Human	U10 – Team member assigned to other activities.
11	Sprint observation	Socio-Human	U11 – Partial or total absence of team members because of Covid-19.
12	Project management system	Technological	U12 – Inadequate electrical installation for the data center.
13	Project management system	Environment	U13 – Lack of financial resources despite the budget forecast.
14	Project management system	Technological	U14 – A suspected invasion of the Pandora system.

help of the Pandora project team from three sources of uncertainty. The following action deals with the process of quantifying the uncertainties recorded.

5.3.4 Action 4 - Quantification of epistemic uncertainties

By exploiting the authors' quasi-Systematic Literature Review, approaches for quantifying epistemic uncertainty could be identified. For the need observed in the Pandora project, the approach to quantify the epistemic uncertainty used was to quantify uncertainty by using the Dempster-Shafer Theory of Evidence.

In Evidence Theory, the representation of uncertainty, belief, and plausibility represent a range of potential values for a given parameter without assumptions about the probability of the underlying data.

According to Elmaghraby (2005), project management is prone to uncertainty, whether in estimating resources, terms, or costs. Therefore, the author argues that the question is not how to recognize uncertainty but how to measure it.

DST was chosen as the technique for quantifying uncertainty as it has the advantage of allowing the treatment of uncertainty based on the knowledge of experts (project team), who, based on their experiences, knowledge, and preliminary information about the activities of the design, assign mass to the elements (uncertainties) in the interval [0, 1]. This quantification is associated with each expert's belief in the chosen hypothesis.

Another aspect as to why DST was chosen was the ease with which the authors of this article understood it and the availability of programming libraries using the R language to produce scripts for automating the process of quantifying epistemic uncertainty. The package in R used for coding the script to apply the Theory of Evidence was dst (the name of the package R has the same name as the acronym adopted for the theory) which can be found in its repository on CRAN-R[4] (BOIVIN, 2019).

In this context, this means that the DST can be used to collect the opinions of each project participant (here called experts, as they know the project domain well) to quantify the belief that each of them has as to whether or not the uncertainties identified in the previous actions occur. Following the steps described in Santos e Cardoso-Junior (2018), this action had the following steps:

1st Step: Each specialist in the project team was asked to indicate which uncertainties of the uncertainty table presented in the previous action are more likely to occur according to their current knowledge, prior knowledge, and evidence of similar uncertainties that may have occurred and thus to attribute mass to these elements (uncertainties);

2nd Step: The evidence collected from each expert was recorded in the R script that used the EvCombR[5] package to obtain the results of the belief and plausibility of the occurrence of uncertainties by each expert. In accordance with the indicator in the DST, these masses were

4 https://cran.r-project.org/web/packages/dst/
5 https://search.r-project.org/CRAN/refmans/EvCombR/html/EvCombR-package.html

assigned within the interval [0, 1].

3rd Step: In this step, project uncertainties were monitored so as to follow the evolution of activities and associated uncertainties. If necessary, a new assessment of project uncertainties using DST would be performed.

The script used in this Action Search can be found in the GitHub[6] created to store the source codes in R.

5.3.5 Action 5 - Relate sources of uncertainty

Also, by analyzing the quasi-Systematic Literature Review carried out by the authors, it was possible to identify approaches that aimed not only at quantifying epistemic uncertainty but also at taking into account the interdependence relationships between the various sources of uncertainty identified in the Pandora project. The approach identified in the literature review that best fits this need was the use of Bayesian Networks. A Bayesian Network (BN), also called a Bayesian Belief Network (BBN), Probabilistic Network (PN), or a Causal Network (CN), can be seen as a model that uses Graph Theory and Probability Distribution to model a situation or phenomenon variables, and the states of these variables.

It is essential to know their origin, associated with the Bayes Theorem, in order to understand the importance of Bayesian Networks. Our interest in Bayes' Theorem comes from the fact that it is one of the mathematical ways to characterize learning from experience, i.e., the initial attitude towards evidence and hypotheses can be modified after having the information (evidence) that a given hypothesis can be materialized.

For this thesis, we can understand learning and experience that is perceived from the previous action to be essential for the process of quantifying uncertainty that uses the DST applied in the previous action and the reference to the hypothesis that can be materialized is our uncertainty that can be characterized in something concrete and can impact the project.

The process of building a BN undergoes the following steps Russell e Norvig (2003):

1st Step: Choose a set of relevant variables Xi that describes the domain;

2nd Step: Choose an order for the variables;

3rd Step: Select a variable Xi and add a node in the network;

4th Step: Determine the parent nodes pa(Xi), among the nodes that are already in the network so that the Markov Condition is satisfied;

6 https://github.com/jeffersonjpa/quantifying$_{uncertainty}$

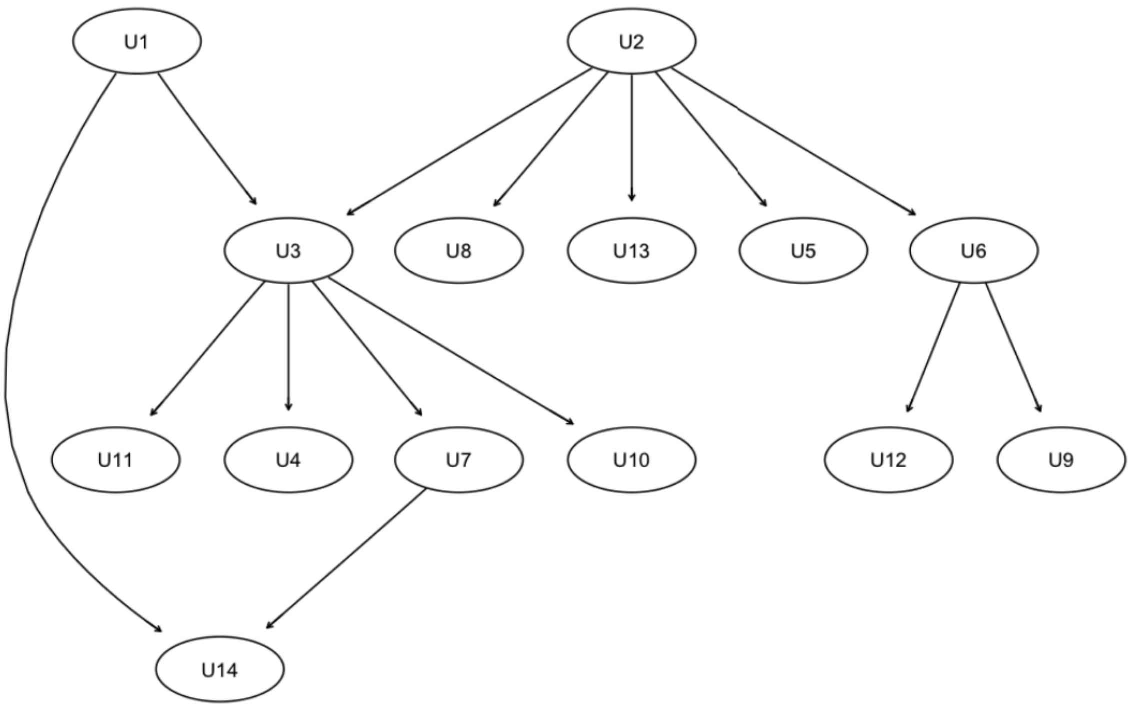
5th Step: Determine the conditional probability table for Xi;

The topology of the network is defined by using the uncertainties identified for the Pandora project and was carried out with the project team to understand which uncertainties (Us) may have some interdependence relationship. The definition of the topology of the BN aims to create graphs that can represent the causal relationships (interdependencies) between uncertainties. The main characteristic of these graphs, called Directed Acyclic Graphs (DAGs), is that they are non-cyclic.

A Conditional Probability Table (CPT) is defined by producing a table of the probability of occurrence of uncertainties in the project. This table of probabilities results from the previous action that used DST to quantify the uncertainties. Thus, the number or interval associated with each uncertainty can be used as CPT and input to the BN.

Figure 10 below shows an example of the BN created during this action. For BNs applied to the uncertainties identified in the Pandora project, code libraries in the R language could also be found. The library for BNs used was bnlearn[7](SCUTARI, 2010).

Figure 10 – Bayesian Network Pandora Project



Source: Author

At the end of this process, 14 uncertainties and their interdependent relationships can be quantified. It would now be possible to prioritize actions associated with the treatment of

[7] https://www.bnlearn.com/

uncertainties since the project team has its measurement and considers the interdependence relationships in the process, which clarifies the relationships between the uncertainties of the Pandora project.

What can be noticed is that we can quantify this interdependence relationship and thus realize to what degree (quantification) an uncertainty may impact or may be impacting another uncertainty as presented by Figure 10.

Questions such as *"How much of the uncertainty U1 impacts or is it associated with the uncertainty U3?"* can be put and answered by using this technique.

This process demonstrates the benefits that the project team can achieve from quantifying epistemic uncertainty in the Pandora project by means of techniques used in the previous actions.

5.3.6 Action 6 - Development of a preliminary model to quantify epistemic uncertainty in software projects

After applying the epistemic uncertainty quantification process using DST and identifying the interdependence relationships between the uncertainties identified, the project team and the researchers had the confidence to propose a model that could be systemically applied in other projects.

The objective of this action was to formalize the creation of a model. The language chosen for formalizing the processes of the model was the Business Process Model and Notation (BPMN)[8].

The model was created taking into account three processes extracted from the previous actions:

1st Step: Identifying uncertainty sources and identifying project uncertainties;

2nd Step: Applying the Dempster-Shafer Theory of Evidence (DST) to quantify the identified uncertainties;

3rd Step: Applying Bayesian Networks to identify the interdependence relationships between the uncertainties identified in the project;

As a result of the first process that deals with identifying sources of uncertainty and identifying uncertainties in the project, the team can count on the support of the sources of uncertainty found in the literature. The output of the first process serves as input to the second

[8] https://www.bpmn.org/

process, which deals with quantifying the uncertainty associated with this set of uncertainties identified in the first process.

The second process produces a set of measured uncertainties that serve as input to the third process, which deals with identifying the existing interrelationships between the sources of uncertainty using the concepts of BNs.

The following action was to apply the newly created model in a sprint. This application served as a pilot for applying the model.

5.3.7 Action 7 -Apply pilot of the developed model to a sprint

After formalizing a model for quantifying epistemic uncertainty in projects, a sprint was chosen as a pilot to apply the model in its preliminary version. The sprint chosen was number 5, which took place between November and December 2020.

The objective of this action was to observe that the use of the model was no longer fragmented as used in previous actions. However, its preliminary version formally joined the output of a process to the input of the following process. At the end of the sprint, the project team and researchers noticed the need to insert two complementary steps into the model.

First, it was discussed that the model should have a stage of interpretation of visible results by analyzing the graph of the BN. As highlighted in the literature, one of the advantages of using BNs is the visual record of the uncertainties is characterized by their nodes (uncertainties) and existing interdependence relationships.

The second discussion revolved around the fact that the model should record the events that occurred in an information base that would function as the project's body of knowledge. This body of knowledge could be used in future projects and would work very similarly to recording lessons learned in projects.

As the last contribution of this action to evaluate the pilot model in a sprint, the need for a model to have a cyclical character along the lines of the PDCA cycle (Plan, Do, Check, Act) (JOHNSON, 2002) was highlighted. The cyclical aspect is characterized by continuously executing the model during the life cycle of the project. It is known that uncertainty can arise at any stage of a project, whether at the beginning or in the final phases.

5.4 EVALUATION AND LEARNING

It is known that according to the theory of the Action Research method, its cycles must be executed until a stopping criterion is reached.

For this Action Research, the constructed model was evaluated by recording the number of new uncertainties identified in the project and the number that represents the measurement of the uncertainty identified. So, this was the stopping criterion for this Action Search.

As a general summary of the Action Research Learning phase, Table 13 shows the scenario before and after applying the Action Research in the SPS and directly compares their respective advances and learning.

Table 13 – Scenario before and after applying the Action Research in the SPS

Before Action Research	After Action Research
Analysis based on expert perception	Analysis based on expert perception using DST to quantify expert perception
Unmapped sources of uncertainty	Uncertainty sources are systematically mapped as there is now a formal process.
Unquantified uncertainties (assuming their potential impact)	Possibility of quantifying project uncertainties, which allows prioritizing those with the most significant potential
Failure to map the relationships that exist between uncertainties and sources of uncertainty	By mapping the relationships between the various sources of uncertainty, one can see dependencies between them.
There was no model or approach to quantify epistemic uncertainty.	Existence of a formal model to quantify epistemic uncertainty

After the Assessment and Learning phase, the research question that guides this thesis can be recapitulated: *"how do agile teams approach the quantification of epistemic uncertainty in software development projects?"*.

5.5 CLOSING REMARKS

This chapter presented the processes of Action Research that aimed to investigate how agile teams deal with epistemic uncertainty in software projects, ending with the presentation of the first version of the Euler model.

6 AN ASSESSMENT OF A MODEL FOR DEALING WITH EPISTEMIC UNCERTAINTY IN AGILE PROJECT MANAGEMENT SOFTWARE

This chapter sets out the process for conducting the empirical application of the Euler model - version 1 (V1) with agile teams. In addition, the method and the results arising from evaluating the V1 of the model are presented.

6.1 EULER MODEL V1 - OVERVIEW

The objective of this chapter is to make a report on the empirical application of the first version (V1) of the Euler model and to evaluate the results. The practical application of this version followed the guide presented in Appendix A. This guide divides the application of the Euler model into its internal processes, which, for V1, are:

1. Identifying the Sources of Uncertainty
2. Applying the Dempster–Shafer Theory of Evidence
3. Applying Bayesian Networks
4. Making a Visual Analysis of Bayesian Networks
5. Feeding the Body of Knowledge

This first version results from applying the Action Research method in the State Prosecution Service of Paraíba (SPS), the objective of which was to investigate how agile teams manage uncertainty in software projects. A description of this Action Research process was published in Barbosa J. F. (2021).

Figure 11 presents an overview of the Euler model (V1) . This diagram emphasizes the cyclical character of the model and the execution in sequence of the internal processes.

Each of the previous processes has its inputs, processing, and outputs well defined as a way to achieve the main objective, namely, quantifying epistemic uncertainty in agile project management software. If this objective is achieved, it is expected that the project manager will be better able to prioritize the uncertainties present in the project, and thus focus on the uncertainties that are considered the most critical, and will be better able to coordinate activities and so achieve greater predictability of the uncertainties and risks associated with the project.

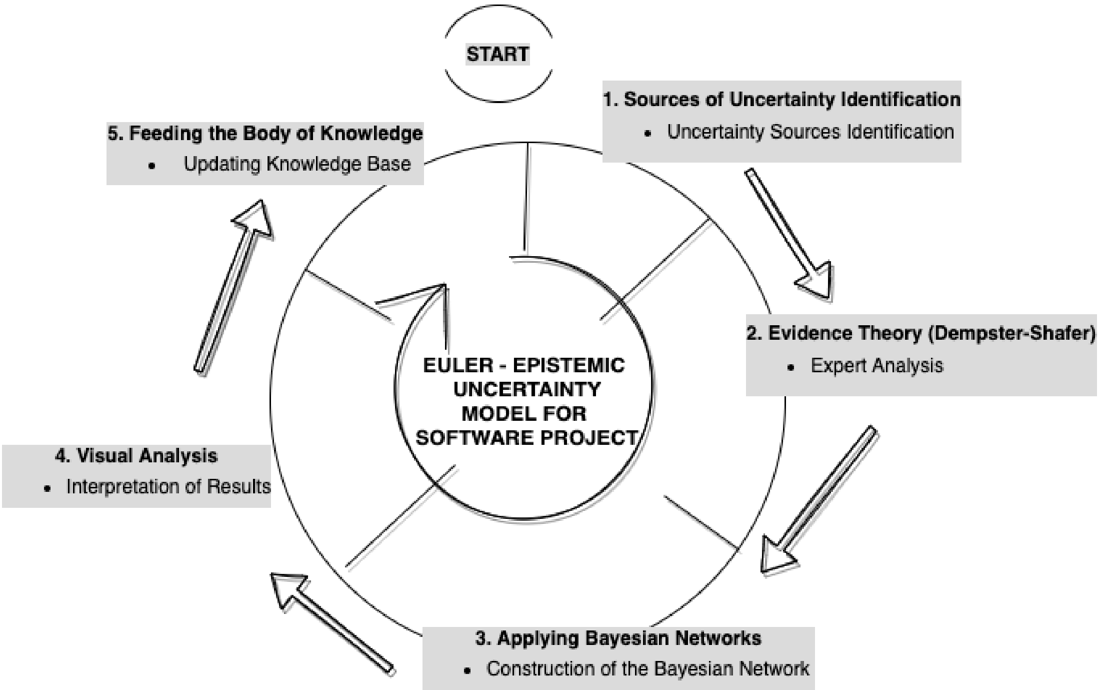


Figure 11 – Euler Overview

The Euler model was written using the notation language BPMN. As can be seen, there are sub-processes within each process for the complete execution of the model's activities and activities associated with a decision process. The need to return to the processes at the beginning of their cycle is verified to identify new uncertainties or proceed to the end of its cycles to complete the process of quantifying uncertainty.

This first step of the model deals with identifying uncertainties in the project. A search was conducted in the literature for sources of uncertainty to assist in identifying uncertainties. These sources can be understood as a driver or guide or an initial step in identifying new project uncertainties. Sources of uncertainty may be Technological, Market-related, Environmental, Socio-Human and Organizational article$_{m}arinho_{s}ampaio.$

The next step of the model, applying the Dempster-Shafer Theory (DST) of Evidence, has the advantage of allowing uncertainty to be dealt with based on the knowledge of experts (an agile team), who, based on their experiences, knowledge, and preliminary information about the activities of the design, assign mass to the elements (uncertainties) in the range [0, 1]. In this step, DST can collect the expert's opinions to quantify the belief that each of them has in whether or not the uncertainties identified in the previous process will occur.

The subsequent step deals with the process of building a Bayesian Network (BN). The following steps are undertaken in this process: (i) Collect Data or the Knowledge of Experts

- The entry process and subsequent construction of a BN can be carried out by collecting data by means of a survey or directly by recording the opinion of experts; (ii) Assessment Variables - This step deals with identifying the variables that will form the BN and become nodes within the topology of the network. The input of this process will be the output of the previous process, as the nodes to be used will be the previously identified uncertainties. The definition of the topology of the BN aims to create graphs (DAGs) that can represent the causal relationships between uncertainties; (iii) Definition of a Conditional Probability Table (CPT) - This step is performed by producing a probability table of the occurrence of uncertainties in the project. This table of probabilities is the result of the previous step that used DST to quantify the uncertainties. Thus, the number or interval associated with each uncertainty can be used in a CPT and input to the BN; (iv) a script was also developed using the R programming language and the bnlearn package to assist in building the BN from the output of the previous process. This script can be found in the GitHub[1] repository created to store the code artifacts produced during this research.

As important as the generation of the Bayesian Network are the BN visualization methods. Network visualization cannot be regarded only as an aesthetic concern. For Tufte (2001), "often the most effective way to describe, explore, and summarize a set of numbers - even an extensive set - is to look at pictures of those numbers." Moreover, the author adds that data graphics can be even simpler and more powerful than any other method.

Potential correlations or causal relationships between variables can be quickly visualized. Studies such as Sundarararajan, Mengshoel e Selker (2013) focus on improving navigation and visualization in large networks with networks that range up to thousands of variables.In order to understand these large networks and make the networks more efficient and better visualized, new methods are needed.

The output of this process is an even richer visual representation of the representation elaborated by a Bayesian Network. This visual representation would not be possible without using the techniques presented in this step. It will now be possible to better visually analyze the whole set of uncertainties and their relationships within the Euler model.

The last process of the Euler model deals with knowledge produced by the processes. The step of feeding a body of knowledge with the result of this process, despite being the last one, is an essential step in order to refine the results presented by the model and to ensure these results are used in future projects. The output of this process is the representation of the

[1] https://github.com/jeffersonjpa/quantifying$_{u}ncertainty$

knowledge produced during its construction and that this knowledge is stored in an accessible way. This knowledge must serve as an input for future application cycles of the Euler model and future applications of the model in new projects.

6.2 PROOF OF CONCEPT RESEARCH

The Proof of Concept Research (PoCR) application aims to assess whether a particular prototype, be it a model, method, or tool, has achieved the desired results. By using this type of evaluation, it is also desirable to identify points of improvement that can take the development of the prototype or model to a new level.

This thesis uses the framework for PoCR proposed by Elliott (2021). The framework is divided into prototypes, proof of concept research demonstrations, and post facto arguments. This thesis adapts Elliot's framework to incorporate a first step for making agile teams aware of concepts of risk and uncertainty that are present in this thesis.

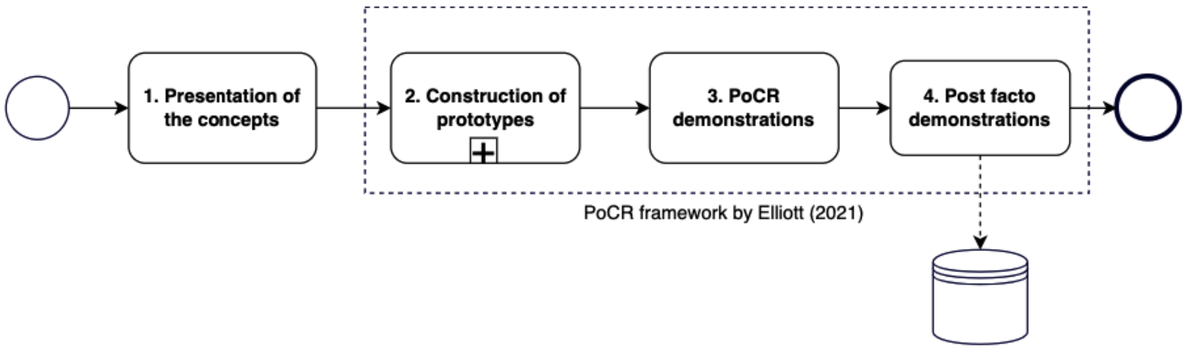


Figure 12 – Proof of Concept Research Steps adapted from Elliott (2021)

According to Elliott (2021), prototypes can be considered new artifacts or have innovative characteristics concerning pre-existing solutions for solving relevant problems in some fields of science. Prototypes can be either physical engineering systems or theories or models. For this thesis, the prototype is considered the first version of the Euler model for quantifying epistemic uncertainty in agile project management software. The following subsections will explain the steps shown in Figure 12.

6.2.1 Step 1. Presentation of concepts

For this evaluation, Elliot's framework was adapted to add an initial step associated with explaining the concepts around risks and uncertainties. This step arose from the researchers' observation when applying the Action Research, where they demonstrated that these concepts (risks and uncertainties) were not common knowledge or disseminated among all practitioners of the process for undertaking the project management of software.

Before the agile teams applied the model in practice, the concepts associated with this research were presented, mainly characterizing the difference between the concepts of risks and uncertainties. For this characterization, references in the literature and examples of uncertainties collected during this research were used. This presentation lasts around 30 minutes and is carried out with the support of a set of slides created for this purpose [2]. It is expected that the output of this process will be that the agile team gains a better understanding of the concepts around managing risks and uncertainties in software projects that use agile practices.

6.2.2 Step 2. Construction of Prototypes

The second step of our evaluation demonstrates the concepts around the prototype mentioned in the first stage. This demonstration shows that the prototype contains the functions needed to solve a given problem and thus can achieve the desired results.

When talking about physical artifacts, the demonstration can be a physical presentation of how the artifact will be used, such as for launching a new rocket. In the case of theories or models considered abstract artifacts, this demonstration can be theoretical but should include confirmation that the artifact produced achieves its objective, which can be the description, presentation, or explanation of a given phenomenon.

6.2.3 Step 3. PoCR Demonstrations

The third step of this PoCR is the application of the Euler model. At this moment, the researcher follows the application of all the steps of the Euler model together with the target organization's agile software development team. The researcher's role is to serve as a facilitator on how to apply the model with the consequent identification and quantification of the

[2] bit.ly/slides$_{p}rotocol_{P}oCR$

uncertainties of the software development project. The output of this PoCR process, which corresponds to the output from applying the Euler model, is a set of uncertainties related to the target project, associated, identified, and quantified for the purposes of prioritization and predictability for better coordination of the software project.

6.2.3.1 Sampling

With regard to conducting an evaluation process in qualitative research, one major methodological decision is how to sample the interviewees. Qualitative research usually works with a small sample of people within a specific context, unlike quantitative research, which works with many cases and seeks some statistical significance miles2018qualitative.

According to Miles, Huberman e Saldaña (2018), a random sample is a good standard for quantitative research but is little used in qualitative research. They also mention that sampling tends to be more strategic when focusing on a specific case with a very characteristic context. They admit that in some contexts, selecting a case to study is made possible as it is the most accessible geographically and temporally (surveys always have a deadline to be completed) and that this form can be understood as sampling for convenience.

Staron (2020) says that subjects of different characteristics should be selected for assessment before and after the research.

To overcome the bias associated with selecting subjects for evaluating the model in practice and the semi-structured interviews, projects were sought with teams from outside the institution in which the first version of the model was created.

It can be seen from the interviewees' profiles that they have different types of years of experience, different positions in the project teams, different levels of education, work in the public (government) and private sector (for profit organizations), and in organizations of different sizes (small, medium, and large).

6.2.3.2 Participants

Seven participants from three companies and projects applied the first version of the Euler model. They are from three groups of agile teams, one coming from a public company while the members of the other two teams are from private companies.

The first team to evaluate the Euler model is from SPS Espírito Santo (SPS ES), northeastern Brazil. SPS ES is a public institution equivalent to SPS PB. The team comprises two members who, in Table 14, are identified as P4 and P5. This team has been working together to develop technology solutions with agile practices for SPS ES for ten years.

The second team to evaluate the Euler model is from a Brazilian startup called FABWORK. FABWORK has been operating for five years and is headquartered in Paraíba, northeast Brazil. FABWORK develops solutions in data science and customizable technologies that improve companies' digital and analytical performance. The team comprises three members who, in Table 14, are identified as P1, P2, and P3.

The third team to evaluate the Euler model is from the Municipality of João Pessoa, the capital of Paraíba. This team has been working on the Pandora project for 12 months, applying agile practices and developing integration solutions between the software of the city of João Pessoa and the Pandora system of the SPS PB. The team comprises two members who, in Table 14, are identified as P6 and P7.

All participants are involved in agile projects and have realized the need for a better approach to identifying and quantifying epistemic uncertainties in their respective projects (Table 14).

Participants were labeled from P1 to P7 as a way to reference them within the process for analyzing results. The assessment was carried out using semi-structured interviews. The participants took part in the semi-structured interviews after using the model in practice to identify and manage uncertainties in real projects in which they were involved at the time of evaluating the model.

The Euler model evaluation sessions were divided into three 90-minute meetings for each project team. In the first meeting (90 minutes), the concepts of Risks and Uncertainties and how to apply the process for identifying uncertainty were presented. In the second meeting (90 minutes), the processes for applying the model were carried out. In the last meeting (90 minutes), the questionnaire was applied while conducting the semi-structured interviews and points of improvement of the Euler model were discussed. These sessions were repeated for each of the three project teams.

Figure 13 presents a summary of the profile of the participants in the process for evaluating the first version of the Euler model.

This survey had seven participants who, in accordance with the summarized results can be characterized as follows: 57.1% of the participants have between 1 and 5 years of experience in software development projects. That of the other participants varies from 6 to 20 years of

Table 14 – Evaluation Participants

Code	Role	Experience (years)	Company Type	Education	Size	Company
P1	Developer	1-5	For Profit organization	Graduate	Small	A
P2	Scrum Master	1-5	For Profit organization	Graduate	Small	A
P3	Scrum Master	1-5	For Profit organization	Graduate	Small	A
P4	Scrum Master	16-20	Government	M.Sc	Large	B
P5	Developer	11-15	Government	M.Sc	Large	B
P6	Developer	6-10	For Profit organization	MBA	Small	C
P7	Product Owner	1-5	Government	Graduate	Small	C

experience in software projects. Regarding the educational level of the participants, 57.1% are graduates, representing the majority, 28.6% have a master's degree, and 14.3% have an MBA degree.

It can be seen from these data that the participants have different levels of experience in software projects, which can benefit our results from different points of view.

Thus, it can be concluded that the respondents are fully capable of carrying out the final evaluation of the Euler model and of contributing to its development by offering criticisms and suggestions for points for improving the model.

At the start of applying the protocol and during its process, the participant is reminded that if he/she does not feel comfortable in terms of his/her knowledge about answering any of the questions, he/she should not leave it unanswered, but should feel comfortable about providing approximate answers because, for this research, it is more critical to give a rough answer than no answer.

All participants are from the Brazilian market. It is well understood that since the solution was built in the Brazilian market for the nuances of Brazilian software development projects, its evaluation should also be conducted with participants from the Brazilian software development market. That said, we are not limiting the application of the model to the Brazilian market. The objective is to maintain consistency between the scenario where the model was conceived (within a Brazilian public agency) and the projects where the Euler model was evaluated

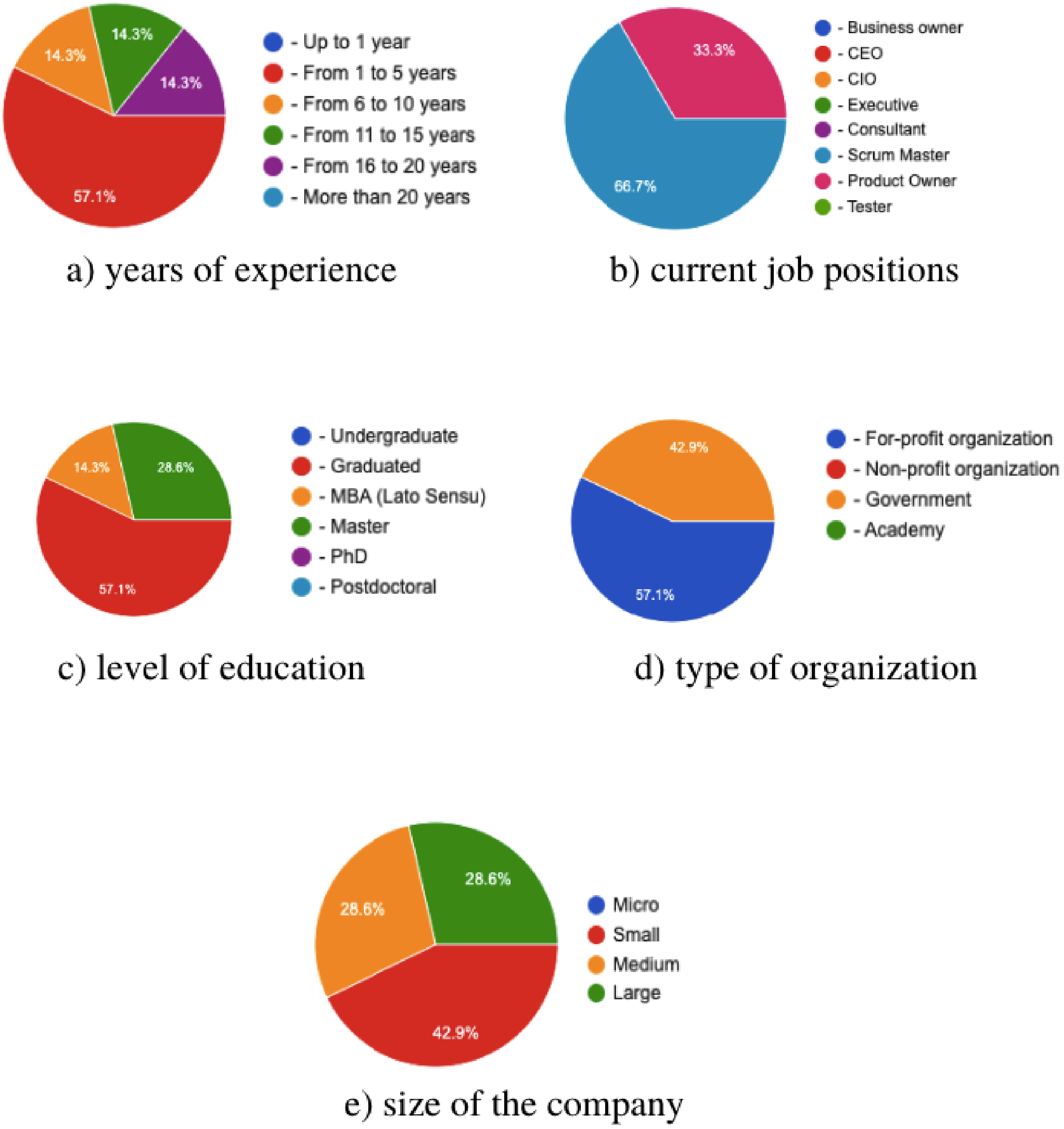


Figure 13 – Participants

(within two other Brazilian public agencies and a Brazilian startup).

6.2.3.3 *Empirical Application of the Euler Model*

For the empirical application of the Euler model, sessions were organized that took place online using the Zoom Video Conferencing Software[3]. This software was chosen because it is free, and it is easy to record the online sessions and therefore to analyze and transcribe the

[3] https://zoom.us/

results.

The sessions began with the researcher's acknowledgment regarding the limited availability of time to participate in the empirical application and evaluation of the model. Next, the concepts of risks and uncertainties are presented, together with examples, to create in the practitioner the knowledge and ability to distinguish between risks and uncertainties.

V1 of the Euler model V1 was presented in detail in chap:actionresearch, so here we will simply present a summary of its processes.

The application of V1 of the Euler model begins by identifying and listing the sources of uncertainty. An Excel spreadsheet in Google Sheets online spreadsheet software was created to support this process [4].

6.2.3.4 *Application of the Euler model at SPS ES*

A summary of the spreadsheet is shown in Table 15. The spreadsheet shows the need to identify uncertainties and sources of uncertainties as a result of the first process of V1 of the Euler model. Each uncertainty listed in the spreadsheet was labeled. U1 means Uncertainty 1, U2 means Uncertainty 2, U3 means Uncertainty 3, and so on). This lean nomenclature will also help in the process for designing and visualizing the Bayesian Network.

Table 15 – Identified uncertainties and mass assignment - MPES

Code	Uncertainty	Expert 1 (m[0,1])	Expert 2 (m[0,1])
U1	Changes in the legal or contractual aspect that governs the project	0.6	0.5
U2	Technology resource shortage - lack of server (hardware)	0.7	0.9
U3	Sudden change of priorities within the project	0.3	0.4
U4	Turnover of team members	0.8	0.5
U5	Need for integration with legacy systems	0.5	0.7
U6	More drastic change of architecture of the current system (product of the project)	0.3	0.5
U7	System overload due to undimensioned accesses (external accesses)	0.4	0.5

[4] https://drive.google.com/drive/folders/1adFSFYmSX-QWG$_X oIVzO1WZdYF4jzRxK?usp = sharing$

Column one (named Code) and Column two (named Uncertainty) of the Table 15 present some of the uncertainties identified by the SPS team in Espírito Santo (SPS ES), a state located in the Northeast of Brazil where there is an agile team to develop the previously mentioned Pandora project. Columns three (labeled Expert 1) and four (labeled Expert 2) indicate the individualized beliefs of each project member. The next step will be calculating a unified belief for the project team.

The second Euler process deals with how the participants (experts) attribute masses. This mass must be between 0 and 1 and represents the degree of evidence related to the occurrence of a particular uncertainty listed in the previous process. At this point, each participant (expert) must attribute a mass based on their evidence (knowledge and experiences) and belief in the occurrence of each uncertainty. Columns three (Experts 1) and four (Expert 2) of Table 15 correspond to the output of process two in the Euler model and is completed by the experts of SPS ES.

After the project team completed the table above, it was asked to assist in the process of building a Bayesian network BN). The motivation, concepts, and step-by-step for creating the network structure were explained.

The process of building a BN is undertaken by using the following steps:

1. Choose a set of relevant variables for the network that describes the domain. In this case, the variables are the uncertainties that we want to relate to the BN;
2. Choose an order for the variables. Based on the team's experience, we identify how uncertainty impacts or relates to other uncertainties;
3. Select a variable and add a node to the network. This is a cyclic process of several iterations;
4. Determine the parent nodes, among the nodes that are already in the network, in this case highlighting the uncertainty that most impacts or most relates to the other uncertainties;
5. Determine the conditional probability table for the network. This step is satisfied by the output of process number 2 of V1 of the Euler model, as the conditional probability table will come from amalgamating the team's evidence by applying the Theory of Evidence.

The result of the manual production of the BN is shown in Figure 14. This BN was built manually with the support of the project team and built in draw.io an online visual software for flowchart maker online diagram[5].

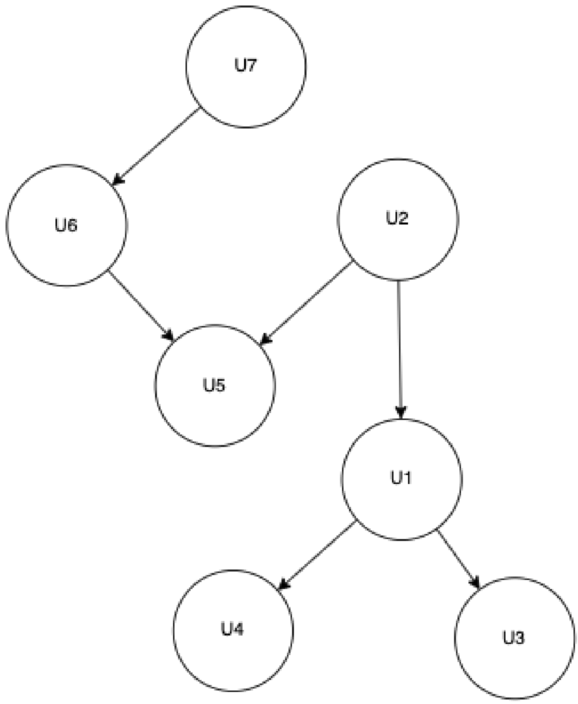


Figure 14 – Bayesian Network - SPS ES

By analyzing each uncertainty seen in Figure 14, we sought to identify relationships between these uncertainties as a way of perceiving the impact that uncertainties have on each other. For example, note from Figure 14 that uncertainty U5 is impacted not only by uncertainty U6 but also by uncertainty U2. This conclusion was reached with the help of the project team. It can thus be seen that the U5 uncertainty deserves special attention from the project team.

As a way of helping the process of applying the Theory of Evidence, a script called Script 1 was created in Python language (Figure 15). The script is available in the code repository created to store the code artifacts of this thesis on Github[6] and APPENDIX D.

The output of the script can be understood as a combination of beliefs and evidence, found individually in Table 15, and which, after applying the Theory of Evidence, represent the beliefs of the project team regarding the occurrence of the uncertainties listed above. This can be seen in Table 16.

5 https://app.diagrams.net/

6 https://github.com/jeffersonjpa/quantifying$_{uncertainty}$

Another analysis that can be done, based on the evidence available to experts, is to calculate the belief in the occurrence of two or more uncertainties combined. For example, the degree of belief in the occurrence of the combined uncertainties U1 and U2 can be obtained by using the formula (6.2).

This can also be done in order to calculate several other combinations of uncertainties. The result would correspond to the experts' degree of belief in the occurrence of the two uncertainties together; likewise for other combinations of occurrences to prioritize combinations of occurrences and not just isolated uncertainties.

Also within the results that can be obtained from applying the Theory of Evidence is plausibility, which can be understood as the risk of the occurrence of analyzed uncertainty. The plausibility function (Pl(A)) can be seen in (6.1). The Pl(A) function presents the degree to which A is plausible (admissible), given the evidence known to experts. Another analysis of plausibility may indicate the degree to which we do not believe in the denial of uncertainty.

$$Pl(U1) = 1 - Bel(\sim U1) \;\; or \;\; Pl(U1) = \sum_{A \cap B \neq \theta} m(B) \tag{6.1}$$

This concept of plausibility is essential for our study because, as presented in Chapter 2, plausibility is associated with the concept of risk. Once the risk is identified, traditional project management processes can be applied to risk management.

$$Bel(U1) = \sum_{U1 \subset U} m(U) \tag{6.2}$$

Table 16 – Result of applying the theory of Evidence

Code	Team Belief (Bel (m[0,1])
U1	0.020
U2	0.043
U3	0.008
U4	0.027
U5	0.024
U6	0.010
U7	0.013

The numbers, one for each uncertainty, representing this joint belief of the project team, were used as input to produce the conditional probability table (CPT) that is input for building the Bayesian network.

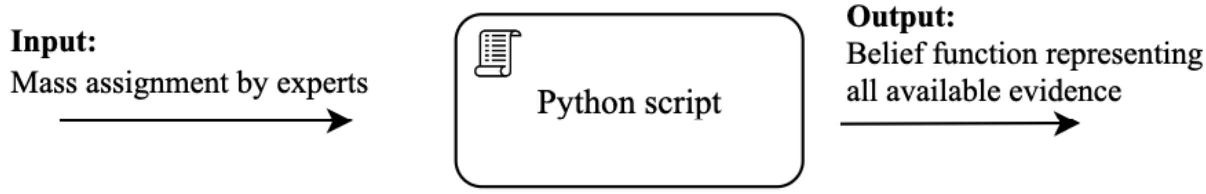


Figure 15 – Script 1 - DST

The next step was to apply the Bayesian network structure (nodes and links), together with the conditional probability table in Script 2 (Figure 16) also available on GitHub[7] and APPENDIX D.

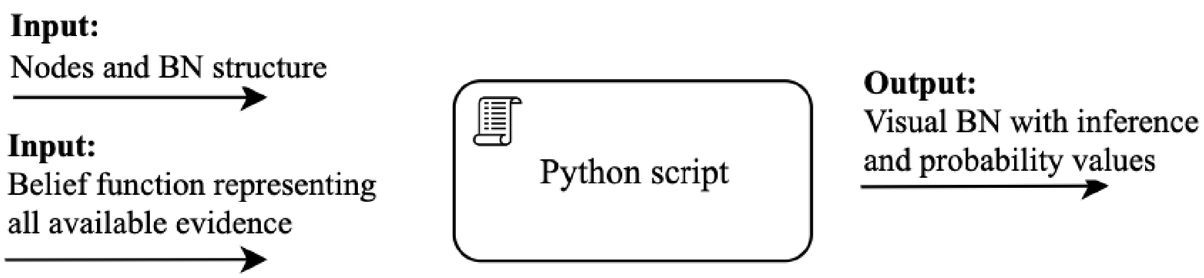


Figure 16 – Script 2 - BN

Figure 17 presents the visual result (now in an automated way) of applying Script 2. The visual analysis of the BN is seen to highlight the uncertainties (nodes) U2, U4, and U5.

The analysis to be performed on the BN starts with the size of the nodes. The node size represents the quantification of the value of belief in uncertainty assessed by the team, meaning that the larger the node size, the greater the team's belief in there being uncertainty. In addition, the use of colors attributed to the nodes is a form of precise visual highlighting of the nodes that need to be observed with higher priority by the risk and uncertainty analysis team. The tone of the color presented a scale that went from redder tones (higher uncertainties) to lighter tones (lower uncertainties).

The results were presented to the SPS ES team and validated in order for it to make sense for the project team to prioritize the highlighted uncertainties.

[7] https://github.com/jeffersonjpa/quantifying$_{uncertainty}$

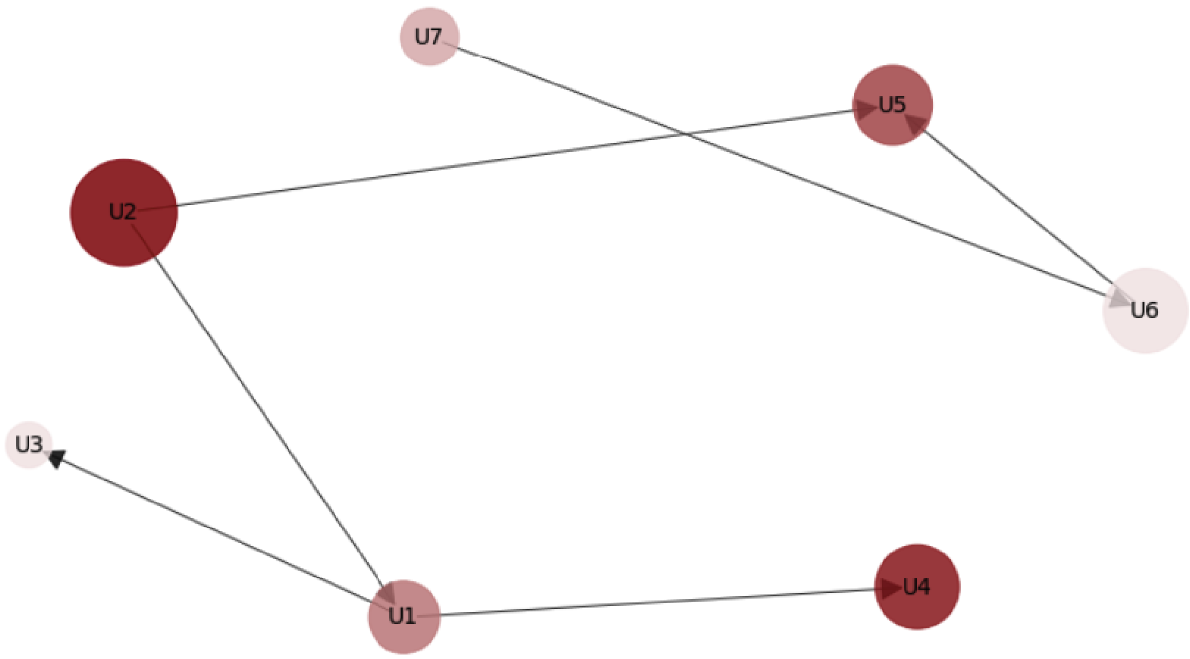


Figure 17 – Bayesian Network Model from SPS ES

The team understood that this analysis could bring a differential to managing uncertainties in software projects. It presents a prioritization relationship between uncertainties and relationships between uncertainties to highlight which uncertainties impact each other.

Thereafter, the empirical sessions of applying the V1 Euler model followed with the participation of the other two agile teams.

6.2.3.5 *Application of the Euler model at Fabwork*

One of the teams belongs to the Information Technology sector of a private company based in Paraíba, and another agile software development team is from the Municipality of João Pessoa - PB. The worksheets with the results of applying Euler V1 in these two other teams can be found in the cloud repository created for this thesis[8]. A summary is given in Table 17

[8] https://bit.ly/QURepository

Table 17 – Identified uncertainties and mass assignment - FABWORK

Code	Uncertainty	Expert 1 (m[0,1])	Expert 2 (m[0,1])	Expert 3 (m[0,1])
U1	Lack of expertise in the technology/programming language used in the project.	0.5	0.55	0.5
U2	Lack of formal process for managing uncertainty.	0.75	0.8	0.75
U3	Adaptation of the model based on the type of projects being executed (a mix of practices)	0.25	0.3	0.5
U4	Turn-over of project team members	0.9	0.7	0.35
U5	Lack of expertise in the business	0.5	0.5	0.5
U6	Lack of financial resources despite the budget forecast by client	0.25	0.25	0.25
U7	Lack of knowledge of the development process / customer interaction	0.75	0.7	0.85

Another summary is given in Table 18

!

Table 18 – Result of applying the Theory of Evidence

Code	Team Belief (Bel (m[0,1])
U1	0.003
U2	0.002
U3	0.000
U4	0.004
U5	0.002
U6	0.000
U7	0.010

Figure 18 presents the visual result (now in an automated way) of applying Script 2. The visual analysis of the BN is seen to highlight the uncertainties (nodes) U7 and U4.

6.2.3.6 Application of the Euler model in the city hall of João Pessoa - PB

A summary is given in Table 19

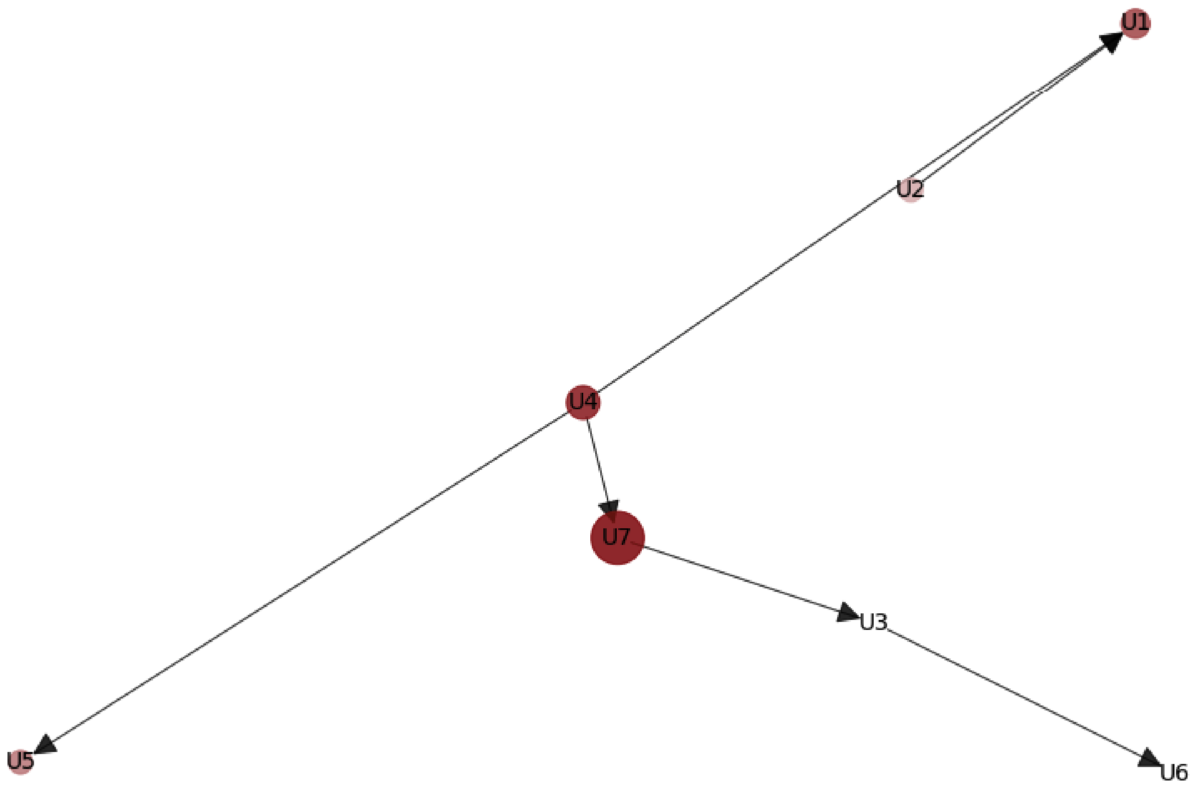


Figure 18 – Bayesian Network Model from FABWORK

Table 19 – Identified uncertainties and mass assignment - City Hall Joa Pessoa

Code	Uncertainty	Expert 1 (m[0,1])	Expert 2 (m[0,1])
U1	Team member turnover	0.8	0.75
U2	Public management changes	0.5	0.5
U3	Error in the integration between APIs	0.4	0.7
U4	Error in requirements elicitation	0.3	0.2
U5	Exploitation of vulnerabilities in the libraries used	0.6	0.6

Another summary is given in Table 20

Table 20 – Result of applying the theory of Evidence - city hall Joao Pessoa

Code	Team Belief (Bel (m[0,1])
U1	0.083
U2	0.034
U3	0.039
U4	0.008
U5	0.050

Figure 19 presents the visual result (now in an automated way) of applying Script 2. The visual analysis of the BN highlights the uncertainties (nodes) U1 and U5.

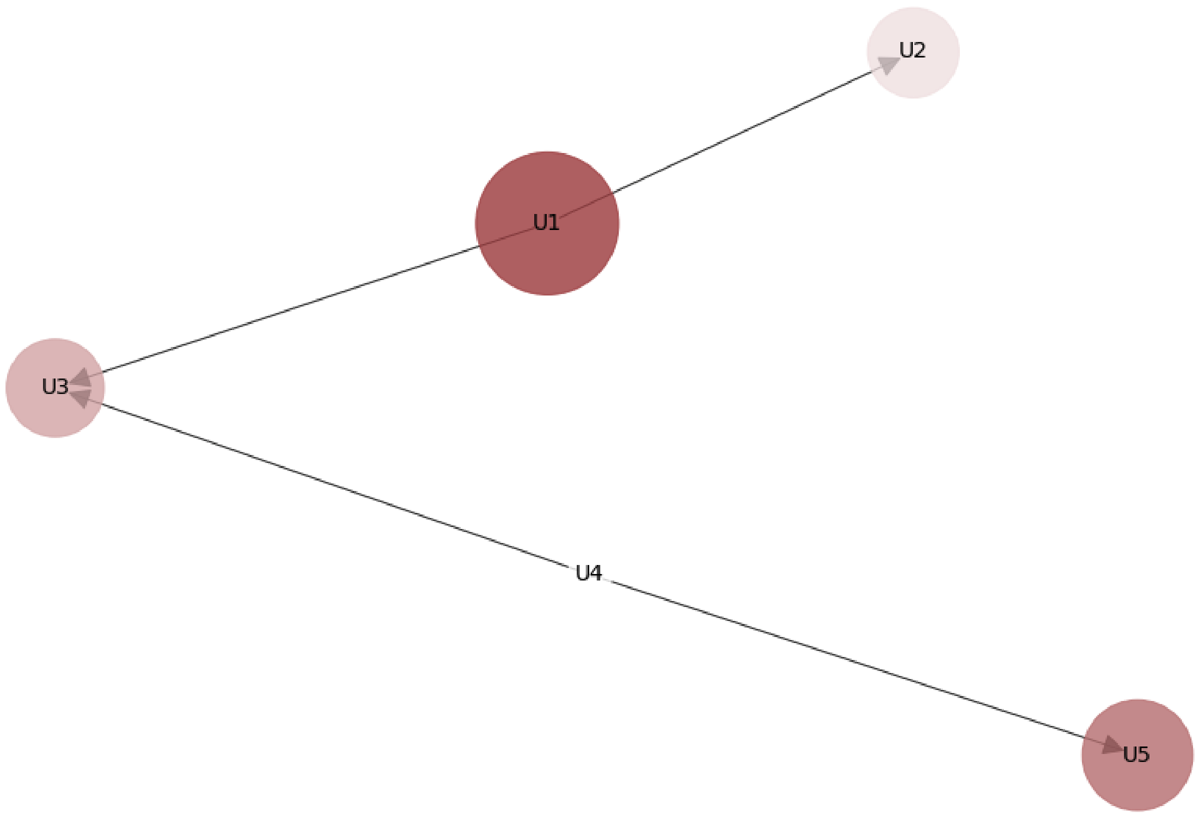


Figure 19 – Bayesian Network Model from city Hall of Joao Pessoa

Finally, the fourth step, post facto demonstrations, we will treat as the moment to discuss the benefits and limitations of applying the model as a way of evaluating the model and thus producing an improved version that we will call Euler version 2.0. This will be presented in the next section.

6.2.4 Step 4. Post Facto Demonstrations

This section introduces the PoCR framework step called post facto demonstrations. This is the moment to discuss the benefits and limitations of the empirical application of the model and thus to produce recommendations for an improved version of the Euler model.

This section starts with the data collection process, and this is followed by analyzing the data collected after transcribing the participants' interviews on the empirical application of the model. It ends with the presentation of the recommendations for adjustments and improvements that emerged from this process.

6.2.4.1 Data Collection

After three agile teams had each carried out a practical application of the processes of the model, it was necessary to evaluate it by applying a questionnaire. This corresponded to the post facto arguments stage of the PoCR. The questions asked were based on Runeson e Höst (2009) and questions found in the protocol developed by Marinho et al. (2015b) to evaluate their approach to the management of uncertainties.

In this step, the questionnaire[9] and semi-structured interviews were applied, and the results from the questionnaire will be used to evaluate the model. An improved version called Euler 2.0 will be built based on the results of this evaluation. The output of this last PoCR process is a set of suggestions and points for improving the Euler model and proving the viability of the solution developed.

The qualitative data will be analyzed after completing an analysis of answers to the open-ended questions. We have applied the phases of the thematic analysis process robson2016real to answers of the open questions. We used a coding guide in conjunction with the Atlas.it tool[10]. This coding guide can be used as a form of alignment between several researchers who aim to analyze the same data set within this research. Involving more than one researcher in analyzing the material collected can minimize threats to the validity of the research, such as inconsistencies in the coding process. However, we emphasize that this analysis was carried out by a single researcher and that this protocol was evaluated by three specialists, each with a Ph.D. in Software Engineering and extensive experience in the area.

[9] https://bit.ly/3t5lWOO

[10] https://atlasti.com/pt

6.2.4.2 Data Analysis

In this section, we describe the data analysis process during which transcripts should be coded using the phases of thematic coding analysis robson2016real. These codes were built with a focus on possible answers to research questions. The coding scheme adopted is: C1 for 'Problems and challenges', meaning an excerpt from the transcript that identifies existing problems in the model that may make its application difficult or impossible in real environments of software projects with agile practices; C2 for 'Suggestions for improvement', meaning an excerpt from the transcript that identifies points for improvement in the Euler model, e.g., suggestions for changing the name or adjustments in parts of the process of applying the Euler model; and C3 for 'New processes or the replacement of existing processes', meaning an excerpt from the transcript that identifies the need to apply new processes or replace existing processes in the model.

Table 21 – Codes and Themes

Id	Code	Description
C1	P — Problems, challenges	It encodes an excerpt from a transcript that identifies existing problems in the model that may make it difficult or impossible to apply it in real environments of software projects with agile practices.
C2	I — Suggestion for improvement	It encodes an excerpt from aa transcript that identifies points for improvement in the Euler model, such as a suggestion for name changes or adjustments to parts of the process of applying the model.
C3	N — New processes or replacing an existing one	It encodes an excerpt from the transcription that identifies the need to apply new processes and replace existing processes in the model.

The first column of Table 21 presents a code identifier of the theme. This identifier will be used when coding the quotes of each participant in the process for evaluating the open questions on the evaluation form of the Euler model. The second column presents the code created to standardize the analysis performed by more than one researcher. Finally, the third column presents the description of the code in order to make its use clear during the qualitative analysis of the answers of the open questions.

6.2.4.3 Results of the Interviews

This section presents the results collected from the empirical application of PoCR. Here we also start a discussion regarding the results found after applying the questionnaire [11].

The scale used in the type of statement seen in Figure 21 and similarly used throughout the questionnaire ranges from Strongly Disagree (1) to Strongly Agree (5).

As to the assertion *The organization where I work (or I worked) is interested in adopting an approach to managing uncertainty. You as project manager would be interested in adopting the approach.*

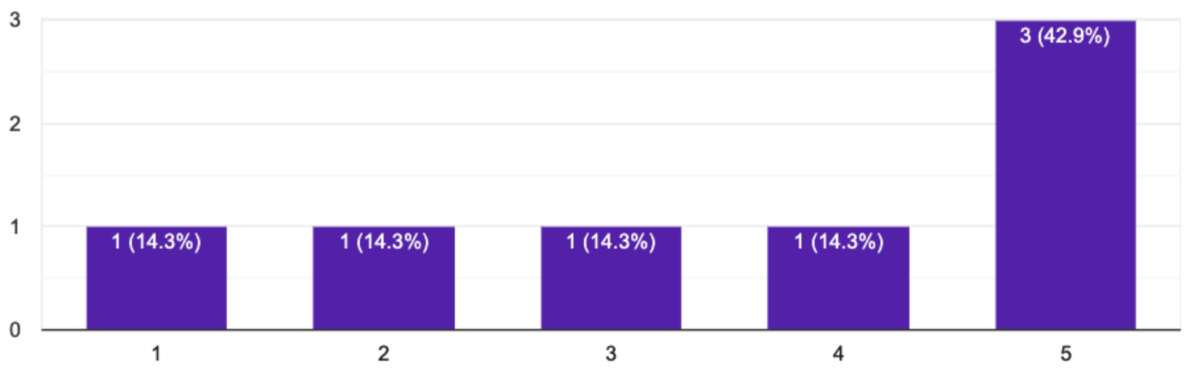


Figure 20 – The organization where I work

It can be seen from the results in Figure 20 that 42.9% of the participants strongly agree about the interest of organizations in adopting an approach with uncertainty management practices in agile software development projects. Thus, it is concluded that there is interest in managing the uncertainties of their projects on the part of organizations and those who adopt agile practices in project management.

As to Figure 21 that shows the answer to the assertion *You understand that using a model to quantify epistemic uncertainties in software projects can bring benefits such as greater predictability of uncertainties associated with the various sources of uncertainty inherent in projects that use agile project management practices and that this can lead to better project coordination.*

It can be seen that 57.1.9% of the participants strongly agree, and 42.9% agree that using a model to quantify epistemic uncertainty can bring benefits such as greater predictability and

[11] https://bit.ly/3t5lWOO

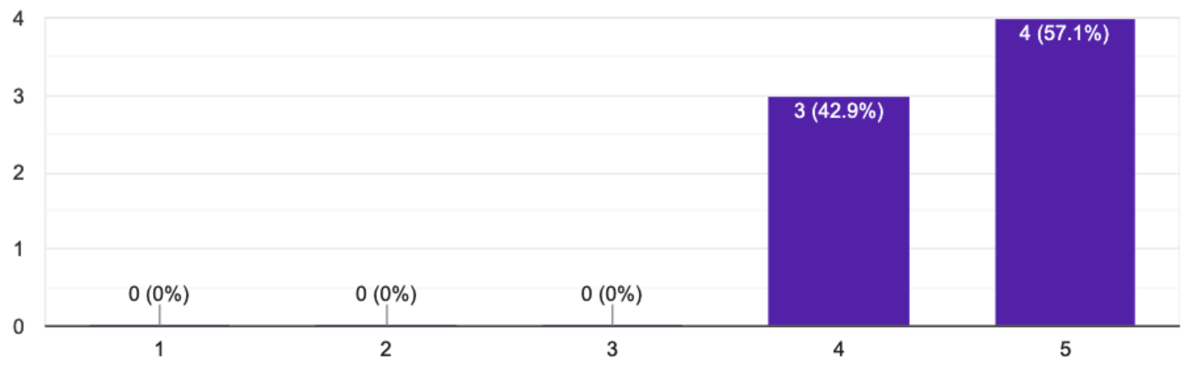


Figure 21 – Using a model to quantify epistemic uncertainties

that this practice can result in the better coordination of projects.

Thus, it is concluded that there are noticeable benefits in the agile project management process with the addition of uncertainty quantification practices. When quantifying, the project manager can prioritize, and make the project more predictable, which results in more security regarding the predictability of the project tasks and the identification of risks, and consequently, better organization of the project.

As to the assertion of the statement *The nomenclature of the phases of the Euler model is straightforward and follows a logical sequence.*

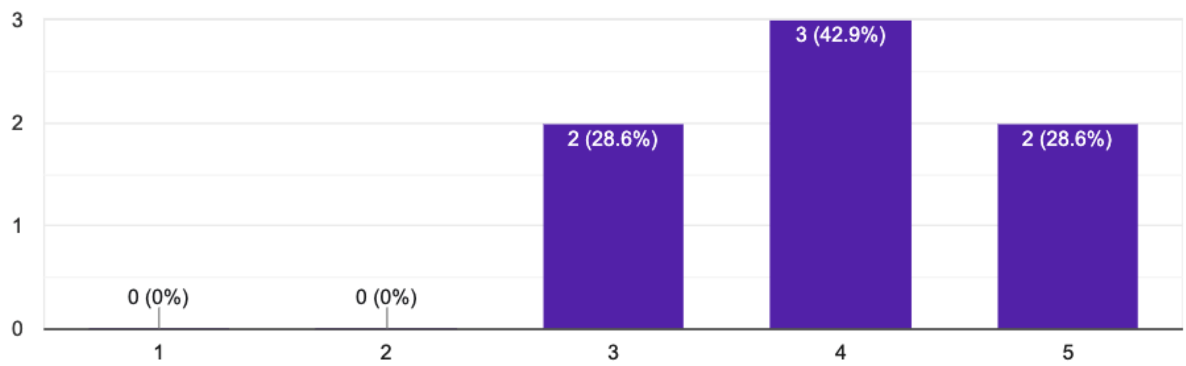


Figure 22 – The nomenclature of the phases of the Euler

It can be seen in Figure 22 that there is a variation in how clear the nomenclature of the phases of the Euler model is. Answers to this question show 28.6% strongly agree that the nomenclature is clear, 42.9% agree that the nomenclature is clear, and 28.6% were neutral regarding the clarity in the nomenclature used by the model.

It is evident there is a need to adjust the existing names in the model to improve the nomenclature and to better communicate the objectives of each phase of the model. All closed questions were supported by open questions so that the participants could develop their reasoning regarding their answers to the question. To exemplify, adjustments were suggested regarding the name of the second phase, 'Evidence Theory (Dempster-Shafer'), which had been, in the first version of the Euler model, the name of the model and its authors (Dempster-Shafer).

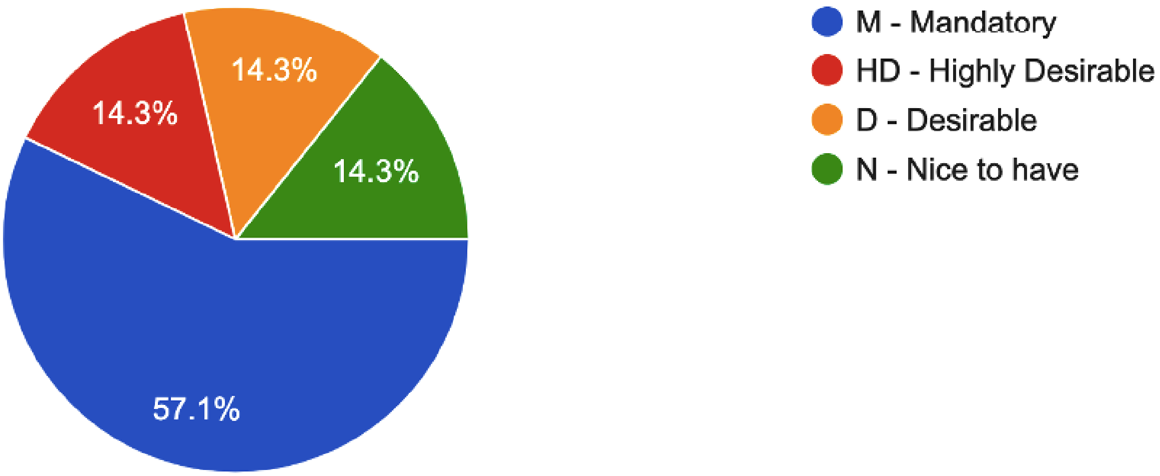


Figure 23 – On the process of Applying Bayesian Networks

About Figure 23 that shows the answer to the question about *Applying the Bayesian Networks process and its importance for modeling the quantification of epistemic uncertainty*, it can be seen from the results that 57.1% of the participants understand that this process is mandatory for the Euler model. This result makes perfect sense if we think that one of the characteristics of the Euler model is to relate sources of uncertainty and quantify them. According to a search carried out on the literature, one of the approaches used to quantify uncertainty and the interrelationship between sources of uncertainty can be Bayesian Networks. Thus, the conclusion is that this Euler model process has evident importance for PoCR participants and is also mandatory from their point of view.

Regarding Figure 24, which presents the result of the statement that the outputs of each phase of the Euler model are clear, as well as how each output feeds the input of the following process, it can be seen that 71.4% of the participants from PoCR strongly agree with this statement. From this answer, it can be concluded that the process is coherent from the point of view of the connection between its processes.

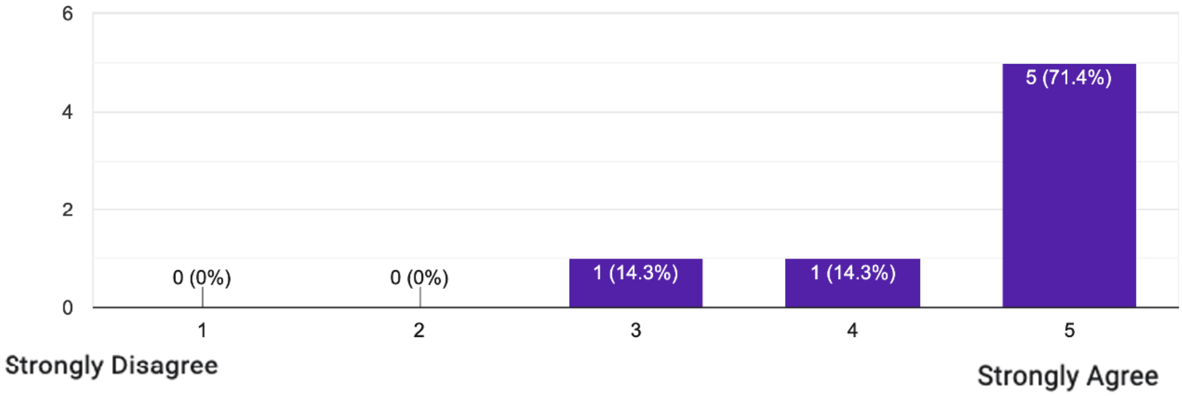


Figure 24 – The output of each phase of the Euler model

Following the thematic analysis process, there is a quote from Participant P1 *"I believe that it will take time for this concept to mature and, hence to carry out the integration with risk management. This will not be an easy process."* It was coded as C3.

A quote from Participant P4 is *"It is necessary to understand what impacts the identification of uncertainty will have throughout the other project stages, such as cost management and time management."* It was coded as C3.

A quote from Participant P2 is *"I understand that the name 'Theory of Evidence' could have a clearer name. Example: Validation of Uncertainties or something along those lines."* It was coded as C2.

A quote from Participant P3 is *"The name of the second phase causes me some difficulty."* It was coded as C3.

A quote from Participant P3 is *"As it was a study of low probability possibilities or the lack of knowledge about the possible uncertainties, they may discourage the implementation of the model. I believe that a great difficulty would be to change the team's mentality to highlight the importance of preventing uncertainty."* It was coded as C1.

6.2.5 Impact of COVID-19 on the application of the model

At the end of 2019, the first signs appeared of a severe acute respiratory syndrome that would soon spread worldwide. In line with the climate of uncertainty, it was hoped that a vaccine would emerge that could protect people and prevent them from suffering from a severe form of the disease.

This syndrome became known as SARS-CoV-2 or Covid-19. Like other extreme crises, such as wars, the Covid-19 epidemic was also a catalyst for numerous changes that were about to happen. Changes in the educational system due to adopting virtual classes, changes in the monetary system due to virtualizing currencies, and changes in how people relate and work due to social distancing. This social distancing was one of the main impacts of Covid-19 that was felt by companies, especially startups da2020agile.

This thesis was developed during the most critical period of the COVID-19 Pandemic, namely the year 2020.

The impact of COVID-19 on this thesis is demonstrated by noting that some uncertainties mention the absence of co-workers due to the disease - "U11 – Partial or total absence of team members because of Covid-19" and "U3 - Insufficient team to meet the growing demands of the project" described in Chapter 5. These are uncertainties that can happen in normal times but the scale of these was magnified during the height of the pandemic.

The Covid-19 pandemic impacted the application and evaluation of the Euler model in aspects of the need to hold online meetings with agile project team members because of the need for social distancing.

Even though needing to comply with the rules of social distancing may have reduced the level of interaction that can be achieved by holding face-to-face meetings, benefits emerged, such as reduced costs as few people traveled to meetings, access to more participants, and research could be carried out in less time by making greater use of virtual tools. Access to geographically remote populations is also considered a significant benefit of this type of virtual interaction cooper2003report.

This advantage of the geographic reach of participants proved to be very useful to this research since two of the evaluation team', despite being members of the project teams of the Joao Pessoa city hall in Paraíba (northeast of Brazil), conducted the evaluations in states in the North and South regions of the country, by virtual means.

An attempt was made to minimize any adverse impacts of online interaction on this thesis by adopting measures such as the availability of a range of schedules by the leading researcher made a range of schedules of his activities available to the project teams in order not to impact the schedules and agenda of the project team. The Zoom online Meeting tool[12] was used, which allows free recording of meetings for later analysis, thus making discussions during meetings more fluid between researcher and project teams. Finally, the meetings were broken

[12] https://zoom.us/

into three shorter sessions in order to make the meetings less tiring and not to overload the teams' agenda.

6.2.6 Euler V2 Recommendations

The first column of Table 22 presents an identifier code of the participant in the process for evaluating the Euler model The second column presents the excerpt quoted from the open questions in the Euler model evaluation questionnaire. Finally, the third column presents the encoding used to classify the quote shown in the previous column.

Table 23 presents a coding as the first column to facilitate the traceability of the analysis process and identification of the recommendations made by the PoCR participants.

For example, in the first line of the column, 'P1 – Q1 – C3', its result should be interpreted as that Participant P1 generated the recommendation as a result of his Quote Q1, which was marked with the code C3 - see Table 21 for the explanation of the meaning of each code.

The same reasoning can be applied to the third line, 'P2 – Q1 – C2', where Participant P2, as a result of his Quote Q1, was marked with the code C2, resulting in the recommendation in the equivalent column.

6.2.6.1 Leveling and awareness

This section presents the structural changes made to the model, such as the addition of new processes and adjustments to process names to make the terms used in the model more meaningful to practitioners and researchers of uncertainty management.

As a result of the empirical evaluation made by applying the Elliott framework (ELLIOTT, 2021) from PoCR in the first version of the Euler model, we have version 2.0, shown in the next Chapter 7.

A total of eight recommendations were collected and implemented, which can be seen in Table 23.

The recommendation we call R1 deals with the adjustments made to the Euler model that suggests a way of clarifying the integration with the risk management process. Thus, before the 'Feeding the Body of Knowledge process, the Risk Management Integration process was added to collect the uncertainties arising on applying the Euler model. It is understood that this adjustment makes the model more predictable and more visible than the current practices

Table 22 – Participant mapping, quotes and theme

Participant ID	Quote	Theme
P1	"I believe that it will take time for this concept to mature and, hence, to carry out the integration with risk management. This will not be an easy process." – Q1	C3
P4	"It is necessary to understand what impacts the identification of uncertainty will have throughout the other project stages, such as cost management and time management." – Q1	C3
P2	"I understand that the name 'Theory of Evidence' could have a clearer name. Example: Validation of Uncertainties or something along those lines." – Q1	C2
P6	"It is necessary to demonstrate the applicability of the model in different markets, for small or large companies, cash or customized software, and whether or not it has well-defined processes. " – Q1	C2
P3	"As it was a study of low probability possibilities or the lack of knowledge about the possible uncertainties, they may discourage the implementation of the model. I believe that a great difficulty would be to change the team's mentality to highlight the importance of preventing uncertainty." - Q1	C1
P5	"I believe that it will take time for this concept to mature and, hence, to carry out the integration with risk management. This will not be an easy process." - Q1	C3
P3	"The name of the second phase causes me some difficulty." - Q2	C3
P7	"Perhaps the term Bayesian Networks is not completely clear to everyone, making it difficult to understand." - Q1	C3

of managing uncertainties known to practitioners and researchers.

It is understood that this step closes a complete cycle of the uncertainty management process in agile software development projects, thereby contributing to the project's success.

Regarding recommendation R2, we understand that this is also reflected at the moment when we made the integration of uncertainty management processes with risk management processes more explicit. In the same way that a new risk represents inputs in other phases of the project, such as costs, time and human resources (to exemplify a few), a new uncertainty must also generate inputs in these same processes. As the management of uncertainties also

Table 23 – Traceability and Recommendations

Code	Traceability	Euler 2.0 Recommendations
R1	P1 – Q1 – C3	Make clear the integration with the risk management process.
R2	P4 – Q1 – C3	It is necessary to demonstrate how managing uncertainty impacts other stages of project management.
R3	P2 – Q1 – C2 P3 – Q2 – C3	Search for a more meaningful name for the process associated with Theory of Evidence.
R4	P6 – Q1 – C2	It is necessary to demonstrate which adaptations are necessary in the model so that it can be applied in other types of companies, projects or software development processes.
R4	P3 – Q1 – C1	It is thus understood that this recommendation is reflected in the creation of the leveling process and awareness of the importance of this subject for the company, thereby constantly demonstrating its benefits. However, we know that this is not an easy process and that any change in mindset requires time.
R6	P5 – Q1 – C3	This recommendation is another reinforcement that it is necessary, during the leveling process, to emphasize the existing benefits in managing uncertainties and the form of integration with risk management. It should be made clear that the process for managing uncertainty is something more significant than the risk management process. This process brings greater security and predictability to project management as a whole.
R7	P3 – Q2 – C3	This recommendation aims to clarify the name of the phase related to the process of collecting expert opinions and building a single opinion associated with the project team. As a consideration, the name of this stage has been changed to something more meaningful, namely 'Expert Analysis of Uncertainties'.
R7	P7 – Q1 – C3	As to this recommendation, the name of the phase 'Applying Bayesian Networks' has been adjusted to 'Analysis of interdependence between uncertainties,' thus making the purpose of this process clearer.

generates costs, it also lacks a timeline record by which we can handle it and it also needs an owner with responsibility for dealing with the uncertainties raised during the uncertainty management process.

6.2.6.2 Risk Management Integration

Recommendation R3 was reflected in the name change of the Evidence Theory application process. The research result presented in the Results section demonstrates the importance of this phase, so the question here was not to consider its importance. The point was to give the process a more meaningful name so that even someone unfamiliar with the Theory of Evidence could understand what is expected of this process. It is understood that this change made the model more flexible, as it is known that there are other ways of quantifying uncertainty that can be used to replace the application of the Dempster-Shafer Theory of Evidence.

Following the PoCR framework, the Euler model was applied to three agile teams from three different companies, two from private companies and one from a public company. All three teams apply agile practices to software development projects. Therefore, and in response to the R4 recommendation, the authors understand that it is not yet possible to suggest adaptations to the model outside the scope described above. So that the suggestions do not come only from a theoretical field, the model must be applied to different companies, such as mixed capital companies and types of projects with different development processes that use, for example, traditional methods of managing projects. On the other hand, the purpose of this work would not be to present an empirical process of evaluating a model. Therefore, this recommendation highlights this, but the authors understand that there is no structural modification to meet this recommendation.

In response to recommendation R5, it is understood that greater emphasis should be given to the Leveling and Awareness process. This emphasis must be made so that the difference between the concepts of risks and uncertainties becomes more explicit, but mainly the benefits of applying uncertainty management processes, including quantifying epistemic uncertainty in agile projects.

Recommendation R6 reflects the same concerns as Recommendations R1 and R2 but with more emphasis on the perception that the risk management process is an integral part of the uncertainty management process. This recommendation is reflected in the structure of the model through the clear visualization of the risk management process as part of the uncertainty management model.

In recommendation R7, the same thought applied to recommendation R3 applies here, namely, that the issue discussed is not the importance of the process of building a Bayesian network. The question is the most significant naming choice. Thus, there was a change in the nomenclature of this process in order to reflect its purpose better. This change also resulted in the advantage of making the model more flexible with the possibility of adopting other ways of checking the interdependence between sources of uncertainty in agile projects.

It is essential to highlight that the Euler model is a cyclical process and must be carried out constantly, and is concluded only at the end of the project.

6.3 CLOSING REMARKS

This chapter presented the results of evaluating the first version of Euler's model through an empirical evaluation method. As a result, gaps and suggestions for improvement were collected to encourage the construction of a second version of the model. The next chapter will present this second version of the Euler model.

7 EULER - A MODEL FOR DEALING WITH EPISTEMIC UNCERTAINTIES IN AGILE SOFTWARE PROJECT MANAGEMENT

The previous chapter tried to present the evaluation results of the first version of the Euler model. The result of this evaluation carried out empirically in three teams that work with agile software development practices, served as a subsidy for the production of the second version of the model that will be presented in this chapter. Among the appendices of this chapter is a guideline that aims to help the process of implementing the model.

This chapter seeks to help organizations that practice agile and uncertainty management in software projects. A model for quantifying epistemic uncertainty in software projects is presented. This is done by specifying their processes and highlighting the inputs and outputs of each process.

7.1 INTRODUCTION

Making decisions about what to expect in the future is of paramount importance for project managers. A project plan is nothing more than a realization of the desire to anticipate or control uncertainties in projects. The size of this challenging situation becomes even clearer when one realizes that the further into the future that decisions made today assume what the future will bring, the more that the decision-making process will contain uncertainties. Making decisions about what to expect in the future is of paramount importance for project managers. A project plan is nothing more than a realization of the desire to anticipate or control uncertainties in projects.

This study was conducted based on various research methods which were applied to the context of a problem and the research question which was defined in Chapter 1 as *How do project managers deal with epistemic uncertainty in a quantitative perspective in agile software project management?*

The model built and presented in this chapter combines the application of Evidence-Based Software Engineering methods described in Chapter 4, Action Research presented in Chapter 5, empirical application, and evaluation of the first version of the Euler model called V1 presented in Chapter 6.

A guideline was developed to assist in applying the model with agile teams. This guideline can be found in APPENDIX A.

7.2 OVERVIEW OF EULER ARCHITECTURE

According to pidd1997tools, a model can be defined as an external and explicit representation of a part of reality or a phenomenon of interest to people who can use this model to understand, change or manage and control this part of the reality studied.

According to kuhne2004, in Software Engineering, a model can be traditionally represented by an artifact formulated in a modeling language such as Unified Modeling Language (UML) or Business Process Model and Notation (BPMN)[1]. These models can describe a system by using various types of diagrams. Generally speaking, these models are based on diagrams and are typically assembled visually.

Since processes are built into the internal construction of a model, it can focus on the part of reality that needs to be understood and managed. For the author of this thesis, this part of reality can be understood as the quantification of epistemic uncertainty in software projects, and so constructing such a model that also uses diagrams to manage uncertainty in projects is justified.

The Action Research cycles, set out in barbosa2021emdirecao, produced a model and the knowledge acquired about quantifying epistemic uncertainty in software projects. This model was called Euler. The name of the model in English is **E**pistemic **U**ncertainty Mode**L** for Softwar**E** P**R**ojects. The acronym of the model, EULER, is in honor of the mathematician Leonhard Paul Euler who lived in the 18th century and made discoveries in various fundamental areas of mathematics, such as calculus and graph theory.

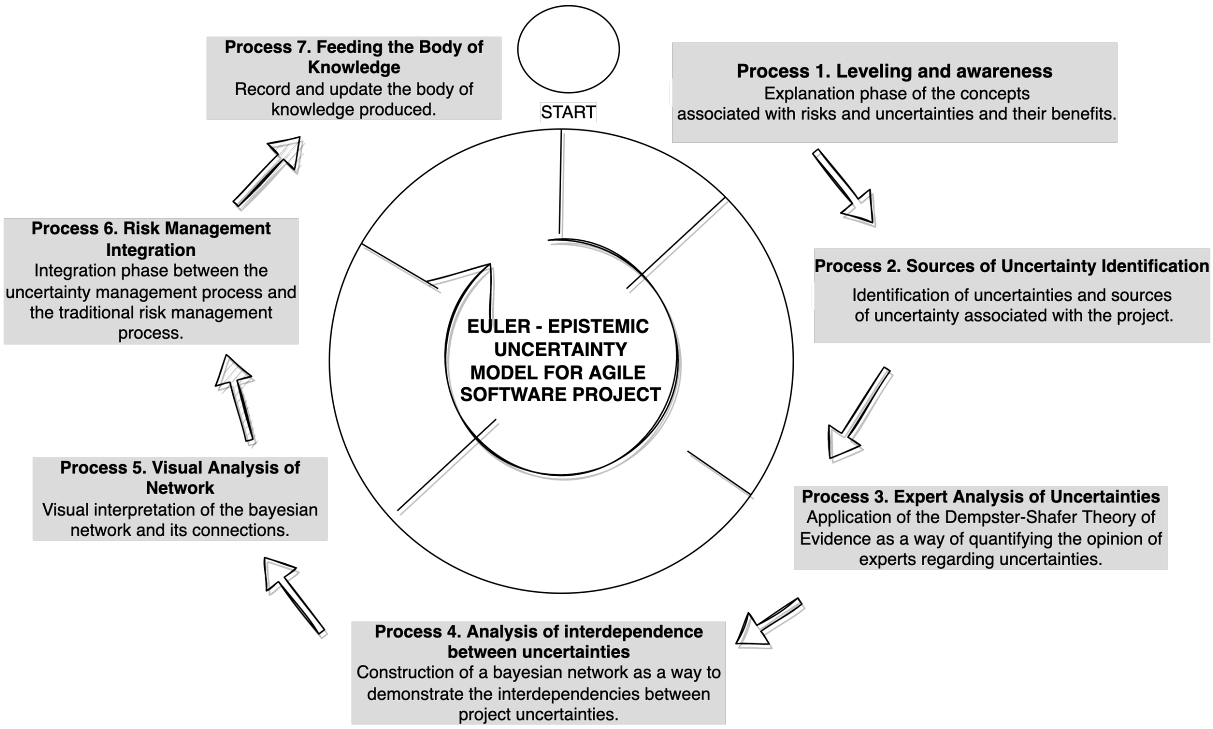
The Euler model comprises seven subprocesses, 21 tasks, three artifacts, and three roles. Figure 25 gives an overview of the architecture of this model.

This view highlights the cyclical character of the model and the sequential execution of the processes belonging to the model.

The subsections that follow will present each part of the model in detail.

[1] https://www.bpmn.org/

Figure 25 – Architecture Overview

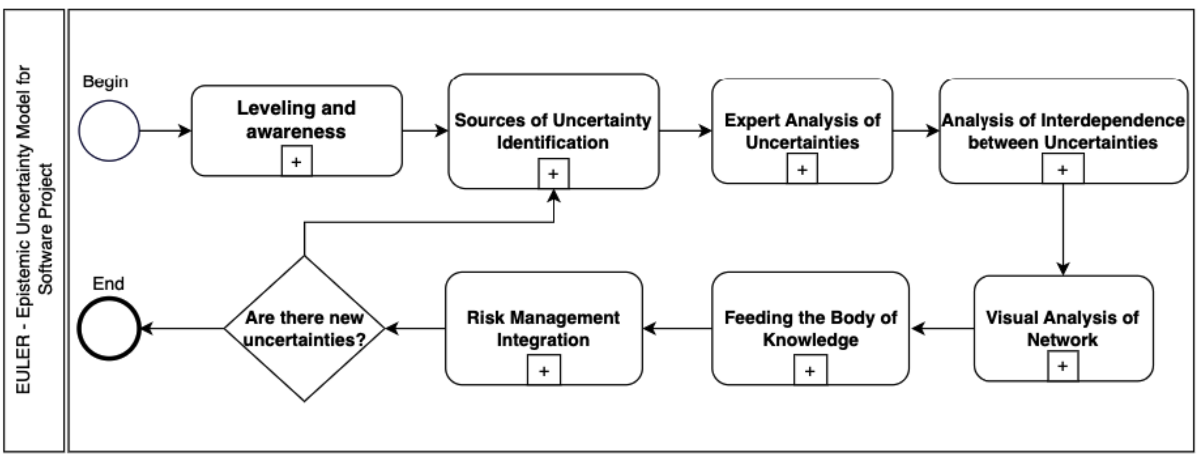


Source: Author

7.3 EULER INTERNAL PROCESSES

The Euler process model was written using the notation language BPMN (Business Process Model and Notation). As can be seen, there are sub-processes within each process for the complete execution of the model's activities and activities associated with a decision process. The need to return to the processes at the beginning of the cycles is verified to identify new uncertainties or proceed to the end of its cycle to close the process of quantifying uncertainty.

Figure 26 – Overview of the Euler process



Source: Author

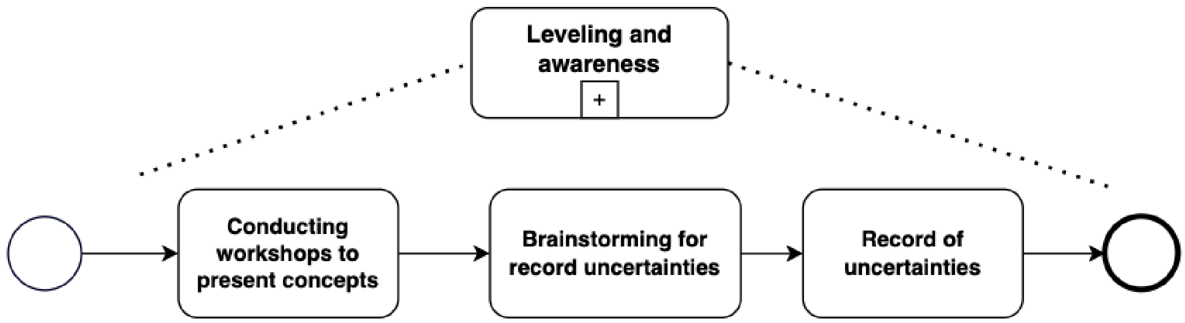
7.3.1 Leveling and Awareness

Regarding its first version, V1, the Euler model was changed so as to include a new process called Leveling and Awareness. The objective of this new process is to level the practitioners of the process of managing uncertainty regarding the concepts associated with risks and uncertainties in software projects.

This objective can be achieved by taking practical examples of risks and uncertainties from the body of knowledge and the literature to reduce team members' doubts before they seek to identify project uncertainties and sources of uncertainty.

This leveling step proved necessary when taking advantage of the three opportunities for the empirical application of the Euler model within the PoCR (Proof of Concept Research) empirical evaluation process.

Figure 27 – Leveling and Awareness Process



Source: Author

The activities of this process include:

1. Conducting workshops to present concepts on risks and uncertainties, the presentation of examples being a way to illustrate and reinforce the concepts before starting the process of identifying uncertainties;

2. Next, with the concept of uncertainty, recall and record the uncertainties that occurred in previous projects and how the teams dealt with them. This activity aims to identify and record lessons learned from dealing with uncertainties in previous projects.

3. Register the lessons learned in an appropriate repository within the project's infrastructure to produce a body of knowledge that can be consulted for future projects.

We did not identify any supporting artifact for this process that could be highlighted because a simple Excel spreadsheet with the formatting of rows and columns can be used to record the uncertainties listed. Likewise, if members of the project team understand that they need something more dynamic during the workshop, they can consider using, e.g., Canvas, Flip charts, or another appropriate tool.

The output of this process should be the leveling of all participants in the process for managing uncertainty concerning the concepts of risks and uncertainties.

7.3.2 Identification of the Source of Uncertainty

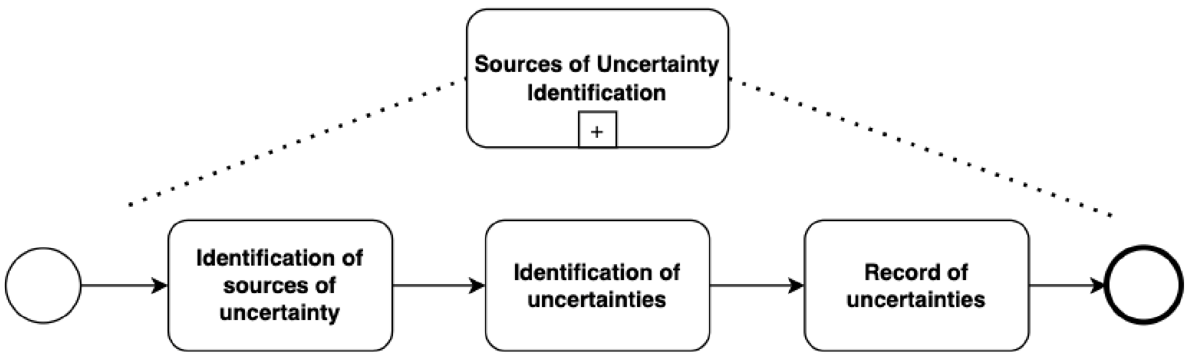
This second step of the model deals with identifying uncertainties in the project. A search was made in the literature for sources of uncertainty to assist in identifying uncertainties. These sources can be understood as a driver or guide or an initial step in identifying project uncertainties.

To help the process, sources of uncertainty that could support identifying uncertainties found in article$_{m}arinho_{s}ampaiowereused.$

Sources of uncertainty include Technological, Market, Environmental, Socio-Human, and Organizational matters. The sources of uncertainty can be used to guide the members of the project team on how to search for and identify uncertainties.

Figure 28 presents the Process for Identifying Sources of Uncertainty.

Figure 28 – Process for Identifying Sources of Uncertainty



Source: Author

After identifying sources of uncertainty, one proceeds to the process of identifying the registration of project uncertainties. This record of uncertainties can be made on an Excel spreadsheet or on the institution's project management system.

The output of this process must be a set of identified and duly registered uncertainties so that the entire project team is aware of the existing uncertainties.

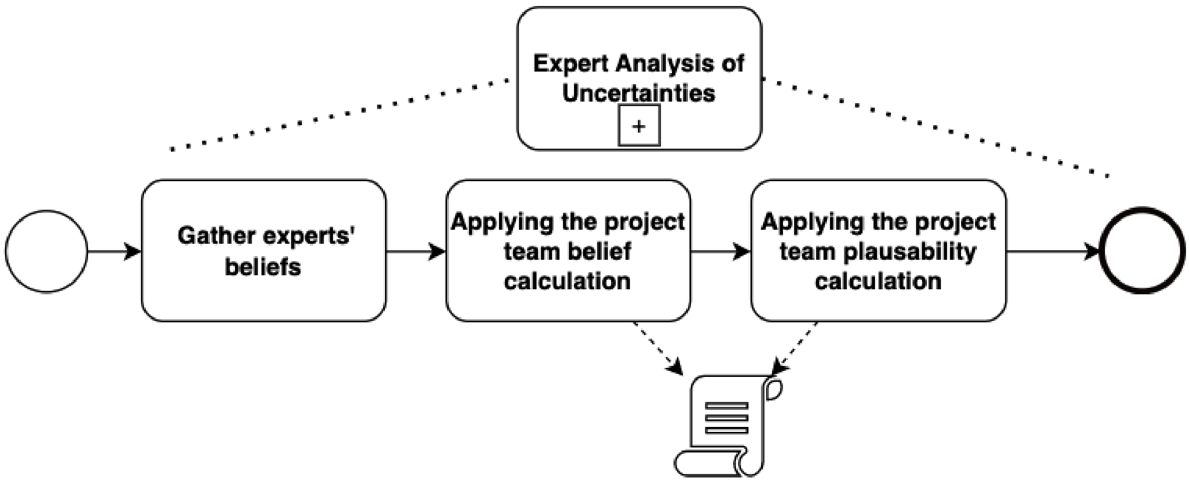
7.3.3 Experts' Analysis of Uncertainties

The Dempster-Shafer Theory of Evidence (DST) allows the treatment of uncertainty based on the knowledge of experts (agile team), gained from their experiences, knowledge, and preliminary information about the activities of the design, assign mass to the elements(uncertainties) in the range [0, 1].

In this step, the DST requires the collection of the expert's opinions in order to quantify the belief that each of them has on whether or not the uncertainties identified in the previous step occur.

Figure 29 presents the process by which the Experts' Analysis of Uncertainties is made

Figure 29 – The Process for Experts' Analysis of Uncertainties



Source: Author

According to santosDempster2018, this process had the following activities:

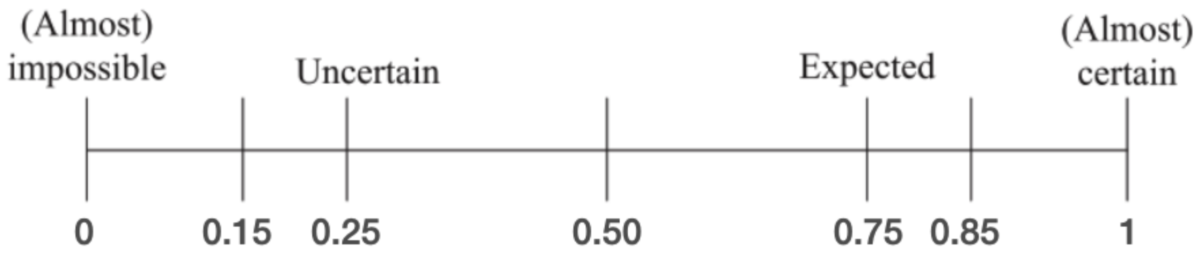
1. Each expert in the project team was asked to indicate which uncertainties, present in the uncertainty table, are most likely to occur according to their current knowledge, prior knowledge, and evidence of similar uncertainties that may have occurred. Next, they attribute mass to these elements (uncertainties).

2. Then, they must apply the script produced to assist in the process of calculating belief and plausibility using the formula presented in the DST.

3. In this step, the monitoring of project uncertainties was carried out in order to monitor the evolution of activities and associated uncertainties, If necessary, a new assessment of project uncertainties using DST is carried out.

Regarding activity one, a modification to the original scale was needed to meet the DST requirement that experts must assign a mass that is between intervals [0, 1]. Figure 30 presents an example of the scale used to assist elicit the Conditional Probability Table (CPT) from experts.

Figure 30 – Probability Scale



Source: Adapted from knochenhauer2013using

As to activity two, the belief will be used as input for the following process, serving as input for producing the conditional probability distribution tables of the Bayesian network.

Furthermore, according to coutto2007problema, plausibility can be understood as being the risk of the occurrence of a specific event. Plausibility represents the degree to which the uncertain event is admissible, given the evidence known to the experts.

7.3.3.1 The Role of Expert Knowledge

For Booker, Anderson e Meyer (2001), the knowledge provided by experts is essential for understanding, estimating, and propagating different types of uncertainties. However, experts are human and subject to cognitive and motivational biases.

One way to minimize these biases is by using elicitation techniques and alternative forms of expert elicitation present in theories to deal with uncertainties.

With regard to this, the following advice is offered by Hanea, Hemming e Nane (2021)

- Expert elicitation is a decision support tool;
- Expert judgment should not replace empirical data;
- Expert elicitation requires methodological rigor.

This process is based on mathematical calculations on the Theory of Evidence presented in Chapter 2. These calculations can be performed using Excel spreadsheets or other more elaborate electronic calculation software. However, as a way to help calculate belief and plausibility based on the evidence collected from each expert, a script was developed in Python code using the PyDS library[2], which is a Python library for performing calculations in the Dempster-Shafer Theory of Evidence.

The input for this script is the set of uncertainties mapped in the previous process or their respective mass assigned within the interval [0, 1] by each of the project team experts. This assignment of mass can also follow the pattern found in Figure 30 to give the project team more options.

The output of this script is the calculation of the project team's degree of belief in the occurrence of particular uncertainty, which is one of the advantages of using the DST, and thus presents a measure of the degree of uncertainty of the entire project team. This output serves as input to the following process. This script can be found in APPENDIX D.

7.3.4 Analysis of Interdependence between Uncertainties

Figure 31 shows the processes of the model for the Analysis of Interdependence between Uncertainties.

This stage of the process deals with the process of building a Bayesian Network (BN). How to build a BN undergoes the following activities:

1. *Data or Knowledge of Experts*. The process and subsequent construction of a BN can be done by using data collected from a survey or directly by using the opinions of experts. This process deals with the form adopted to collect data and build the BN.

2. Assessment Variables. This step deals with identifying the variables composing the BN and the emergence of nodes within the topology of the network.The input of this process will be the output of the previous process, as the nodes to be used will be the previously identified uncertainties. The aim of defining the topology of the BN is to create graphs (DAGs) that can represent the causal relationships between uncertainties.

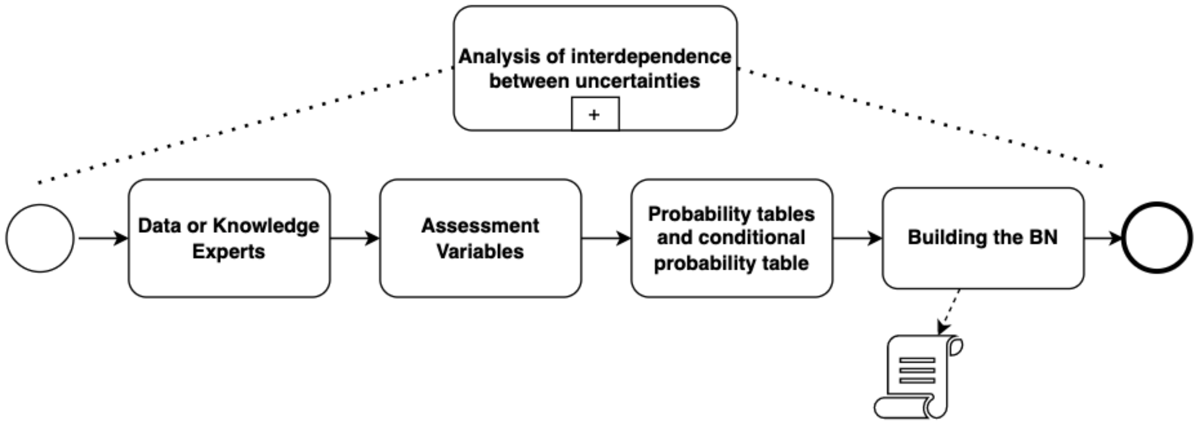
3. Definition of a Conditional Probability Table (CPT). This step is undertaken by producing a probability table of the occurrence of uncertainties in the project. This table of probabilities results from the step that used DST to quantify the uncertainties. Thus,

[2] https://github.com/reineking/pyds

the number or interval associated with each uncertainty can be used as a CPT and input to the BN.

4. For this step, a script was also developed using the R programming language and the bnlearn package to assist in building the BN from the output of the previous process. This script can be found in the GitHub[3] repository created to store the code artifacts produced in this thesis.

Figure 31 – Analysis of Interdependence between Uncertainties



Source: Author

There will be a network of relationships between the uncertainties and the quantification of uncertainties propagated throughout the BN. It will be possible to understand which uncertainties are impacting each other and to quantify and prioritize actions associated with mitigating these uncertainties in the project. It will be possible to realize to what degree (quantification) an uncertainty may impact another uncertainty.

These processes help to answer questions such as:

1. Which uncertainties are most impacted by other uncertainties?

2. Which uncertainties should be prioritized in the process of managing uncertainty?

Building a BN can be done manually, but the CPT calculation becomes increasingly complex as the number of nodes increases. For this reason, a script was created to assist in building the BN.

The input of this script is the set of nodes and their respective relationships. It is important to remember that, for the purposes of this thesis, the BN's input nodes are the uncertainties mapped in process number 2 of the Euler model.

[3] https://github.com/jeffersonjpa/quantifying$_{u}ncertainty$

The output of this script is the BN view and the network's conditional probability tables of the nodes (uncertainties). An essential addition as an output of this script is the visualization of the built network with highlighted visual details such as node size and colors indicating essential highlights. Some of these highlights can be the size of the node which is based on the size of the degree of objective uncertainty in the previous process.

One of the main objectives is to prioritize uncertainties that the project manager must deal with as a priority in agile projects. This knowledge is expected to prioritize the uncertainties with the most significant impact, greatest project predictability, and thus lead to better project coordination.

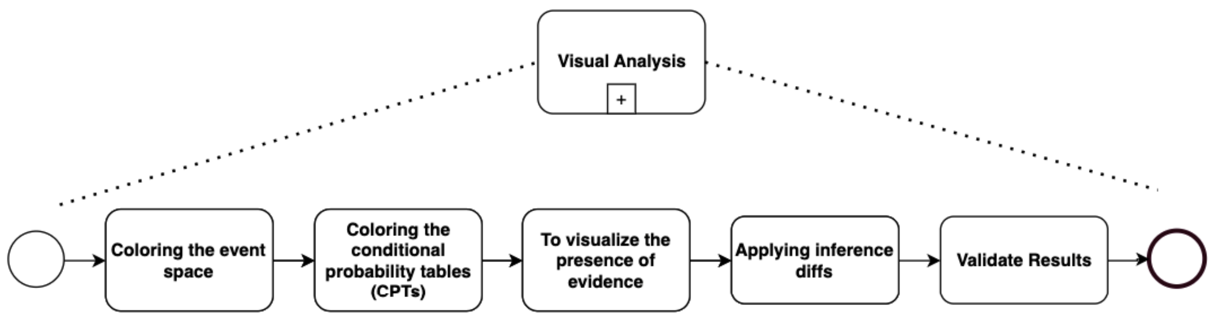
The output of this process serves as the input to the following process. This script can be found in APPENDIX D.

7.3.5 Visual Analysis of the Network

As important as the generation of the BN are the BN visualization methods. Network visualization cannot be seen only as an aesthetic concern. For tufte2001visual, "often the most effective way to describe, explore, and summarize a set of numbers - even an extensive set - is to look at pictures of those numbers." Moreover, the author adds that data graphics can be even simpler and more powerful than any other method.

Figure 32 shows the processes for being able to produce a Visual Analysis of the Euler model.

Figure 32 – Visual Analysis

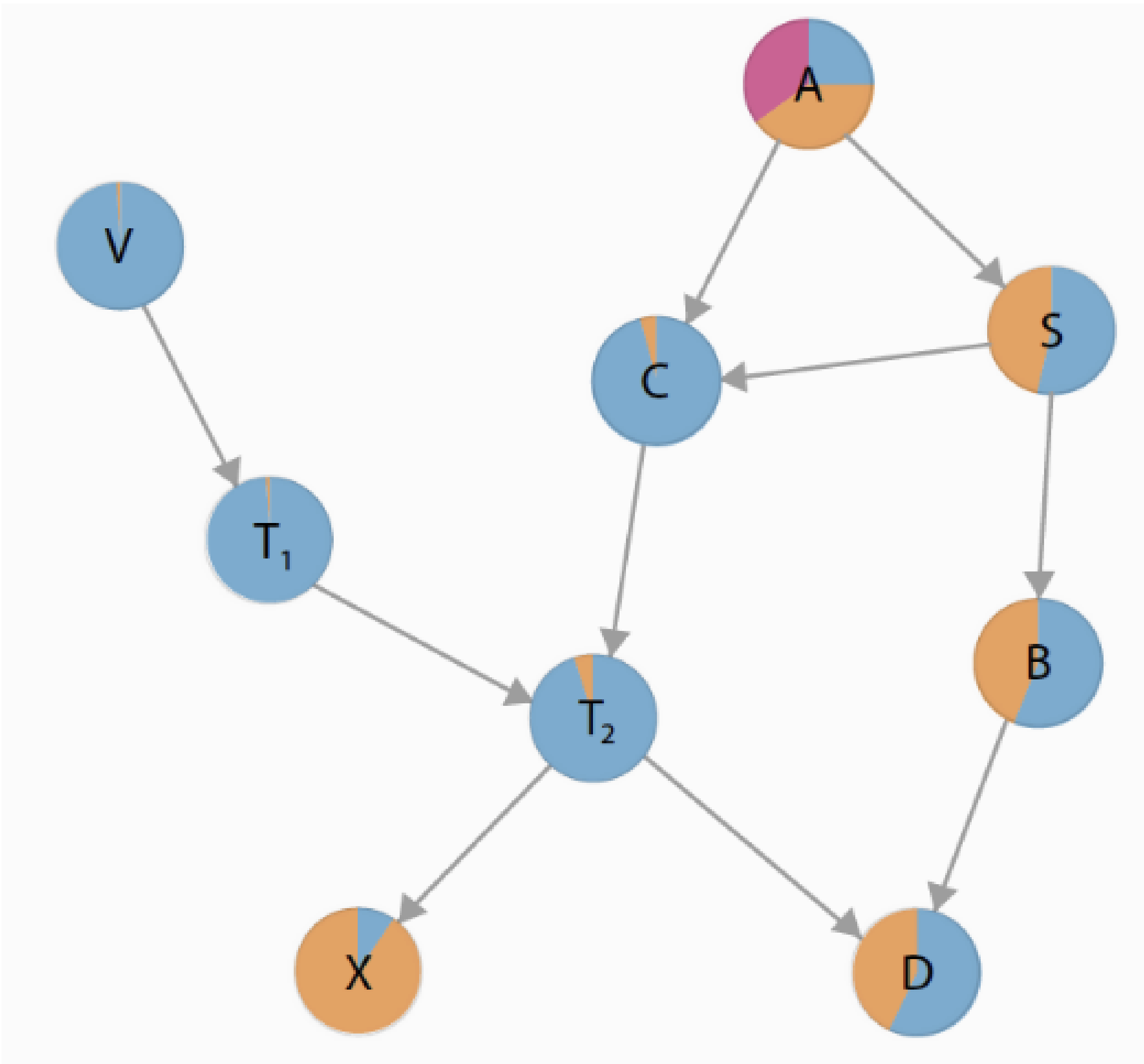


Source: Author

Potential correlations or causal relationships between variables can be quickly visualized. Studies like sundarararajan2013multi focus on improving navigation and visualization in large networks with networks that range up to thousands of variables. In order to understand these large networks and make the networks more efficient and better visualized, new methods are needed.

Figure 33 shows the visual representation of a marginal distribution. Marginal distribution or CPT was directly inserted in the nodes by having them represent small pie charts so as to visualize the proportional size of the mass attributed to a certain variable. In our specific case, the mass attributed to a particular uncertainty, started at the 12 o'clock position, and slices were allocated in clockwise order.

Figure 33 – Visual Marginal Distributions



Source:champion2017visualizing

This step of the process deals with the visualization of insights in the Bayesian Network built in the previous process. This process undergoes the following activities:

1. Coloring the event space. Possibilities related to the BN visualization step go beyond the visualization of its nodes. zapata1999visualization explore the use of colors to highlight

nodes and communicate information related to BNs.

2. This step deals with coloring the CPT.

3. To visualize the presence of evidence.

4. Applying inference diffs. This is a more sophisticated analytical visualization technique that extends pie charts in BN visualization Champion e Elkan (2017).

As to the first activity, the forms of use of applied visual techniques of the BN involve the variation of the thickness of the lengths of the BN edges, thereby indicating the degree of influence existing between a parent node and a child node using thicker thicknesses as a way to highlight the influence of one edge over another edge (ZAPATA-RIVERA; NEUFELD; GREER, 1999).

As to the second activity, visual techniques in CPT, cossalter2011visualizing propose 2D miniatures of heat maps associated with the edges of the BN instead of the traditional colorless matrixes. Other forms of BN visualization can include "bubble lines" connecting nodes in the network to CPT, making it easier for users to maintain their orientation and understanding of the network.

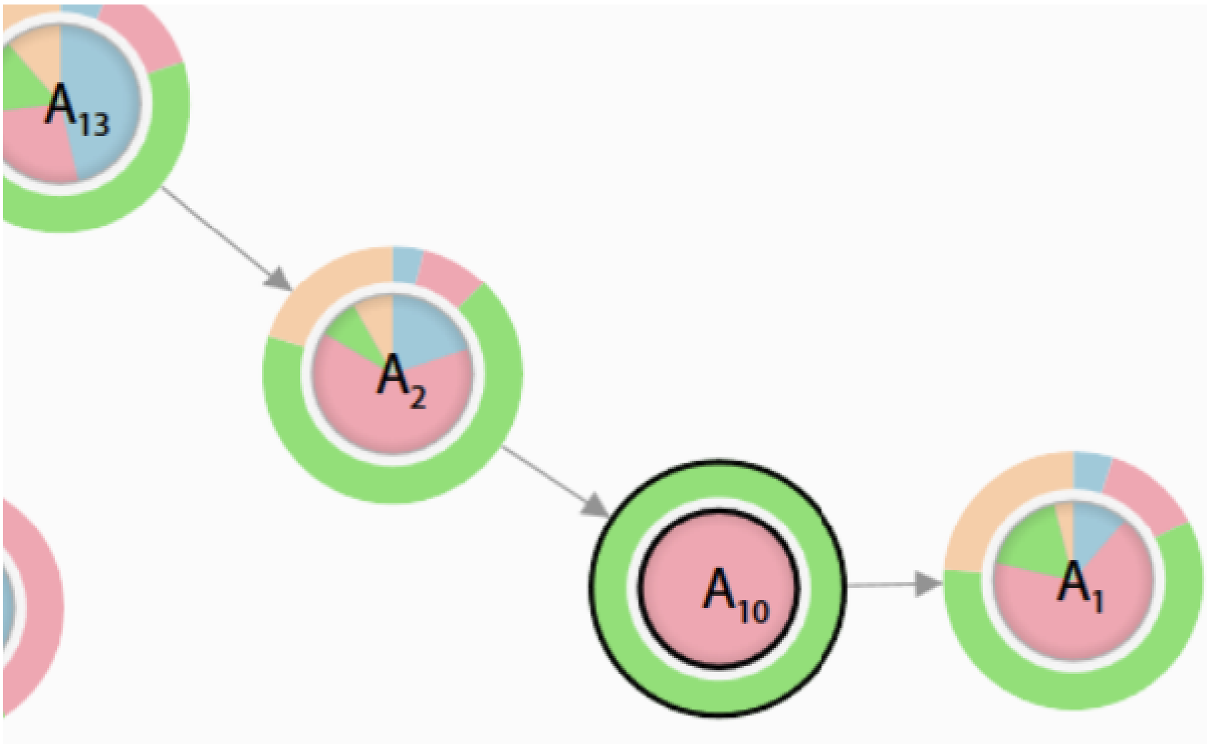
As to the third activity, the purpose of highlighting the visual interpretation of Bayesian Networks is to capture more dynamic visualizations and insights that shed new light on the information flow. Highlights that were not previously available without the colors or visual attributions.

There are inference diffs among the techniques for visual highlighting of BN found in the literature.

An inference diff is the set of pairs of conditional probability distributions for each random variable according to two sets of evidence.

As a way of assisting the analysis of the produced graph, create this Table 14 that in its first column reports the characteristic of the graph (size or color), the second column interprets its result (higher or lower), and the third column reports the range of possibilities for a property. Likewise, the size does not represent the quantification of the value of the credence in uncertainty, and the hue of the color represents a scale that goes from more reddish tones (more significant uncertainties) to lighter tones (lesser uncertainties).

Figure 34 – An inference diff between two evidence sets



Source:Champion e Elkan (2017)

Table 24 – Visual evaluation of the graph

Feature	Explanation	Range
Size	The larger the size of the team, the greater the belief of the team in the existence of uncertainty	[0-1]
Color	Goes from more reddish tones (greater uncertainties) to lighter tones (less uncertainties)	Reddish tones to lighter tones

The output of this process is an even richer visual representation of the representation elaborated by a Bayesian Network. This visual representation is only possible using the techniques presented in this step. It will now be possible to better visually analyze the whole set of uncertainties and their relationships within the Euler process.

7.3.6 Integrating Risk Management

Placing the risk management process at the end of the Euler model processes reflects the view that this process is part of the process of managing uncertainty, which is a more extensive process than the previous one. This view agrees with that developed in article$_{m}arinho_{s}ampaio$.

Once the uncertainties can be identified, quantified, and mapped through their interrelationships, the next step would be managing the risks that these uncertainties may represent for the project.

This can be done using traditional risk management techniques such as those in the PMBoK Guide (PMI, 2021).

1. Risk Identification - Attempting to identify threats early in the project puts the project manager one step ahead to prevent, when possible, and control, when necessary, threats to project objectives. This risk identification process is already undertaken in the Euler model when calculating belief and plausibility in the Expert Analysis of Uncertainties process. Remembering what happened in Chapter 2, we explain that the plausibility function can be interpreted as the risk of occurrence of event A. Therefore, at the end of the referred process, we quantified the belief by calculating the Belief Function and the risks by calculating the Plausibility Function formula.

2. Qualitative and Quantitative Risk Analysis - This step deals with assessing the importance of each risk in order to categorize, measure, and prioritize individual risks. This objective can be achieved by applying some form of risk measurement, e.g. by applying probabilistic techniques to identify and perceive correlations between risks and their interdependencies (PMI, 2019b). The output of this process gives us the perception of the degree of exposure of the project to risks.

3. Plan and implement risk response - In this process, practical risk response actions must be planned and implemented. These actions must consider prioritizing the risks and uncertainties defined so that they can be more effective.

4. Monitor Risk - The last of the activities in this process is risk monitoring. These activities reassess, at a specified frequency, the status of previously identified risks. This process aims to identify and reassess emerging, secondary, and residual risks.

Figure 35 shows the processes of Risk Management.

The actions described in item three must also consider the dual nature of risk management which seeks to deal both with risks and opportunities.

According to PMI (2019b), some of the possible actions for risk management are: escalate, avoid, mitigate, transfer, and accept; possible actions for opportunities are: escalate, explore, share, improve and accept.

Figure 35 – Risk Management

Source: Author

7.3.7 Feeding the Body of Knowledge

The last process of the Euler model deals with knowledge produced by the process. The step of feeding a body of knowledge with the result of this process. Despite being the last step, it becomes an essential one for refining the results presented by the model and for using these results in future projects.

Four subprocesses divide this process that deals with the feeding of the body of knowledge.

1. Choosing the solution for storing the knowledge produced while applying the Euler model. This sub-process must consider that the tool to store the knowledge produced while applying the model must store and return the information present in its body of knowledge in a timely way. Timely means the query interface for previously recorded uncertainties and ways to resolve these uncertainties must be quickly and easily accessible;

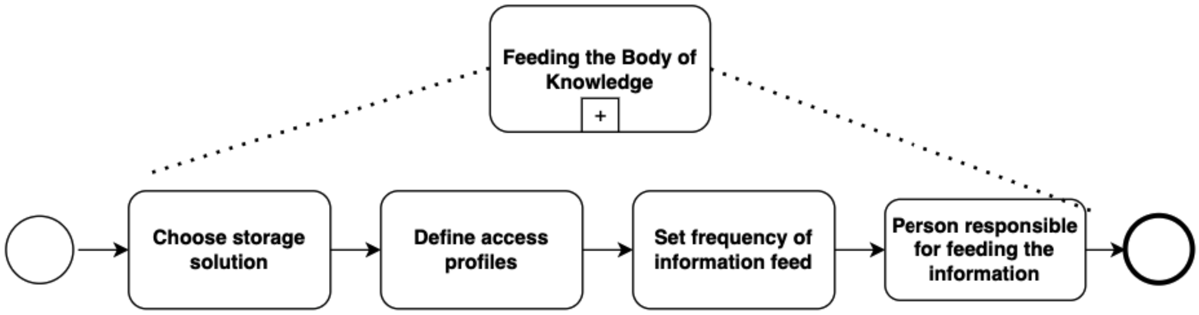
2. Definition of access profiles to the body of knowledge within the project team. All company stakeholders should have access. However, some information can be considered strategic for the institution. Therefore, not all information should be accessible to all profiles;

3. Definition of how often this body of knowledge will be fed and who will be responsible for feeding the body of knowledge.

4. Indication of the team member responsible for curating the information produced. This person or profile should be responsible for keeping the knowledge base up to date or for ensuring other stakeholders keep the knowledge base up to date.

Figure 36 shows the processes of Feeding the Body of Knowledge.

Figure 36 – Subprocesses for feeding the Body of Knowledge



Source: Author

The output of this process is the representation of the knowledge produced and stored. This knowledge must serve as input for future application cycles of the Euler model and future applications of the model in new projects.

7.4 CLOSING REMARKS

The challenge present in the daily routine of the project manager is that decisions that must be made now about the future are inherently uncertain.

In addition, uncertainty in organizational environments is usually measured as perceived by the stakeholders involved in the project. Because of this, it is difficult to assess the accuracy of the experts' input in project management.

In order to fill these gaps, this chapter presented a model for managing epistemic uncertainties in software projects that apply agile methodologies—primarily using epistemic uncertainty quantification techniques and taking into account the existing relationships between the various sources of uncertainty in software projects.

8 FINAL CONSIDERATIONS

The last chapter of this thesis presents the answers to the research questions introduced in the first chapter and the research contributions, limitations, and future work.

8.1 ANSWERS TO RESEARCH QUESTIONS

Applying agile project management practices has grown significantly, especially when it comes to software projects. Agility coexists with uncertainty, since one of the principles of agile projects is the possibility of rapid change. This characteristic of agile projects makes it essential to know and apply approaches for managing uncertain events in projects.

Epistemic uncertainty is associated with the lack of information about certain project phases which is very characteristic of projects that adopt agile methodologies. Thus, applying epistemic uncertainty management in software projects can represent a competitive advantage for companies involved in agile software development.

Therefore, throughout this thesis, a model for managing epistemic uncertainties in software projects was built, evaluated, and presented in an initial version (V1) and in its current version (v2). Therefore, it is now possible to answer the main research question **RQ - How do agile teams quantify epistemic uncertainties in software projects?** and the other questions listed in Chapter 1.

The related subquestion **RQ1.1 What approaches to quantifying epistemic uncertainty in software projects are already in the literature?** was used to investigate the literature in search of approaches to quantify epistemic uncertainty in software projects, and found nine approaches for quantifying epistemic uncertainty, namely: (i) The Dempster-Shafer Theory of Evidence, (ii)Coherent upper and lower probability, (iii)The NUSAP Scheme, SAPIUM - Systematic Approach for model Input Uncertainty quantification Methodology, (iv)Interval Analysis, (v)Bayes Linear Method, (vi)Bayesian Belief Networks, (vii)C-FASEE - Consistent Fuzzy Analogy-based Software Effort Estimation, (viii)Sensitivity Analysis, (ix)Possibility Theory.

For the related subquestion **RQ1.2 How do we measure the causal relationship between sources of epistemic uncertainty in software projects?**, we find that when measuring causal relationships, Bayesian Networks emerge from the literature with the application of this concept in several practical cases and are applied not only to project management but to several other areas to quantify uncertainties in identifying their interrelationships.

After analyzing the approaches found in the literature, Action Research was undertaken to investigate, in a real project scenario, how agile teams approach the quantification of epistemic uncertainty in software development projects and to seek an answer to the related subquestion **RQ1.3 How do agile teams approach quantifying epistemic uncertainty in software projects?**.

During the development of the actions in Action Research, the following actions that deal with epistemic uncertainty in agile teams emerged as ways to deal with epistemic uncertainty in software projects that use agile methodologies: Action 1 - Research approaches for quantifying uncertainties in projects; Action 2 - Identify sources of uncertainty; Action 3 - Review the team's experiences of uncertainties in projects; Action 4 - Quantify epistemic uncertainties; Action 5 - Relate sources of uncertainty; Action 6 - Develop a preliminary model to quantify epistemic uncertainty in software projects; and Action 7 -Apply pilot of the model developed for a sprint.

After executing the Action Research and building the first version of the model called V1, it was possible to start searching for the answer to the final related subquestion, **RQ1.4 What are the impacts of a model based on quantitative approaches to help manage epistemic uncertainty in software projects?**.

We answered this research question by applying the model empirically and evaluating it in three agile teams from three different institutions: public and private. Adjustments made included renaming some processes and creating new strategies to integrate with the risk management present in project management methodologies.

The conclusion was that the model is relevant and can be used to manage uncertainties in agile software projects. It allows the concepts associated with risks and uncertainties to be disseminated to the project team, thereby identifying and prioritizing uncertainties as what is also related to the uncertainties is to perceive the impact of one uncertainty on another or several other uncertainties. In addition, there is the visual highlight at the end of the construction process of the Bayesian Network.

Thus, since it is possible to quantify and prioritize the uncertainties existing in the project and to perceive their interdependence relationships, a greater predictability of project activities is obtained, resulting in even more excellent safety, increased probability of success, and better project coordination.

One of the works presented in this thesis in the related works section is the study by Khodakarami, Fenton e Neil (2007), where they identify causal relationships between the sources of uncertainty existing in projects, specifically managing project schedules. However, this study does not present anything related to quantifying uncertainty collected in the experts' judgment.

In Khodakarami, Fenton e Neil (2007), the authors conclude that a classic project management problem deals with project uncertainty and that more advanced techniques are needed to capture different aspects of project uncertainty. In addition, this approach uses Bayesian Networks without highlighting its results or the formation of its network that can help project managers interpret the Bayesian Network.

In comparison, the Euler model presents two of its processes the quantification of uncertainty collected from the opinion of experts in the project and a process related to the highlight of the Bayesian network as a way to visualize uncertainties and their relationships better.

Another important study presented in the related works section is Santos e Cardoso-Junior (2018), where the authors deal with planning execution deadlines in projects. To solve this problem, Dempster-Shafer's Theory of Evidence will be applied because it allows uncertainties based on experts' knowledge to be dealt with.

This study presents a complementary approach to the approach presented previously. This study has a way of quantifying the uncertainty collected from the experts' opinions but has no way of looking for interrelationships between the identified uncertainties.

This gap was filled by the Euler model, which sought to unite the two approaches to achieve the quantification of uncertainties through the experts, as well as the interrelationship between these uncertainties with the highlight of the network at the end of the process.

Regarding other approaches to managing uncertainties in projects, the Euler model can be characterized as taking a more quantitative approach in contrast to more qualitative approaches, such as the approach developed by Marinho(MARINHO; SAMPAIO; MOURA, 2017).

However, these authors believe that the quantitative aspect can be associated with the qualitative aspect as a way for the two approaches to complement each other.

Generally speaking, both approaches aim to reduce uncertainties, thus allowing them to be seen as risks and so traditional risk management techniques can be applied.

Another similar aspect on comparing the two approaches is that they both agree that uncertainty management is something more significant and that risk management is a complementary part of the process of managing uncertainty.

Some processes of Marinho's approach (MARINHO; SAMPAIO; MOURA, 2017) such as "sense-making" aim to make "organizations and individuals work uncertainties, ambiguities, changes and problem situations, thus generating inventions and new situations that result in actions that lead to solving the problem and environmental stability", and the "early signs process" which can be defined as "The first inaccurate indications of impending and impactful events; all that is known, is that there are threats and opportunities; those will undoubtedly emerge, but their form, nature, and origin are not yet known". These can be observed in a complementary way in the Euler model by undertaking the process of 'Leveling and Awareness' and the process of 'Analysis of Interdependence between Uncertainties'.

Leveling and Awareness present practical examples of risks and uncertainties taken from the body of knowledge and the literature to reduce the doubts of team members and thus form an awareness in the project team of the importance of the topic and uncertainty management practices for the objectives of a project in alignment with what the Sensemaking process aims at in Marinho's approach.

Analysis of Interdependence between Uncertainties aims to construct a Bayesian network to demonstrate the interdependent relationships between uncertainties identified in the project. The perception of these relationships that were not apparent and that now have visual and quantitative prominence can result in the perception of new uncertainties, in tune with what is done in the process for detecting early signs present in Marinho's approach.

Thus, the Euler model can support Marinho's approach in the steps listed above in a complementary way, and they can be considered as being an auxiliary approach to each other. So, we observed that there are alignments between the processes of one and approach to managing uncertainties in software projects.

The Euler model was created based on the relationship between agility and uncertainty. As I said in the introduction of this chapter, agility coexists with uncertainty since one of the two main principles is an adaptation to rapid changes, and these rapid changes can bring uncertainties to the project.

Despite its origin in an agile environment, it is not possible to say that the Euler model is restricted to software development projects that use agile project management methodologies to the detriment of traditional project management methodologies, which also have high uncertainties associated with your activities.

That said, we understand that a model with auxiliary management and uncertainty can be helpful for all of the nature of projects, whether traditional or agile. Therefore, the type of project that has its nature associated with more uncertain environments would benefit more from a model for uncertainty management. Although Euler has been built to be practical, and it can be observed through the availability of an application guide, there is still an overhead associated with its application.

The same type of observation can be extended for applying Euler only in software development projects. There are peculiarities in software development projects in the same way as in civil construction projects. However, the two types of projects have people, resources, and activity management in common. All these types of projects can also benefit from a model for managing uncertainty in the same way as software projects.

One of the limitations mentioned in the pertinent section is the reduced number of members of the age teams (three to four). We also understand that this is a limitation of the research, not of the model. The sources of uncertainty in projects may be associated with various aspects of project management, from its schedule, requirements, deliveries, and quality management. People management is an important aspect, but it is not the only source of uncertainty in projects. There are various sources of uncertainty in projects that are not associated with the reduced number of members of the age teams. Likewise, we understand that the reduced number is not decisive for the use or not of the model. Therefore studies with older teams with more members need to be carried out with comparative analyses to obtain more conclusive data.

8.2 CONTRIBUTIONS TO RESEARCH

One of the practical results of this thesis is that it makes available, both for the academy and for practitioners of uncertainty management, a model that can help project managers to tackle how to deal with uncertainties that arise in software projects that use agile methodologies.

We understand that by allowing the project manager to quantify the uncertainty, he/she can have a more accurate idea of the degree of uncertainty in a project. Thus, the project manager can make better decisions during project management. Actions to reduce uncertainty can be better targeted and more assertively applied to those uncertainties that significantly impact the project.

Another expected result is that the community of practitioners and academics will contribute to expanding knowledge associated with managing project uncertainty.

This thesis highlights the importance of searching for approaches to deal with the interdependence between uncertainties and sources of uncertainty, as well as, ways to quantify experts' perception of the project based on applying Bayesian logic.

Some of the results obtained by developing this research were published in vehicles of high visibility among researchers and practitioners of Software Engineering, namely:

- Barbosa, J. F., de Moura, H. P., Marinho, M. L. M. (2020, October). Towards a Quantitative Model to Deal with Uncertainty Management in Software Projects. In Anais Estendidos do XI Congresso Brasileiro de Software: Teoria e Prática (pp. 91-99). SBC.
- Barbosa, J. F., M. M. L. M. . d. M. H. P. (2021). Em direção a um modelo para quantificação da incerteza epistêmica em projetos de software: uma pesquisa-ação. Revista Ibérica de Sistemas e Tecnologias de Informação, 44:67–83.
- Barbosa, J. F., M. M. L. M. . d. M. H. P. (2023). An empirical evaluation of a model to deal with epistemic uncertainty in agile software project management. SBSI 2023 (Brazilian Symposium on Information Systems.

Other results that have potential for publication are to be found in Chapter 4, which deals with the quasi-Systematic Literature Review; in Chapter 6, which deals with the empirical application and evaluation of the model; and in Chapter 7, which presents the Euler model.

Other findings co-related to this thesis and the theme of project management have been published:

- da Silva Araujo, J. I., Moser, P. C., Domingos, E. R., Afonso, V. A. S. D., Barbosa, J. F. (2018). Benefícios e Limitações da Simplicidade em Projetos Inovadores de Software: uma revisão sistemática da literatura. GESTÃO. Org, 16(7), 279-292.
- Arruda, G. M., Barbosa, J. F., de Oliveira, P. E. A., da Silva, D. C., Hollanda, A. C. R. (2019). Análise da maturidade em gestão de projetos dos laboratórios de pesquisa em engenharia de uma universidade do nordeste brasileiro. Brazilian Journal of Development, 5(11), 2726-27247.

For the author of this research, professionally and academically, it was of great value to develop research in an area that has been constantly growing and is of great importance to project managers.

The creation of relationships with other researchers can also be considered a noteworthy result of this research, as such relationshipscan result in several different types of research that can contribute to the growth of the Project Management and Software Engineering area.

8.3 LIMITATIONS OF THE RESEARCH

One of the known limitations of qualitative research is the inability to reproduce and generalize its results.

Despite all efforts regarding rigor and combating threats to research viability, we know some biases associated with researcher participation may have emerged in the research process.

Regarding the research method associated with a quasi-Systematic Literature Review, one of the concerns is always about collecting sufficient and quality primary studies to answer research questions. Another concern for the researcher is to know if the literature coverage was done despite the growing number of periodicals considered predatory that affect the quality of scientific research produced in Computer Science.

Practices such as manual searching and snowballing can reduce this limitation, but the concern with this type of problem still exists. In addition to all these limitations, there is the bias that the researcher can introduce when selecting articles and when analyzing them.

Another limitation is associated with the fact that the model was developed based on research carried out in an agile team with a small number of members (three or four members). This type of systems development team has some characteristics that may differ from the format of systems development in mega projects or large projects, such as the complexity of dealing with many tasks, uncertainties, and risks.

8.4 THREATS TO VALIDITY

The validity and reliability of a study depend upon the investigators' ethics. Every researcher needs to pay particular attention to the results of the validation. This item is related to the level of trust the researcher can place in the search process to answer the research question (MERRIAM; TISDELL, 2015).

As stated in Travassos, Gurov e Amaral (2002), there are four types of validity results for a study: construct validity, internal validity, external validity, and conclusion validity.

8.4.1 Construct Validity

The first type of threat to the validity of this research is construct validity. This threat is related to the design of the research and how it was created to measure the research results. By definition, Action Research needs a team to be inserted in the context of the problem to be investigated. Construct validity is extremely sensitive to any bias (STARON, 2020).

The most crucial threat related to this validity is the fact that the author of this research is also the Scrum Master of the project under study, so his observations and interpretations may be based on his values and expectations. However, to mitigate this threat, a researcher with a Ph.D. in Software Engineering and extensive experience in the area was invited to evaluate the results and check if the research objectives were followed.

8.4.2 Internal Validity

This type of threat to research validity aims to validate the operationalization of Action Research. Action Research necessarily includes conducting an in-depth analysis of the Action Taking phase. It is necessary to assess the threats associated with data collection and the planning for using the measurement instrument. This Action Research was conducted and the information for the cycles was captured from semi-structured interviews with software project team members, by observing the team, and via the platform used for project management and task recording, which was the software Redmine[1]. The Scrum Master and external researcher analyzed all data to mitigate this threat. Another point that could threaten the internal validity of this research would be choosing respondents by subjective selection. All project participants were interviewed to avoid this type of threat.

8.4.3 External Validity

This threat is related to the ability to generalize the results found in the Action Research. This type of threat is related to only one project being studied, focusing on only a single agile team. However, we believe that since this Action Research was carried out outside the controlled environment of a laboratory, it was developed within the context of a company. We can conclude that the solution found can be applied to contexts similar to that found as this

[1] https://redmine.org/

research object.

8.4.4 Conclusion Validity

This threat is related to presenting correct conclusions from the data collected, which is challenging since Action Research combines quantitative and qualitative data. This validity aims to ensure that the results found are consistent and that they can be replicated systematically. To mitigate the threats to this validity, all quantitative and qualitative data of the study were stored in a database, organized, and made public on a cloud platform [2] for future access and replication of this study, while preserving the participants' anonymity.

8.5 FUTURE STUDIES

Considering the scope of this research, several opportunities for developing the scope of similar research studies in future are suggested. The following are some points for further investigation:

- Apply the Euler model to a larger number of agile development teams with different formats, such as colocated and distributed teams.
- Develop a computational tool, in addition to the script already developed, with a web interface to support how to apply the model and make it more accessible to different organizations.
- Look for other approaches oriented to mathematical models, such as Dynamical Systems, to apply more complex mathematical models to quantify uncertainty in projects.
- Perform a robustness analysis of the Euler model to identify improvement points and gaps to be solved in later model versions.
- Adapt the model so that it can be applied to other areas of project management knowledge.

[2] http://bit.ly/QuantUncertainty

REFERENCES

ARKSEY, H.; O'MALLEY, L. Scoping studies: towards a methodological framework. *International journal of social research methodology*, Taylor & Francis, v. 8, n. 1, p. 19–32, 2005.

BARBOSA J. F., M. M. L. M. . d. M. H. P. Em direção a um modelo para quantificação da incerteza epistêmica em projetos de software: uma pesquisa-ação. *Revista Ibérica de Sistemas e Tecnologias de Informação*, Cambridge University Press, v. 44, p. 67–83, 2021.

BASU, S. *Evaluation of hazard and risk analysis*. 1st edition. ed. [S.l.]: Academic Press, Cambridge, MA, USA, 2016. ISBN 978-0128038215.

BECK, K.; BEEDLE, M.; BENNEKUM, A. V.; COCKBURN, A.; CUNNINGHAM, W.; FOWLER, M.; GRENNING, J.; HIGHSMITH, J.; HUNT, A.; JEFFRIES, R. et al. Manifesto for agile software development. Snowbird, UT, 2001.

BERNSTEIN, P.; BERNSTEIN, P. *Against the Gods: The Remarkable Story of Risk*. Wiley, 1998. ISBN 9780471295631. Disponível em: <https://books.google.com.br/books?id=uTje6PYAijUC>.

BOEHM, B. W. Software risk management: Principles and practices. *IEEE Softw.*, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 8, n. 1, p. 32–41, jan. 1991. ISSN 0740-7459. Disponível em: <http://dx.doi.org/10.1109/52.62930>.

BOIVIN, C. *Introduction to Belief Functions: The Monty Hall Game*. [S.l.], 2019.

BOOKER, J. M.; ANDERSON, M. C.; MEYER, M. A. The role of expert knowledge in uncertainty quantification (are we adding more uncertainty or more understanding?). In: *Seventh Army Conference on Applied Statistics*. [S.l.]: American Institute of Aeronautics and Astronautics, 2001. p. 155–161.

BRAUN, V.; CLARKE, V. Using thematic analysis in psychology. *Qualitative research in psychology*, Taylor & Francis, v. 3, n. 2, p. 77–101, 2006.

BRECHNER, E. *Agile project management with Kanban*. [S.l.]: Pearson Education, 2015.

BROWNING, T.; RAMASESH, R. Reducing unwelcome surprises in project management. v. 56, p. 53–62, 03 2015.

CHAMPION, C.; ELKAN, C. Visualizing the consequences of evidence in bayesian networks. *arXiv preprint arXiv:1707.00791*, 2017.

CHAPMAN, C.; WARD, S. *How to manage project opportunity and risk*. [S.l.: s.n.], 2011.

CHENARANI, A.; DRUZHININ, E. A quantitative measure for evaluating project uncertainty under variation and risk effects. *Engineering, Technology & Applied Science Research*, v. 7, n. 5, p. 2083–2088, 2017.

CHOUAIBI, S.; FESTA, G.; QUAGLIA, R.; ROSSI, M. The risky impact of digital transformation on organizational performance – evidence from tunisia. *Technological Forecasting and Social Change*, v. 178, p. 121571, 2022. ISSN 0040-1625. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0040162522001032>.

CRUZ, I. R.; TROFFAES, M. C.; SAHLIN, U. A suggestion for the quantification of precise and bounded probability to quantify epistemic uncertainty in scientific assessments. *Risk Analysis*, Wiley Online Library, v. 42, n. 2, p. 239–253, 2022.

CRUZES, D. S.; DYBÅ, T. Recommended steps for thematic synthesis in software engineering. *International Symposium on Empirical Software Engineering and Measurement*, n. April 2014, p. 275–284, 2011. ISSN 19493789.

DAMNJANOVIC, I.; REINSCHMIDT, K. Introduction to uncertainty and risk. In: ______. *Data Analytics for Engineering and Construction Project Risk Management*. Cham: Springer International Publishing, 2020. p. 3–22. ISBN 978-3-030-14251-3. Disponível em: <https://doi.org/10.1007/978-3-030-14251-3_1>.

DANTAS, E.; NETO, A. S.; PERKUSICH, M.; ALMEIDA, H.; PERKUSICH, A. Using bayesian networks to support managing technological risk on software projects. In: SBC. *Anais do I Workshop Brasileiro de Engenharia de Software Inteligente*. [S.l.], 2021. p. 1–6.

DEMPSTER, A. P. Upper and lower probabilities induced by a multivalued mapping. In: *Classic works of the Dempster-Shafer theory of belief functions*. [S.l.]: Springer, 2008. p. 57–72.

DINGSØYR, T.; NERUR, S.; BALIJEPALLY, V.; MOE, N. B. A decade of agile methodologies: Towards explaining agile software development. *Journal of systems and software*, Elsevier, v. 85, n. 6, p. 1213–1221, 2012.

DYBÅ, T.; DINGSØYR, T. Empirical studies of agile software development: A systematic review. *Information and software technology*, Elsevier, v. 50, n. 9-10, p. 833–859, 2008.

DYBA, T.; DINGSOYR, T.; HANSSEN, G. K. Applying systematic reviews to diverse study types: An experience report. In: IEEE. *First international symposium on empirical software engineering and measurement (ESEM 2007)*. [S.l.], 2007. p. 225–234.

DöNMEZ, D.; GROTE, G. Two sides of the same coin – how agile software development teams approach uncertainty as threats and opportunities. *Information and Software Technology*, v. 93, p. 94–111, 2018. ISSN 0950-5849. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0950584917304652>.

EBERT, C.; KUHRMANN, M.; PRIKLADNICKI, R. Global software engineering: Evolution and trends. In: IEEE. *2016 IEEE 11th International Conference on Global Software Engineering (ICGSE)*. [S.l.], 2016. p. 144–153.

ELLIOTT, S. Proof of concept research. *Philosophy of Science*, Cambridge University Press, v. 88, n. 2, p. 258–280, 2021.

ELMAGHRABY, S. E. On the fallacy of averages in project risk management. *European Journal of Operational Research*, Elsevier, v. 165, n. 2, p. 307–313, 2005.

EZGHARI, S.; ZAHI, A. Uncertainty management in software effort estimation using a consistent fuzzy analogy-based method. *Applied Soft Computing*, v. 67, p. 540–557, 2018. ISSN 1568-4946. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1568494618301443>.

FILHO, M. B. D. C.; SOUZA, J. C.; SCHILLING, M. T. Sobre o problema da integração generalizada de dados. *Sba: Controle & Automação Sociedade Brasileira de Automatica*, SciELO Brasil, v. 18, p. 24–43, 2007.

FONTANA, R. M.; JR, V. M.; REINEHR, S.; MALUCELLI, A. Progressive outcomes: A framework for maturing in agile software development. *Journal of Systems and Software*, Elsevier, v. 102, p. 88–108, 2015.

FREIRE, A.; PERKUSICH, M.; SARAIVA, R.; ALMEIDA, H.; PERKUSICH, A. A bayesian networks-based approach to assess and improve the teamwork quality of agile teams. *Information and Software Technology*, Elsevier, v. 100, p. 119–132, 2018.

HANEA, A. M.; HEMMING, V.; NANE, G. F. Uncertainty quantification with experts: present status and research needs. *Risk Analysis*, Wiley Online Library, feb 2021. ISSN 0272-4332.

HAO, Y.; LU, H.; SHI, Y.; GENG, H.; XI, X.; FENG, S. Uncertainty Evaluation of Water Inrush in Karst Tunnels Based on Epistemic Uncertainty with Possibility Theory. *Mathematical Problems in Engineering*, v. 2020, 2020. ISSN 15635147.

HE, Y.; HUSSAINI, M. Y. Mixed aleatory and epistemic uncertainty propagation using dempster–shafer theory. *Journal of Computational and Applied Mathematics*, v. 429, p. 115234, 2023. ISSN 0377-0427. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0377042723001784>.

HELTON, J. C.; OBERKAMPF, W. L. Alternative representations of epistemic uncertainty. *Reliability Engineering and System Safety*, v. 1, n. 85, p. 1–10, 2004.

HESTER, P. Epistemic uncertainty analysis: an approach using expert judgment and evidential credibility. *Journal of Quality and Reliability Engineering*, Hindawi, v. 2012, 2012.

HOST, M.; RAINER, A.; REGNELL, B. *Case study research in software engineering*. [S.l.]: Wiley, 2012.

HOWELL, D.; WINDAHL, C.; SEIDEL, R. A project contingency framework based on uncertainty and its consequences. *International Journal of Project Management*, Elsevier, v. 28, n. 3, p. 256–264, 2010.

JAKEMAN, J.; ELDRED, M.; XIU, D. Numerical approach for quantification of epistemic uncertainty. *Journal of Computational Physics*, Elsevier, v. 229, n. 12, p. 4648–4663, 2010.

JCGM, J. et al. Evaluation of measurement data—guide to the expression of uncertainty in measurement. *Int. Organ. Stand. Geneva ISBN*, v. 50, p. 134, 2008.

JOHNSON, C. N. The benefits fo PDCA. *Quality Progress*, American Society for Quality, v. 35, n. 5, p. 120, 2002.

KEELE, S. et al. *Guidelines for performing systematic literature reviews in software engineering*. [S.l.], 2007.

KENDIG, C. E. What is proof of concept research and how does it generate epistemic and ethical categories for future scientific practice? *Science and Engineering Ethics*, Springer, v. 22, n. 3, p. 735–753, 2016.

KHODAKARAMI, V.; FENTON, N.; NEIL, M. Project scheduling: Improved approach to incorporate uncertainty using bayesian networks. *Project Management Journal*, v. 38, p. 39–49, 06 2007.

KITCHENHAM, B.; CHARTERS, S. Guidelines for performing systematic literature reviews in software engineering. In: . Washington, DC, USA: IEEE Computer Society, 2007. p. 273–281. ISBN 0-7695-2163-0. Disponível em: <http://dl.acm.org/citation.cfm?id=998675.999432>.

KIUREGHIAN, A. D.; DITLEVSEN, O. Aleatory or epistemic? does it matter? *Structural safety*, Elsevier, v. 31, n. 2, p. 105–112, 2009.

Knight Frank, H. Risk, uncertainty and profit. ,): Hart, Schaffner & Marx, 1921.

KNOCHENHAUER, M.; SWALING, V. H.; DEDDA, F.; HANSSON, F.; SJÖKVIST, S.; SUNNEGAERD, K. *Using Bayesian Belief Network (BBN) modelling for rapid source term prediction. Final report*. [S.l.], 2013.

KOVACH, S. *Detecção de fraudes em transações financeiras via Internet em tempo real.* Tese (Doutorado) — Universidade de São Paulo, 2012.

LENFLE, S.; LOCH, C. Lost roots: How project management came to emphasize control over flexibility and novelty. *California Management Review*, SAGE Publications Sage CA: Los Angeles, CA, v. 53, n. 1, p. 32–55, 2010.

LISTER, T. Risk management is project management for adults. *IEEE Software*, v. 14, n. 3, p. 20, 22–, May 1997. ISSN 0740-7459.

LOCH, C. H.; DEMEYER, A.; PICH, M. T. *Managing the Unknown: A New Approach to Managing High Uncertainty and Risk in Projects*. New York, NY, USA: John Wiley & Sons, Inc., 2006. ISBN 0471693057.

MACHERIDIS, N.; NILSSON, C.-H. Managing project complexity. *A managerial view, Lund Institute of Economic Research-Working Paper Series, Department of Business Administration, Lund University, LUND, Sweden*, 2004.

MAGDALENO, A. M.; WERNER, C. M. L.; ARAUJO, R. M. de. Reconciling software development models: A quasi-systematic review. *Journal of Systems and Software*, v. 85, n. 2, p. 351–369, 2012. ISSN 0164-1212. Special issue with selected papers from the 23rd Brazilian Symposium on Software Engineering. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0164121211002287>.

MANZANO, M.; MENDES, E.; GóMEZ, C.; AYALA, C.; FRANCH, X. Using bayesian networks to estimate strategic indicators in the context of rapid software development. In: *Proceedings of the 14th International Conference on Predictive Models and Data Analytics in Software Engineering*. New York, NY, USA: Association for Computing Machinery, 2018. (PROMISE'18), p. 52–55. ISBN 9781450365932. Disponível em: <https://doi.org/10.1145/3273934.3273940>.

MARCONI, M. d. A.; LAKATOS, E. M. *Fundamentos de metodologia cientifica /*. 8. ed.. ed. São Paulo:: Atlas,, 2017.

MARINHO, M.; LIMA, T.; SAMPAIO, S.; MOURA, H. P. de. Uncertainty management in software projects - an action research. In: *Proceedings of the XVIII IberoAmerican Conference on Software Engineering, CIbSE 2015, Lima, Peru, Apr 22-24, 2015.* [S.l.: s.n.], 2015. p. 323.

MARINHO, M.; SAMPAIO, S.; LUNA, A.; LIMA, T.; MOURA, H. Dealing with uncertainties in software project management. In: *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*. [S.l.: s.n.], 2015. p. 889–894.

MARINHO, M.; SAMPAIO, S.; LUNA, A.; LIMA, T.; MOURA, H. Dealing with uncertainties in software project management. In: IEEE. *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*. [S.l.], 2015. p. 889–894.

MARINHO, M.; SAMPAIO, S.; MOURA, H. An approach related to uncertainty in software projects. In: IEEE. *2013 IEEE International Conference on Systems, Man, and Cybernetics*. [S.l.], 2013. p. 894–899.

MARINHO, M.; SAMPAIO, S.; MOURA, H. Uncertainties in software projects management. In: IEEE. *2014 XL Latin American Computing Conference (CLEI)*. [S.l.], 2014. p. 1–10.

MARINHO, M. L.; SAMPAIO, S.; MOURA, H. Managing uncertainty in software projects. *Innovations in Systems and Software Engineering*, v. 14, 08 2017.

MARINHO, M. L. M. Uncertainty management in software projects. Universidade Federal de Pernambuco, 2015.

MATALONGA, S.; RODRIGUES, F.; TRAVASSOS, G. H. Characterizing testing methods for context-aware software systems: Results from a quasi -systematic literature review. *Journal of Systems and Software*, Elsevier Inc., v. 131, p. 1–21, sep 2017. ISSN 01641212. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S016412121730095X>.

MCLAIN, D. Quantifying project characteristics related to uncertainty. *Project Management Journal*, SAGE Publications Sage CA: Los Angeles, CA, v. 40, n. 4, p. 60–73, dec 2009. ISSN 8756-9728.

MENDES, E.; PERKUSICH, M.; FREITAS, V.; NUNES, J. a. Using bayesian network to estimate the value of decisions within the context of value-based software engineering. In: *Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018*. New York, NY, USA: Association for Computing Machinery, 2018. (EASE'18), p. 90–100. ISBN 9781450364034. Disponível em: <https://doi.org/10.1145/3210459.3210468>.

MERRIAM, S.; TISDELL, E. *Qualitative Research: A Guide to Design and Implementation*. Wiley, 2015. (The Jossey-Bass higher and adult education series). ISBN 9781119003656. Disponível em: <https://books.google.com.br/books?id=xGMdCgAAQBAJ>.

MEYER, A. C. L. D.; LOCH, C. H.; PICH, M. T. Managing project uncertainty: from variation to chaos. *MIT Sloan Management Review*, Sloan Management Review Association, v. 43, n. 2, p. 60, 2002.

MILES, M. B.; HUBERMAN, A. M.; SALDAÑA, J. *Qualitative data analysis: A methods sourcebook*. [S.l.]: Sage publications, 2018.

MKRTCHYAN, L.; PODOFILLINI, L.; DANG, V. N. Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability engineering & system safety*, Elsevier, v. 139, p. 1–16, 2015.

MOURA, H. Software project framework. *Federal Univerity of Pernambuco, Recife, Pernambuco, Tech. Rep., 2011*, 2011.

NEAPOLITAN, R. E. et al. *Learning bayesian networks*. [S.l.]: Pearson Prentice Hall Upper Saddle River, NJ, 2004. v. 38.

OBERKAMPF, W. L.; DIEGERT, K. V.; ALVIN, K. F.; RUTHERFORD, B. M. Variability, uncertainty, and error in computational simulation. *ASME-PUBLICATIONS-HTD*, ASME, v. 357, p. 259–272, 1998.

PADALKAR, M.; GOPINATH, S. Six decades of project management research: Thematic trends and future opportunities. *International Journal of Project Management*, v. 34, p. 1305–1321, 10 2016.

PAUL, P. M. *Qualitative research and evaluation methods*. 4. ed. Saint Paul, MN: SAGE Publications, Inc, 2015.

PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B.; SILVA, G. L. *Estatística Bayesiana*. 2a. ed. Lisboa: Fundação Calouste Gulbenkian, 2018.

PEARL, J. Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, v. 29, n. 3, p. 241–288, 1986. ISSN 0004-3702. Disponível em: <https://www.sciencedirect.com/science/article/pii/000437028690072X>.

PEARL, J. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. [S.l.]: Elsevier, 2014.

PERMINOVA, O.; GUSTAFSSON, M.; WIKSTRÖM, K. Defining uncertainty in projects–a new perspective. *International journal of project management*, Elsevier, v. 26, n. 1, p. 73–79, 2008.

PMI, P. *The Standard for Risk Management in Portfolios, Programs and Projects*. [S.l.]: Newtown Square, PA: Project Management Institute, 2019.

PMI, P. *The Standard for Risk Management in Portfolios, Programs and Projects*. [S.l.]: Newtown Square, PA: Project Management Institute, 2019.

PMI, P. M. I. *A Guide to the Project Management Body of Knowledge (PMBOK Guide) – Seventh Edition and The Standard for Project Management*. Newtown Square, PA: Project Management Institute, Inc., 2021. (PMBOK guide). ISBN 978-1-62825-664-2.

RADU, L.-D. Effort prediction in agile software development with bayesian networks. In: *ICSOFT*. [S.l.: s.n.], 2019. p. 238–245.

RAMASESH, R. V.; BROWNING, T. R. A conceptual framework for tackling knowable unknown unknowns in project management. *Journal of Operations Management*, v. 32, n. 4, p. 190 – 204, 2014. ISSN 0272-6963. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0272696314000199>.

RAVITCH, S.; CARL, N. *Qualitative Research: Bridging the Conceptual, Theoretical, and Methodological*. SAGE Publications, 2015. ISBN 9781483351759. Disponível em: <https://books.google.com.br/books?id=V6tiCgAAQBAJ>.

REVIE, M.; BEDFORD, T.; WALLS, L. Supporting reliability decisions during defense procurement using a bayes linear methodology. *IEEE Transactions on Engineering Management*, v. 58, n. 4, p. 662–673, Nov 2011. ISSN 1558-0040. Cited By 18. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-80054863659&doi=10.1109%2fTEM.2011.2131655&partnerID=40&md5=1bfdb6c25bdbd40d9f2e310421a3b7e3>.

ROBSON, C.; MCCARTAN, K. *Real world research*. [S.l.]: John Wiley & Sons, 2016.

RUNESON, P.; HÖST, M. Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering*, Springer, v. 14, n. 2, p. 131–164, 2009.

RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 2. ed. [S.l.]: Pearson Education, 2003. ISBN 0137903952.

SANTOS, A. dos; CARDOSO-JUNIOR, M. Uso da teoria da evidência de dempster-shafer na avaliação da incerteza de prazos em projeto de pd. v. 5, p. 19–29, 12 2018.

SCHUBERT, J. Conflict Management in Dempster-Shafer Theory by Sequential Discounting Using the Degree of Falsity. *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, p. 298–305, 2008.

SCHUBERT, J. Conflict management in Dempster-Shafer theory using the degree of falsity. *International Journal of Approximate Reasoning*, v. 52, n. 3, p. 449–460, 2011. ISSN 0888613X.

SCHWABER, K.; SUTHERLAND, J. The scrum guide. *Scrum Alliance*, v. 21, n. 19, p. 1, 2011.

SCUTARI, M. Learning Bayesian Networks with the bnlearn R Package. *Journal of Statistical Software*, v. 35, n. 3, p. 1–22, aug 2010. ISSN 1548-7660. Disponível em: <http://www.jstatsoft.org/v35/i03/http://arxiv.org/abs/0908.3817>.

SCUTARI, M.; DENIS, J.-B. *Bayesian networks: with examples in R*. [S.l.]: CRC press, 2014.

SERRADOR, P.; PINTO, J. K. Does agile work?—a quantitative analysis of agile project success. *International journal of project management*, Elsevier, v. 33, n. 5, p. 1040–1051, 2015.

SHAFER, G. *A mathematical theory of evidence*. [S.l.]: Princeton university press, 1976. v. 42.

SHAH, H.; HOSDER, S.; WINTER, T. A mixed uncertainty quantification approach using evidence theory and stochastic expansions. *International Journal for Uncertainty Quantification*, Begel House Inc., v. 5, n. 1, 2015.

SILVA, L. G. d. O.; ALMEIDA-FILHO, A. T. de. A new PROMETHEE-based approach applied within a framework for conflict analysis in Evidence Theory integrating three conflict measures. *Expert Systems with Applications*, Elsevier Ltd, v. 113, p. 223–232, dec 2018. ISSN 09574174.

SILVA, L. G. D. O.; De Almeida-Filho, A. T. A multicriteria approach for analysis of conflicts in evidence theory. *Information Sciences*, v. 346-347, p. 275–285, 2016. ISSN 00200255.

SILVA, L. G. de O.; ALMEIDA-FILHO, A. T. de. A new promethee-based approach applied within a framework for conflict analysis in evidence theory integrating three conflict measures. *Expert Systems with Applications*, Elsevier, v. 113, p. 223–232, 2018.

STARON, M. Action research as research methodology in software engineering. In: ______. *Action Research in Software Engineering: Theory and Applications*. Cham: Springer International Publishing, 2020. p. 15–36. ISBN 978-3-030-32610-4. Disponível em: <https://doi.org/10.1007/978-3-030-32610-4_2>.

SUNDARARARAJAN, P. K.; MENGSHOEL, O. J.; SELKER, T. Multi-focus and multi-window techniques for interactive network exploration. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Visualization and Data Analysis 2013*. [S.l.], 2013. v. 8654, p. 86540O.

SWILER, L. P.; PAEZ, T. L.; MAYES, R. L. Epistemic uncertainty quantification tutorial. In: *Proceedings of the 27th International Modal Analysis Conference*. [S.l.]: American Institute of Aeronautics and Astronautics, 2009.

TAIPALUS, T.; SEPPäNEN, V.; PIRHONEN, M. Uncertainty in information system development: Causes, effects, and coping mechanisms. *Journal of Systems and Software*, v. 168, p. 110655, 2020. ISSN 0164-1212. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0164121220301205>.

TEGELTIJA, M.; OEHMEN, J.; KOZIN, I. Risk management challenges in large-scale energy pss. *Procedia CIRP*, v. 64, p. 169–174, 2017. ISSN 2212-8271. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2212827117301671>.

THOMÉ, A. M. T.; SCAVARDA, L. F.; SCAVARDA, A.; THOMÉ, F. E. S. de S. Similarities and contrasts of complexity, uncertainty, risks, and resilience in supply chains and temporary multi-organization projects. *International Journal of Project Management*, Elsevier, v. 34, n. 7, p. 1328–1346, 2016.

TRAVASSOS, G. H.; GUROV, D.; AMARAL, E. A. G. do. Introdução à engenharia de software experimental. COPPE/UFRJ, Rio de Janeiro, Relatório técnico: RT-ES-590/02, 2002.

TRAVASSOS, G. H.; SANTOS, P. S. M. dos; MIAN, P. G.; NETO, P. G. M.; BIOLCHINI, J. An Environment to Support Large Scale Experimentation in Software Engineering. In: *13th IEEE International Conference on Engineering of Complex Computer Systems (iceccs 2008)*. IEEE, 2008. p. 193–202. ISBN 978-0-7695-3139-7. Disponível em: <https://ieeexplore.ieee.org/document/4492892/>.

TUFTE, E. *The visual display of quantitative information*. [S.l.]: Cheshire: Graphic Press.–2001.–213 p, 2001.

WARD, S.; CHAPMAN, C. Transforming project risk management into project uncertainty management. *International journal of project management*, Elsevier, v. 21, n. 2, p. 97–105, 2003.

WHEELER, T. A. *Treatment of uncertainties associated with PRAs in risk-informed decision making (NUREG1855).* [S.l.], 2010.

WIDEMAN, R. M. *Project and program risk management: a guide to managing project risks and opportunities*. EUA: Newton Square:, 1992.

WOHLIN, C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: *Proceedings of the 18th international conference on evaluation and assessment in software engineering*. [S.l.: s.n.], 2014. p. 1–10.

YIN, R. *Case Study Research: Design and Methods*. SAGE Publications, 2013. ISBN 9781483322247. Disponível em: <https://books.google.com.br/books?id=OgyqBAAAQBAJ>.

ZADEH, L. A. Review of a mathematical theory of evidence. *AI magazine*, v. 5, n. 3, p. 81–81, 1984.

ZAPATA-RIVERA, J.-D.; NEUFELD, E.; GREER, J. Visualization of bayesian belief networks. In: PRESS. *IEEE Visualization 1999 Late Breaking Hot Topics Proceedings*. [S.l.], 1999. p. 85–88.

ZHANG, J.; YIN, J.; WANG, R. Basic framework and main methods of uncertainty quantification. *Mathematical Problems in Engineering*, Hindawi, v. 2020, 2020.

APPENDIX A – A GUIDELINE TO APPLY EULER'S MODEL

According to Ebert, Kuhrmann e Prikladnicki (2016), guidelines are lists of pieces of advice, best practice recommendations, and success factors based on empirical evidence. The guideline presented here was based on the empirical application of the Euler model in three teams formed to undertake agile software development projects.

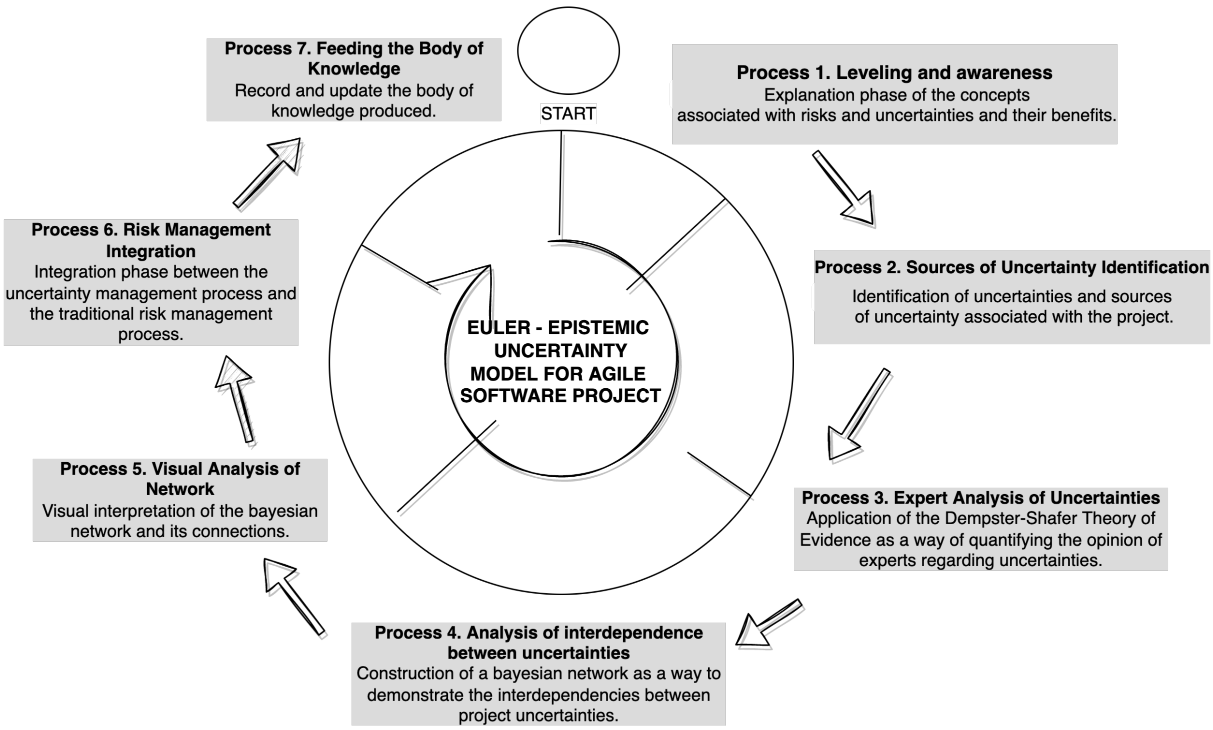
This guideline includes structural definitions of the Euler model, its sub-processes, tasks, goals, exit criteria, tools or techniques, and the definition of the roles for those applying the model. The search for tools or techniques was based on the PMBoK Guide (PMI, 2021) and some of the practices described in this guideline arose during the empirical application of the model, such as the need to hold a workshop or similar event to present and disseminate concepts related to risks and uncertainties in software projects.

It is expected that using this guideline will make applying the Euler model more fluid and transparent for the project manager and the team in charge of managing uncertainties in software projects.

For space and formatting reasons, the portrayal of the entire process of the Euler model has been divided here into three parts: Euler Tasks with Roles - Part 1, Part 2, and Part 3. A full version of the process can be found in the repository [1] created to store the artifacts produced by this thesis. Figure 37 gives an overview of the architecture of Euler's model and the roles needed. Figure 38, Figure 39, and Figure 40 show all the tasks of Euler's model with roles.

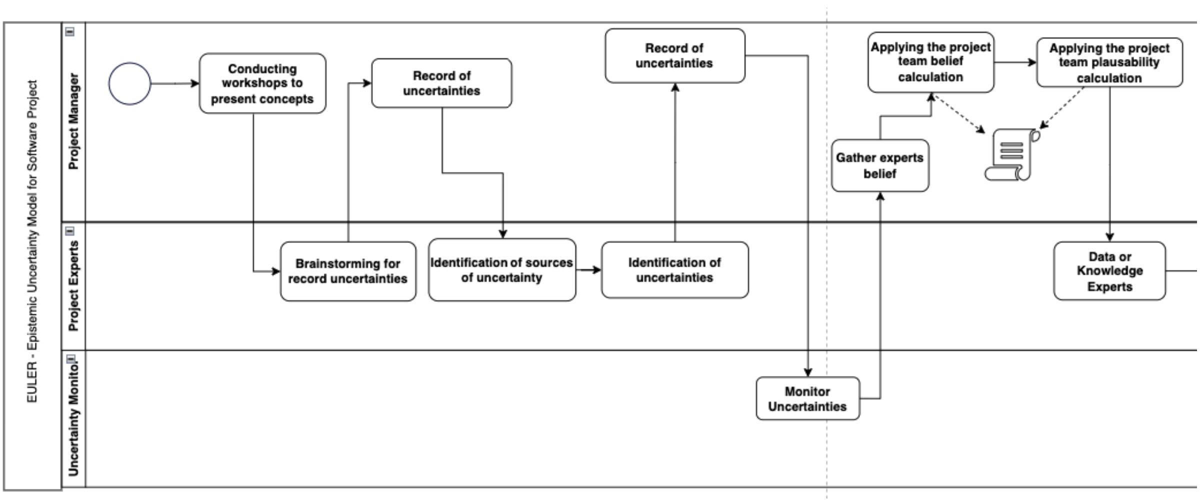
[1] https://bit.ly/QURepository

Figure 37 – Architecture Overview



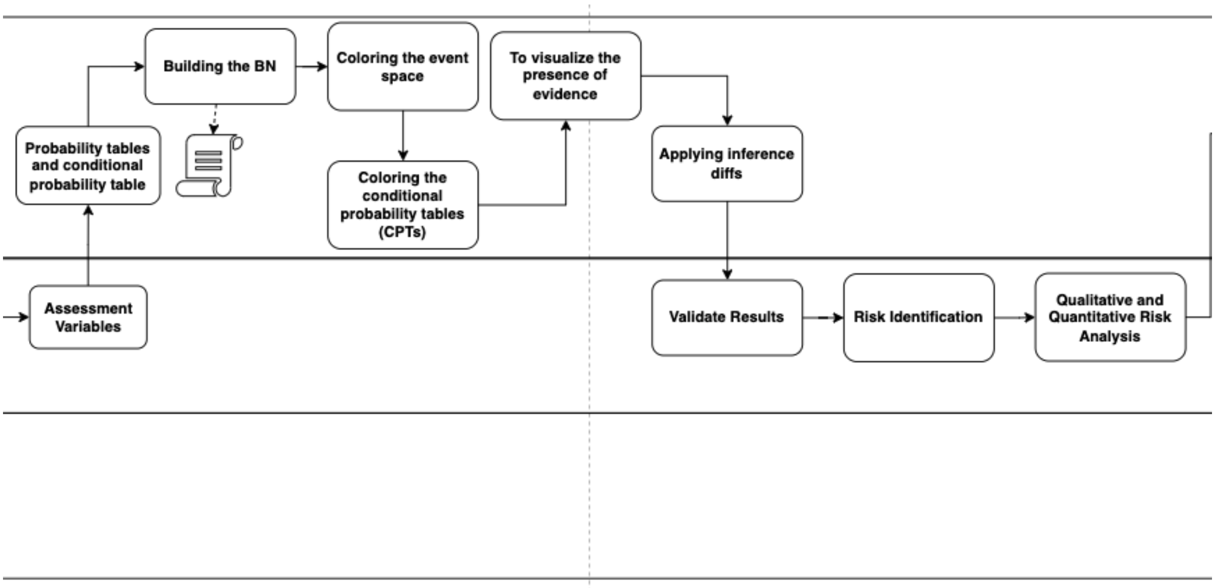
Source: Author

Figure 38 – Euler Tasks with Roles - Part 1



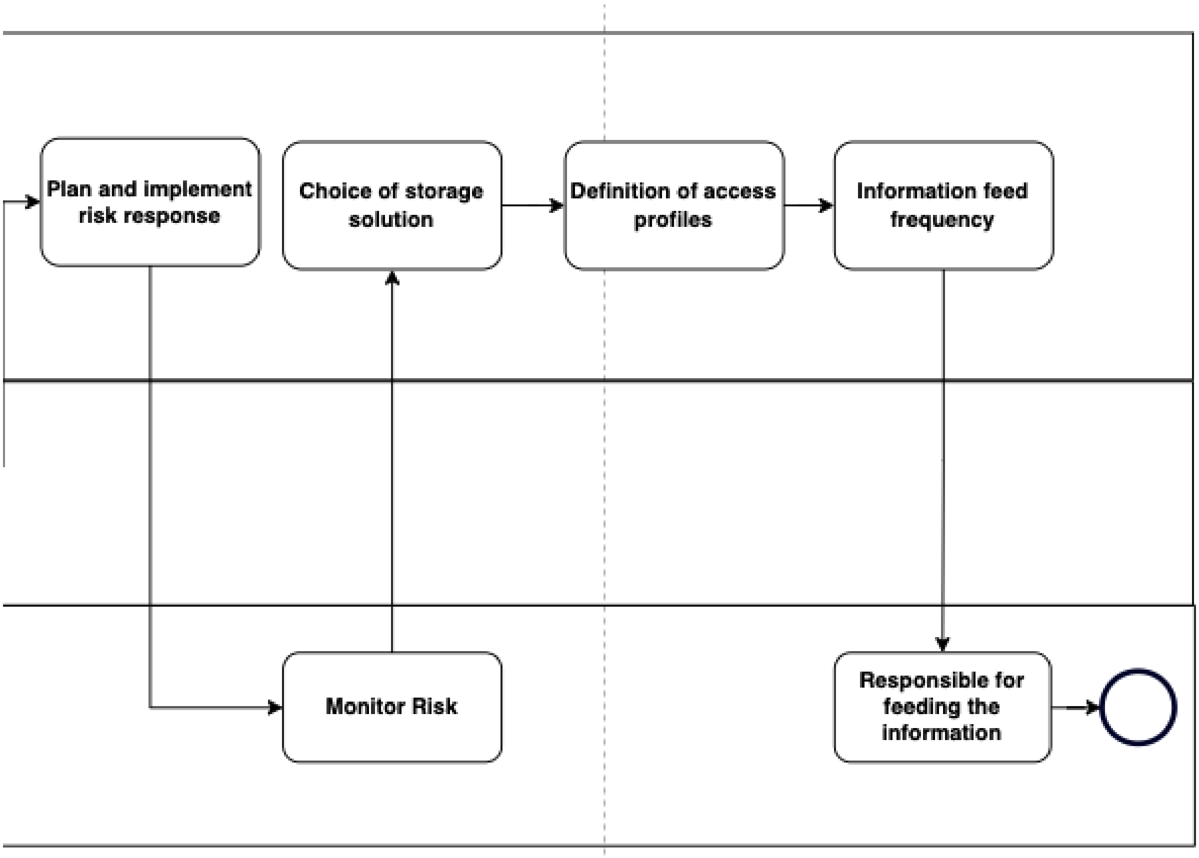
Source: Author

Figure 39 – Euler Tasks with Roles - Part 2



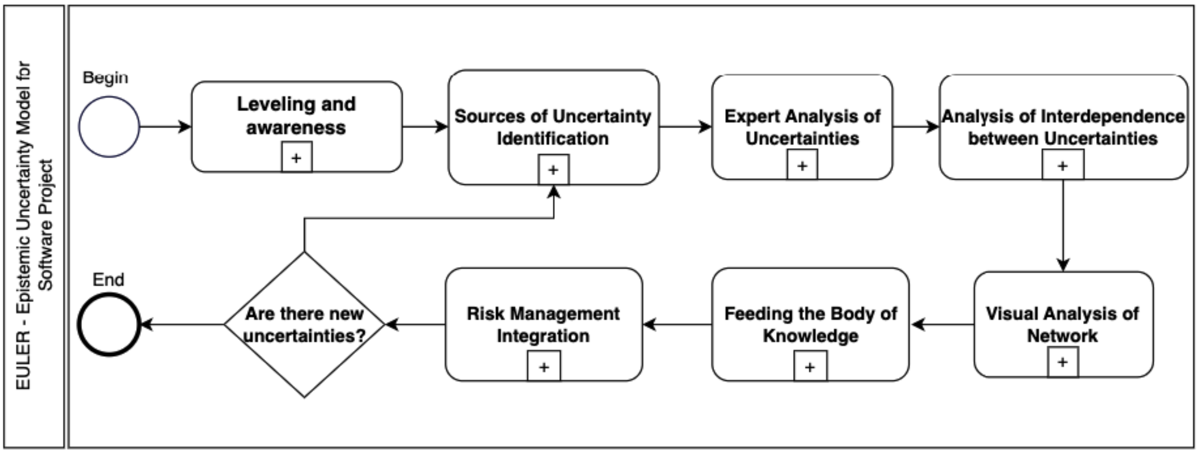
Source: Author

Figure 40 – Euler Tasks with Roles - Part 3



Source: Author

Figure 41 – Euler Processes Overview

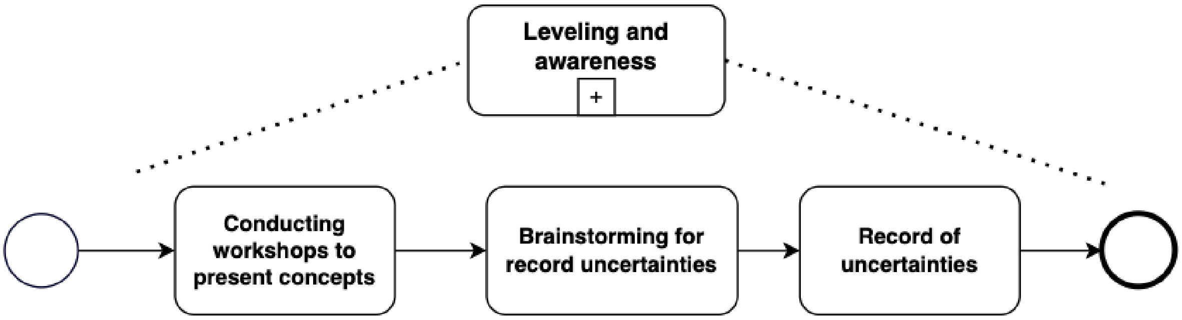


Source: Author

A.1 LEVELING AND AWARENESS

This process can be seen in Figure 42.

Figure 42 – Leveling and Awareness Process



Source: Author

Goal: To present practical examples of risks and uncertainties taken from the body of knowledge and the literature to reduce team members' doubts.

Exit Criteria: Concepts of risks and uncertainties understood.

Tools or Techniques: Workshops to align the concepts of risks and uncertainties..

Challenges: The concept of risk is not something trivial when talking about project management, even less the concept of uncertainty. Therefore, the presentation of context or scenarios is essential so that the concepts presented can be assimilated by the project team.

Tasks:

1. Conducting workshops to present concepts on risks and uncertainties;

2. Remember and record the uncertainties that occurred in previous projects and how the teams dealt with them;

3. Register the findings of the task in an appropriate repository within the project's infrastructure to produce a body of knowledge.

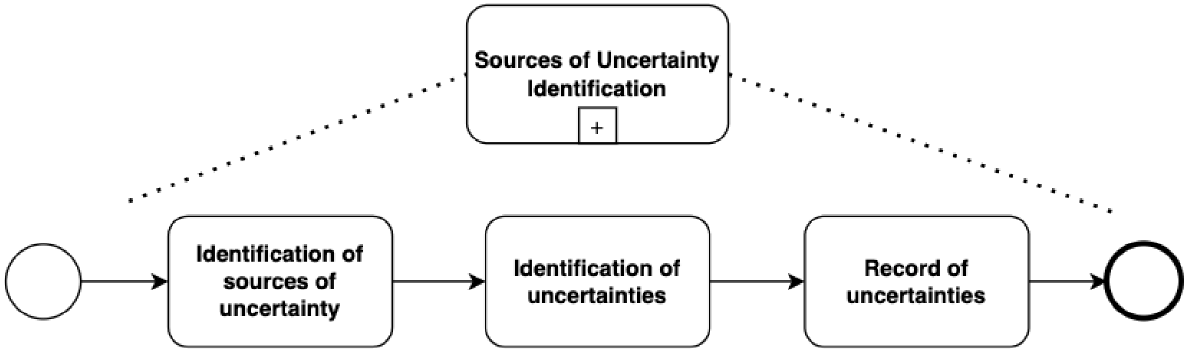
Outcomes: The leveling of everyone in the project team in the uncertainty management process concerning the concepts of risks and uncertainties.

A.2 SOURCE OF UNCERTAINTY IDENTIFICATION

This process can be seen in Figure 43.

Goal: The identification of uncertainties existing in the project.

Figure 43 – Sources of the Uncertainty Identification Process



Source: Author

Exit Criteria: List of identified uncertainties.

Tools or Techniques: Brainstorming sessions.

Challenges: Identification of uncertainties is one of the biggest challenges in this process because the concept of uncertainty is not intuitive. One way to combat this challenge is to look for sources of uncertainty and uncertainty mapped in previous or similar projects.

Tasks:

1. Identification of sources of uncertainty;

2. Identification of uncertainties;

3. Record of uncertainties;

Outcomes: A set of identified and registered uncertainties so that the entire project team is aware of the existing uncertainties.

A.3 EXPERT ANALYSIS OF UNCERTAINTIES

This process can be seen in Figure 44.

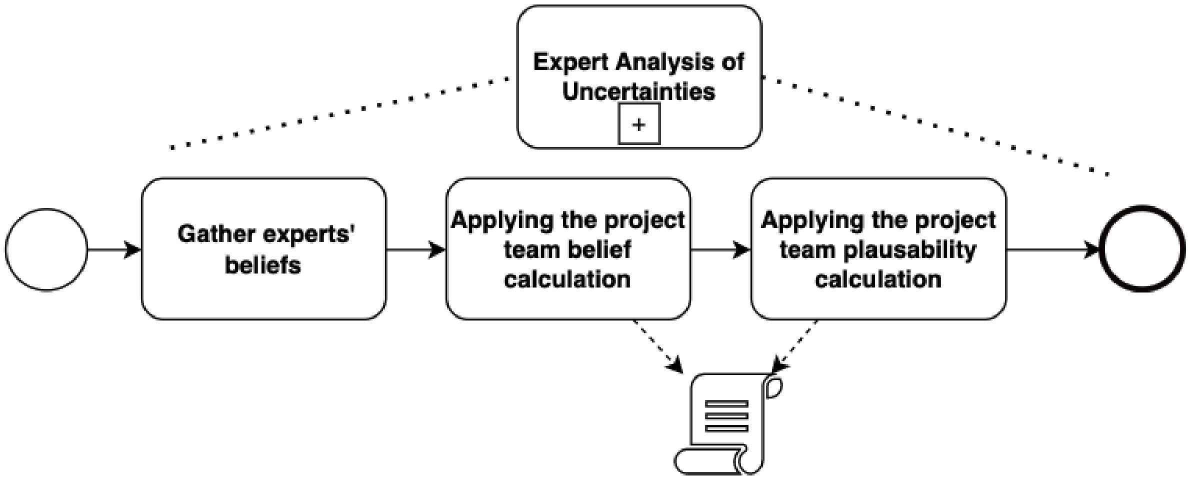
Goal: Collect the expert's opinions to quantify the belief that each of them has in whether the uncertainties identified in the previous step will or will not occur.

Exit Criteria: List of quantified uncertainties.

Tools or Techniques: Expert judgment and application of the Theory of Evidence (Dempster-Shafer Theory).

Challenges: Quantifying the belief in the occurrence of uncertainty based on an expert's evidence can be quite challenging in an uncertainty management process. One way to combat

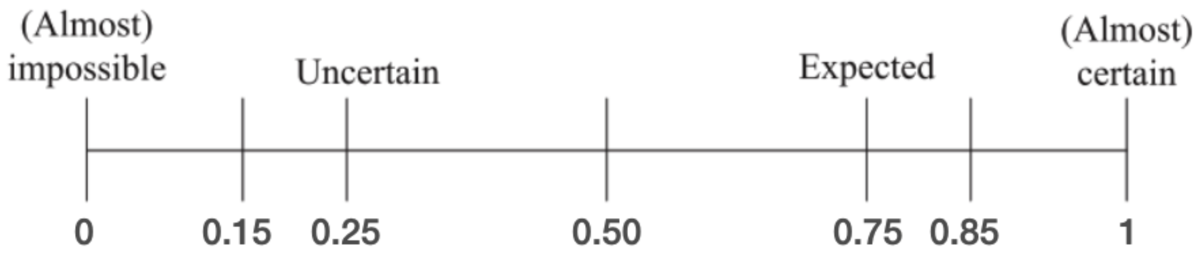
Figure 44 – Expert Analysis of Uncertainties



Source: Author

this challenge is to create a standard scale, as shown in Figure 45, that can guide experts during the assignment of mass activity.

Figure 45 – Probability Scale



Source: Adapted from Knochenhauer et al. (2013)

Tasks:

1. Each expert in the project team was asked to indicate which uncertainties of the uncertainty table presented in the previous step are more likely to occur according to their current knowledge, prior knowledge, and evidence of similar uncertainties that may have occurred and thus assign mass to these elements (uncertainties);

2. Apply the script produced to assist in the process of calculating the project team's level of belief in and the plausibility of the occurrence of uncertainties;

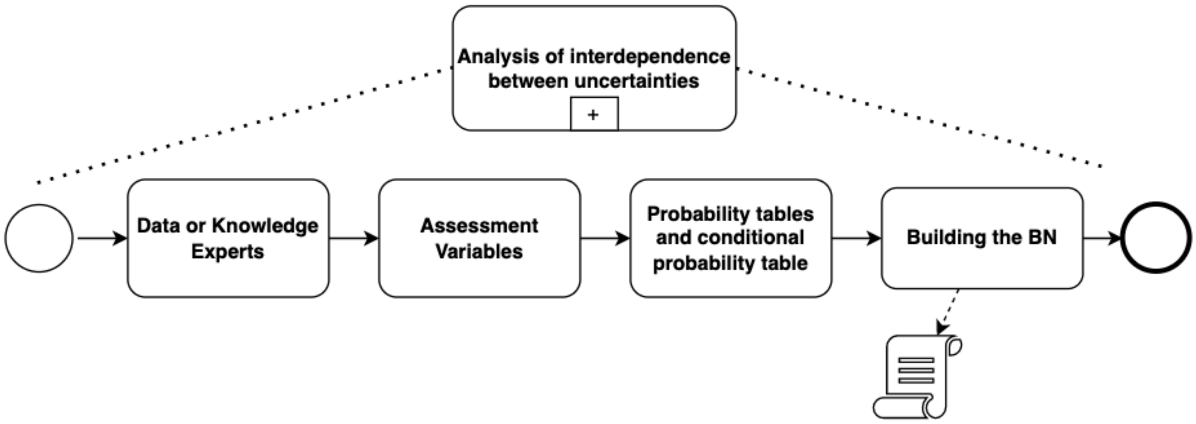
3. Monitor project uncertainties in order to monitor the evolution of activities and associated uncertainties, and if necessary,conduct a new assessment of project uncertainties using DST.

Outcomes: The calculation of the project team's degree of belief in the occurrence of particular uncertainties.

A.4 ANALYSIS OF INTERDEPENDENCE BETWEEN UNCERTAINTIES

This process can be seen in Figure 46.

Figure 46 – Analysis of Interdependence between Uncertainties



Source: Author

Goal: Construction of a Bayesian Network to demonstrate the interdependent relationships between uncertainties identified in the project.

Exit Criteria:Bayesian Network built.

Tools or Techniques: Cause and effect or Ishikawa diagrams, and Python script .

Challenges: Identifying corresponding nodes that impact each other is one of the biggest challenges in this process. One way to combat this challenge is to seek a systematic perception of the uncertainties existing in software projects.

Tasks:

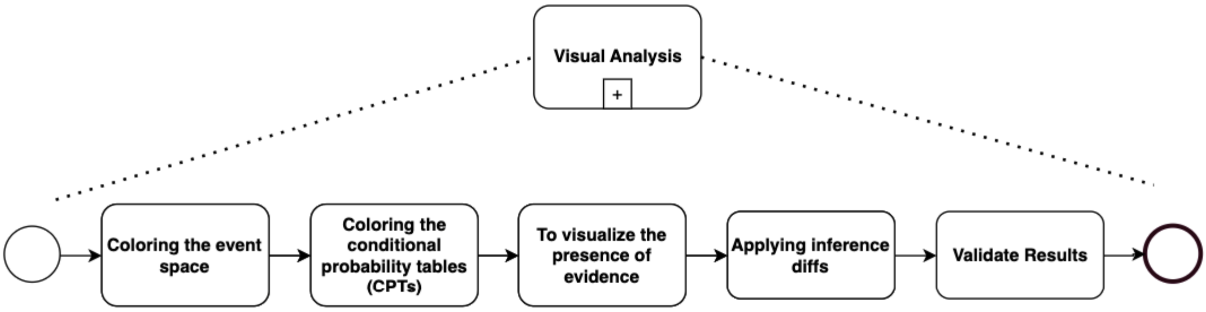
1. Construction of a Bayesian Network;
2. Assessment Variables;
3. Definition of a Conditional Probability Table (CPT);
4. Application of script to build the Bayesian Network.

Outcomes: Visual representation of a Bayesian Network with nodes and their relationships.

A.5 VISUAL ANALYSIS OF NETWORK

This process can be seen in Figure 47.

Figure 47 – Visual Analysis



Source: Author

Goal: Visualization of potential correlations or causal relationships between variables.

Exit Criteria: Bayesian Network built with visual highlights.

Tools or Techniques: Visual representation of a marginal distribution, and Python Script.

Challenges: When it comes to the visual aspect, several challenges arise. Which colors highlight higher values, and which colors are for lower values? What criteria to use to indicate the size of a node (a form of visual highlight indicating a greater degree of uncertainty)? Creating a standard color palette is recommended before starting the process.

Tasks:

1. Coloring the event space;
2. This step deals with coloring the CPT;
3. To visualize the presence of evidence;
4. Applying inference diffs.

Outcomes: A richer visual representation of the Bayesian Network.

A.6 RISK MANAGEMENT INTEGRATION

This process can be seen in Figure 48.

Goal: Since uncertainties are now known to a level called risks, applying traditional risk management processes completes the overall uncertainty management process.

Figure 48 – Risk Management

Source: Author

Exit Criteria: List of analyzed and prioritized risks with response plan.

Tools or Techniques: Checklists, interviews, and scenario analysis.

Challenges: Some of the challenges that can occur in the risk management process are associated with the perception of risk by each project management team member. One way to combat this challenge can be to apply the techniques of the Theory of Evidence (Dempster-Shafer Theory).

Tasks:

1. Risk Identification;

2. Qualitative and Quantitative Risk Analysis;

3. Plan and implement risk response;

4. Monitor Risk.

Outcomes: Production of a list of risks with corresponding response plans and the election of a project team member to monitor the risk.

A.7 FEEDING THE BODY OF KNOWLEDGE

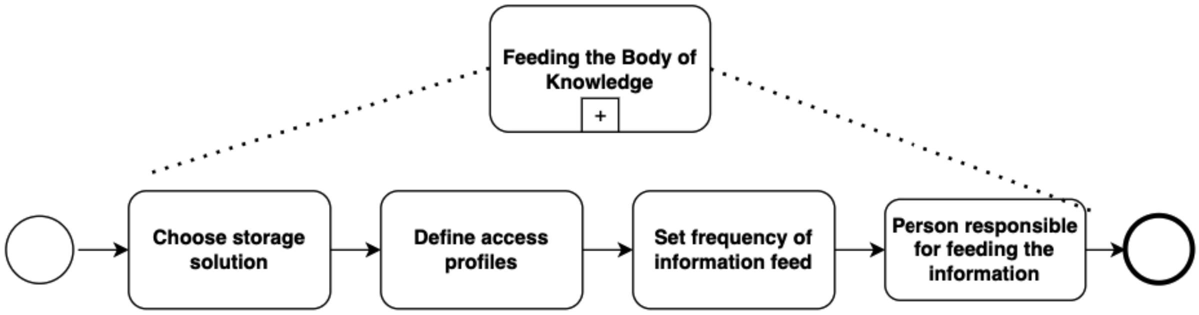
This process is shown in Figure 49.

Goal: Produce a body of knowledge that can serve as a central repository of knowledge and can serve as a point for querying and refining the results presented in the various cycles of execution and application of the Euler model.

Exit Criteria: Record of tacit knowledge produced at the end of each task.

Tasks:

Figure 49 – Feeding the Body of Knowledge subprocesses



Source: Author

1. Choose the solution for storing the knowledge produced during the application of the Euler model. This sub-process must consider that the tool to store the knowledge produced during the application of the model must store and return the information present in its body of knowledge on time. Timely means the query interface for previously recorded uncertainties and ways to resolve these uncertainties must be easily accessible;

2. Define access profiles to this body of knowledge. All company stakeholders should have access. However, some information can be considered strategic for the institution. Therefore, not all information should be accessible to all profiles;

3. Define how often this body of knowledge will be fed and who will be responsible for feeding the body of knowledge.

4. Indicate the team member responsible for curating the information produced. This person or profile should be responsible for keeping the knowledge base always up to date or making other stakeholders keep the knowledge base up to date.

Outcomes: Knowledge presentation produced during the process of Euler model application. This knowledge must serve as an input for future application cycles of the Euler model and future model applications in new projects.

APPENDIX B – SEMI-STRUCTURED INTERVIEW PROTOCOL - ACTION RESEARCH

B.1 INTRODUCTION

This document presents an overview of the research protocol used to conduct the semi-structured interviews as a form of data collection from an Action Research task. This Action Research was carried out within the State Prosecution Service (SPS), based in Paraiba, in Northeast of Brazil to investigate how agile teams approach the quantification of epistemic uncertainty in software development projects in a real project scenario.

B.2 OVERVIEW

The interview begins with a general presentation on the doctoral research subject and the objective of the Action Research. At this point, concepts about risk, uncertainties, and their differences and similarities were explained to obtain the best possible return from the interviewees. During this general presentation, the researcher explained to the participants that the data provided would remain anonymous. The first questions were related to the characterization and work experience of the interviewees.

B.3 QUESTIONS

This section describes the questions used in the semi-structured interview process. The first column presents an identifier for the question to reference it throughout the text better. The second column describes the issue. The third column presents the purpose of the question and, thus, what is intended to be collected from the interview participant. Finally, the last column presents the type of question (whether open or closed). As the definition of semi-structured interviews says, some questions may have undergone some change to collect better or better describe the issue. The researcher applied all of the questions in person or online, but in real-time (synchronously).

Table 25 – Question 1 - How many projects does the respondent currently work

ID	Question Description	Purpose	Response Type
Question 1	How many projects does the respondent currently work on in the SPS, and how many years of experience does the participant have in Software Engineering/Systems Development projects?	To characterize how many projects the participant works or has worked on and how many years of experience he has in software development projects. So, to characterize the participant's experience in Software Engineering. Experience that supports confidence in the answers provided.	Open

Table 26 – Question 2 - How is our current software development process

ID	Question Description	Purpose	Response Type
Question 2	How is our current software development process characterized?	To prospect how the participant understands the process used by the SPS if there is a process for software development. To check if there is transparency in the software development process by the project management.	Open

Table 27 – Question 3 - How to characterize risk or uncertainty management

ID	Question Description	Purpose	Response Type
Question 3	Within the current software development process, how to characterize risk or uncertainty management?	To prospect how the participant understands or sees the management of risks and uncertainties within the software development process. To identify the importance of the risk and uncertainty management process for the participant.	Open

Table 28 – Question 4 - Do you understand that our agile management process is tuned to deal with uncertainty

ID	Question Description	Purpose	Response Type
Question 4	Do you understand that our agile management process is tuned to deal with uncertainty? If not, what actions could we take to improve the current software development process to deal with the uncertainties inherent in agile projects?	To prospect empirical ways to deal with uncertainty management in software projects.	Open

Table 29 – Question 5 - What approaches could be used to prioritize project uncertainties

ID	Question Description	Purpose	Response Type
Question 5	What approaches could be used to prioritize project uncertainties?	To prospect empirical ways to prioritize uncertainties and thus to assist the decision-making process in managing project uncertainties.	Open

Table 30 – Question 6 - The participant wishes to add something more about the topic

ID	Question Description	Purpose	Response Type
Question 6	When it is announced that the interview is coming to an end, the participant is asked if he/ she wishes to add something more about the topic and to indicate someone who can contribute to the topic discussed	To discover additional ideas and additional participants with whom to conduct new interviews.	Open

B.4 CONFIDENTIALITY

For ethical considerations, each participant was asked to sign a consent form for the interview. This term addresses the use of information provided by interview participants and that this use must be used for academic purposes only.

No remuneration was made available to participants. It was explained to participants that the intangible benefits of their participation in this semi-structured interview can be: (i) making the results available in a synthetic format in emails as a way to help improve the system development process in the SPS; (ii) demonstrating willingness to assist a researcher and collecting work in the development of his doctoral project.

APPENDIX C – PROOF OF CONCEPT RESEARCH PROTOCOL

C.1 CONTACT LETTER SUBJECT

Subject: This is an invitation to participate in applying and evaluating a model for the Quantification of Epistemic Uncertainty in development projects that use agile methodologies.

Principal investigator: Dr. Hermano Perrelli de Moura, Professor Center of Informatics (CIn), Federal University of Pernambuco (UFPE), Recife, Brazil.

Co-investigator: Marcelo Luiz Monteiro Marinho, Professor at Federal Rural University of Pernambuco (UFRPE), Recife, Brazil.

Co-investigator: Jefferson Ferreira Barbosa, student at CIn/UFPE, pursuing a Doctoral degree in Computer Science.

C.1.1 Overview

This document presents an overview of the research protocol used to apply and evaluate the Euler model built during doctoral research at CIn-UFPE that aimed to investigate how agile teams approach the quantification of epistemic uncertainty in software development projects in scenarios of real projects.

C.1.2 Purpose

The objective of the study is to discuss some relevant aspects of the benefits of using a model to quantify epistemic uncertainty in software projects that use agile methodologies. This process is structured as a Proof of Concept Research (PoCR) according to the structure suggested by Elliott (2021).

We are specifically interested in investigating how agile teams approach the quantification of epistemic uncertainty in software development projects in scenarios of real projects.

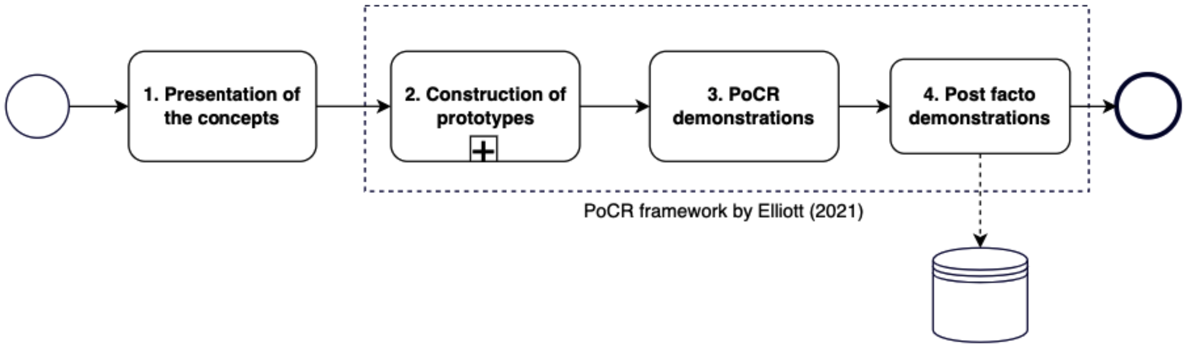
We intend to use the knowledge acquired during this study to improve the model built so that industry can use it to improve organizations' performance and the processes of uncertainty management in software projects.

We are inviting you because you have professional and/or academic experience compatible with the sample profile designed for this study. Respondents also qualify as representative agents of the phenomena under study and have the relevant expertise for this research.

C.1.3 Procedure of the Study

A diagram presented in Figure 50 was developed to understand better and present an overview of the application of this PoCR protocol.

Figure 50 – PoCR Overview



Source: Author

The application of this Proof of Concept Research protocol follows the specific procedures below:

1. Initially, a presentation of the concepts associated with this research was made, mainly characterizing the difference between the concepts of risks and uncertainties. For this characterization, references in the literature and examples of uncertainties collected during this research were used. This presentation lasts around 30 minutes and is carried out with the support of a set of slides created for this purpose[1]. The output of this process would be a better understanding, by the agile team, of the concepts around the management of risks and uncertainties in software projects that use agile practices;

2. The application process of this PoCR, following the framework proposed by Elliott (2021), dashed in the figure, is divided into three steps. The first is to construct prototypes, which, in this case, is the Euler model built for the quantification of epistemic uncertainty as shown in Figure 51 (in its first version). This stage resulted in a preliminary version of the model being built.

[1] https://bit.ly/slides$_{protocol_{PoCR}}$

3. The second step of this PoCR is when the Euler model is applied. At this moment, the researcher follows all the steps for applying the Euler model together with the target organization's agile software development team. The researcher's role is to serve as a facilitator on how to apply the model with the consequent identification and quantification of the uncertainties of the software development project. The output of this PoCR process, which corresponds to the output of applying the Euler model, is a set of uncertainties related to the target project, recorded, identified, and quantified for the purposes of prioritization and for predictability for better coordination of the software project by the manager of the software project;

4. Finally, the third step consists of post facto demonstrations. In this protocol, we will treat this as the moment to discuss the benefits and limitations of the application of the model as a way of evaluating the model and thus producing an improved version of it that we will call Euler version 2.0. In this step, the questionnaire described in the INTERVIEW PROTOCOL section below will be applied, and the result of the questionnaire will be used to evaluate the model. An improved version called Euler 2.0 will be built based on the results from the evaluation. The output of this last PoCR process is a set of suggestions and improvement points for the Euler model and a proof of the viability of the solution developed.

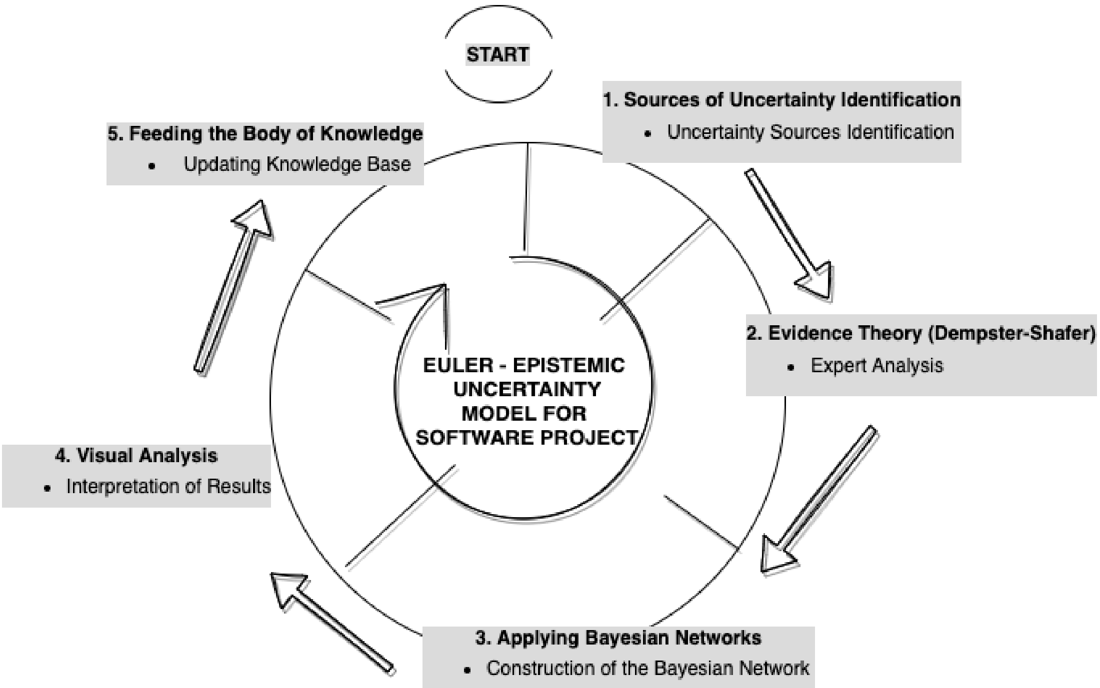
Both when the application of the protocol begins and during this process, the respondent is reminded that if he does not feel comfortable in terms of having the knowledge needed to answer any of the questions, he/she should not leave any questions unanswered, but should feel comfortable about providing approximate answers because for this research, it is more critical to receive a rough answer than no answer.

The application of this protocol was conducted with people from organizations that work or have direct contact with agile management of software development projects

C.1.4 Confidentiality

For ethical reasons, each participant was invited to sign a consent form to participate in this application process and in the evaluation of the Euler model. This term addresses the use of information provided by participants in this PoCR and that such use shall be for academic purposes only.

Figure 51 – Architecture Overview



Source: Author

C.1.5 Remuneration/Compensation

No compensation shall be provided for participation in this study

C.1.6 Benefits of the study

The participant will receive the following intangible benefits from participating in this study:

1. The summary of results can be made available in a synthetic format by e-mail as a way to help improve the project management process in the participating organizations;

2. Goodwill is generated by participating in a study to investigate the application of strategies that can help quantify and manage uncertainties in software development projects, thus contributing to development directives that industry can use.

C.1.7 Contact for information about the study

If you have any concerns or desire further information with respect to this study, you may contact the co-investigator, Jefferson Ferreira Barbosa (jfb2@cin.ufpe.br).

C.1.8 Contact for concerns about the rights of research subjects If

If you have any concerns about your treatment or rights as a research subject, you may contact the Secretary of the Board of Research (SEC-DPQ) in the UFPE Office for Research Affairs and Graduate Studies (PROPESQ) at +55 (81) 2126 7041 or dpq.propesq@ufpe.br.

C.1.9 Consent

Your participation in this study is entirely voluntary, and you may refuse to participate. Your signature on the consent form indicates that you have received a copy of this consent form for your records and future use.

Your signature indicates that you agree to participate in this study and agree with the study requirements. In some cases, the record of verbal consent, captured by the audio recording of the semi-structured interview, is sufficient evidence of voluntary acceptance to participate in the PoCR and equivalent to a signature.

C.2 INTERVIEW PROTOCOL

This section describes the questions used in the evaluation process of the Euler model which was conducted by means of a questionnaire and semi-structured interview.

The first column presents a question identifier to better reference it throughout the text. The second column presents the construct associated with the question. The third column presents the description of the question. The fourth column presents the objective (question purpose) and, thus, what it is intended to collect from the interviewee.

Finally, the last column presents the type of answer expected for the question. Questions can be applied in person or online in real-time (synchronous).

The questions applied in this protocol were based on Runeson et al. (2012) and questions found in the protocol developed by Marinho et al. (2015) to evaluate their approach to managing uncertainties.

C.2.1 Questions related to respondents and their demography

This set of questions aims to demographically qualify the respondent and thus show that he/she was qualified for the study and that the answers provided can be considered at the level of evaluation of the study presented.

This set of questions can also provide a better way to analyze the data produced in this PoCR without necessarily identifying individuals.

Table 31 – Question 1- How many years of experience do you have in software development projects

ID	Construct	Question Description	Purpose	Response Type
1	Demography	How many years of experience do you have in software development projects?	The objective is to characterize the participant's experience in Software Engineering. Experience that supports confidence in the answers provided.	- Up to 1 year - From 1 to 5 years - From 6 to 10 years - From 11 to 15 years - From 16 to 20 years - More than 20 years

Table 32 – Question 2- What is your current job position?

ID	Construct	Question Description	Purpose	Response Type
2	Demography	What is your current job position?	To identify the hierarchical level of the interviewee and thus to qualify their knowledge and level of confidence in the answers provided.	- Business owner - CEO - CIO - Executive - Consultant - Scrum Master - Product Owner - Tester

Table 33 – Question 3- What is your level of education (completed)

ID	Construct	Question Description	Purpose	Response Type
3	Demography	What is your level of education (completed)?	To identify the respondent's education level and thus to qualify their knowledge and confidence level in the answers provided.	- Undergraduate - Graduated - MBA (Lato Sensu) - Master - PhD - Postdoctoral

Table 34 – Question 4- In which of the following groups

ID	Construct	Question Description	Purpose	Response Type
4	Demography	In which of the following groups would the organization in which you work (or have worked recently) be best classified?	To prospect the type of organization in which the respondent and his/her projects are inserted as a way to characterize better the practices of project management, agile management, and management of uncertainties and, in this way, to seek insights into the similarities or differences of uncertainty management practices in these types of organizations.	- For-profit organization - Non-profit organization - Government - Academy

Table 35 – Question 5- How would you rate the size of the company where you work

ID	Construct	Question Description	Purpose	Response Type
5	Demography	How would you rate the size of the company where you work? Use as reference the table from SEBRAE.	To prospect the size of the organization in which the respondent and his/her projects are inserted to better characterize project management practices.	- Micro - Small - Medium - Large

C.2.2 Questions related to organization, project management and uncertainty management practices

This set of questions aims to qualify the project and project management practices such as uncertainty management with which the respondent is associated and thus to show that he or she qualifies for this research since he/she participates in projects that use agile practices.

Table 36 – Question 6- The organization where I work (or I worked) adopts practices to manage uncertainties

ID	Construct	Question Description	Purpose	Response Type
6	Uncertainty Management	The organization where I work (or I worked) adopts practices to manage uncertainties.	To prospect the respondent's level of awareness regarding whether uncertainty management practices applied to his/her organization's projects do or do not exist.	- Strongly Disagree 1 - Disagree 2 - Indifferent 3 - Agree 4 - Strongly Agree 5

Table 37 – Question 7- The organization where I work is interested in adopting an approach to managing uncertainty

ID	Construct	Question Description	Purpose	Response Type
7	Uncertainty Management	The organization where I work is interested in adopting an approach to managing uncertainty. You as project manager would be interested in adopting the approach.	To prospect the respondent's level of awareness regarding whether or not the organization's management is interested in adopting uncertainty management practices applied in their organization's projects.	- Strongly Disagree 1 - Disagree 2 - Indifferent 3 - Agree 4 - Strongly Agree 5

Table 38 – Question 8- Do you have any suggestions regarding the previous question

ID	Construct	Question Description	Purpose	Response Type
8	Uncertainty Management	Do you have any suggestions regarding the previous question?	To discover additional information that can enrich the answer to the previous question.	Open-ended

C.2.3 Questions related to the evaluation of the Euler model

This set of questions aims to evaluate the Euler model for quantifying epistemic uncertainty, concerning its relevance to assisting practitioners and academics in project management, specifically in managing uncertainty in projects that apply agile practices.

Table 39 – Question 9- The development of a model is an extremely necessary contribution to industry and the academy

ID	Construct	Question Description	Purpose	Response Type
9	Uncertainty Management	The development of a model to quantify epistemic uncertainties in software project to help people and organizations to apply strategies to improve project success is an extremely necessary contribution to industry and the academy.	To prospect the respondent's level of awareness concerning the need for a model to quantify epistemic uncertainty in software projects.	- Strongly Disagree 1 - Disagree 2 - Indifferent 3 - Agree 4 - Strongly Agree 5

Table 40 – Question 10- Do you have any suggestions regarding the previous question

ID	Construct	Question Description	Purpose	Response Type
10	Uncertainty Management	Do you have any suggestions regarding the previous question?	To discover additional information that can enrich the answer to the previous question.	Open-ended

Table 41 – Question 11- a model for quantifying epistemic uncertainty is an essential contribution to industry and academia

ID	Construct	Question Description	Response Type	Criteria
11	Model Evaluation	Agile project management, by its nature, is prone to dealing with uncertainty, so a model for quantifying epistemic uncertainty is an essential contribution to industry and academia.	To prospect the respondent's level of awareness concerning need for a model to quantify epistemic uncertainty in agile software projects.	- Strongly Disagree 1 - Disagree 2 - Indifferent 3 - Agree 4 - Strongly Agree 5

Table 42 – Question 12- Do you have any suggestions regarding the previous question

ID	Construct	Question Description	Purpose	Response Type
12	Model Evaluation	Do you have any suggestions regarding the previous question?	To discover additional information that can enrich the answer to the previous question.	Open-ended

Table 43 – Question 13- Using a model to quantify epistemic uncertainties in software projects can bring benefits

ID	Construct	Question Description	Response Type	Criteria
13	Model Evaluation	You understand that using a model to quantify epistemic uncertainties in software projects can bring benefits such as greater predictability of uncertainties associated with the various sources of uncertainty inherent in projects that use agile project management practices and that this can lead to better project coordination.	To prospect the respondent's level of awareness of the benefits associated with applying a model to quantify epistemic uncertainty in agile projects.	- Strongly Disagree 1 - Disagree 2 - Indifferent 3 - Agree 4 - Strongly Agree 5

Table 44 – Question 14- Do you have any suggestions regarding the previous question

ID	Construct	Question Description	Purpose	Response Type
14	Model Evaluation	Do you have any suggestions regarding the previous question?	To discover additional information that can enrich the answer to the previous question.	Open-ended

Table 45 – Question 15- The nomenclature of the phases of the Euler model is clear and follows a logical sequence

ID	Construct	Question Description	Response Type	Criteria
15	Model Evaluation	The nomenclature of the phases of the Euler model is clear and follows a logical sequence.	To prospect the interviewee's level of awareness concerning the nomenclature of the Euler model's phases.	- Strongly Disagree 1 - Disagree 2 - Indifferent 3 - Agree 4 - Strongly Agree 5

Table 46 – Question 16- Do you have any suggestions regarding the previous question

ID	Construct	Question Description	Purpose	Response Type
16	Model Evaluation	Do you have any suggestions regarding the previous question?	To discover additional information that can enrich the answer to the previous question.	Open-ended

Table 47 – Question 17- The output of each phase of the Euler model is clear, as is how this output feeds the next phase

ID	Construct	Question Description	Response Type	Criteria
17	Model Evaluation	The output of each phase of the Euler model is clear, as is how this output feeds the next phase.	To prospect the interviewee's level of awareness concerning the outputs of the Euler model's phases.	- Strongly Disagree 1 - Disagree 2 - Indifferent 3 - Agree 4 - Strongly Agree 5

Table 48 – Question 18- Do you have any suggestions regarding the previous question

ID	Construct	Question Description	Purpose	Response Type
18	Model Evaluation	Do you have any suggestions regarding the previous question?	To discover additional information that can enrich the answer to the previous question.	Open-ended

Table 49 – Question 19- What suggestions for the model would you make to improve the management

ID	Construct	Question Description	Purpose	Response Type
19	Model Evaluation	Thinking as a team member, what suggestions for the model would you make to improve the management of epistemic uncertainties in the project(s) you are currently working on?	To identify improvement points for the model from the team's point of view.	Open-ended

Table 50 – Question 20- Do you have any suggestions for the progress of the research

ID	Construct	Question Description	Purpose	Response Type
20	Model Evaluation	Do you have any suggestions for the progress of the research?	To discover additional ideas and additional participants for conducting new interviews.	Open-ended

C.3 DATA ANALYSIS

This section defines how the qualitative data will be analyzed after completing the open-ended question form.

Below, we have the thematic analysis process phases applied to open questions. We also can find a coding guide used in conjunction with the Atlas.it tool. This coding guide can be used as a form of alignment between several researchers who aim to analyze the same data set within this research. More than one researcher's involvement in analyzing the collected material can minimize threats to the validity of the research, such as inconsistencies in the coding process. However, we emphasize that this analysis was carried out by a single researcher with the validation of this protocol performed by three specialists, each with a Ph.D. in Software Engineering and extensive experience in the area.

Transcripts should be coded using the phases of thematic coding analysis (ROBSON; MC-CARTAN, 2016):

1. Familiarizing yourself with your data.
2. Generating initial codes.
3. Identifying themes.
4. Constructing thematic networks.
5. Integration and interpretation

Table 51 presents a set of codes that must be used during the thematic analysis and coding process. These codes were built with a focus on possible answers to research questions.

Table 51 – Codes for Research Questionnaire

ID Code	Code	Description
1	P — Problems, challenges	It encodes a transcript excerpt that identifies existing problems in the model that may make it impossible to apply it in real environments of software projects with agile practices.
2	I — Improvement suggestion	It encodes a transcription excerpt that identifies points of improvement in the Euler model, such as a suggestion for name changes or adjustments to parts of the process of applying the model.
3	N — New processes or replace existing one	Encodes a transcription excerpt that identifies the need to apply new processes and replace existing processes in the model.

C.4 REFERENCES

1. Barbosa, J. F., Marinho, M. L. M., de Moura, H. P. (2021). Em Direção a um Modelo para Quantificação da Incerteza Epistêmica em Projetos de Software: uma Pesquisa-Ação. Revista Ibérica de Sistemas e Tecnologias de Informação, (44), 67-83.

2. Elliott, S. (2021). Proof of Concept Research. Philosophy of Science, 88(2), 258-280

3. Marinho, M., Sampaio, S., Luna, A., Lima, T., Moura, H. (2015). Dealing with uncertainties in software project management. In 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, 889–894

4. Runeson, P., Host, M., Rainer, A., Regnell, B. (2012). Case study research in software engineering: Guidelines and examples. John Wiley Sons.

APPENDIX D – PYTHON SCRIPTS - EULER MODEL

```python

"""
Script for performing computations in the Dempster-Shafer theory.
@author: Jefferson Ferreira Barbosa
@email: jfb2@cin.ufpe.br, jeffersonjpa@gmail.com
@licence:
"""

# Command for installing the Pyds library
# !git clone https://github.com/reineking/pyds
# !cd pyds; python setup.py install

# Load library
from pyds import MassFunction, powerset

# Command for creating mass assignment by two experts.
# Describe method parameters.

# A Dempster-Shafer mass function (basic probability assignment) based on a
    dictionary.
exp_m1 = MassFunction({'U1':0.6,
                       'U2':0.7,
                       'U3':0.3,
                       'U4':0.8,
                       'U5':0.5,
                       'U6':0.3,
                       'U7':0.4,
                       })

exp_m2 = MassFunction({'U1':0.5,
                       'U2':0.9,
                       'U3':0.4,
                       'U4':0.5,
                       'U5':0.7,
                       'U6':0.5,
                       'U7':0.5,
                       })

# Command to visualize the mass assignment made by Expert 1
exp_m1

#
print("exp_m1:", exp_m1)

```

```
#
print("exp_m1: bpa of {'U1','U2'}=", exp_m1['U1U2'])

#
print("exp_m1: belief of {'U1','U2'}=", exp_m1.bel('U1U2'))

# Plausibility
print("exp_m1: plausibility of {'U1','U1'}=", exp_m1.pl('U1U2'))

# Commonality
print("exp_m1: commonality of {'U1','U2'}=", exp_m1.q('U1U2'))



"""
Script for performing computations in Bayesian Networks.
@author: Jefferson Ferreira Barbosa
@email: jfb2@cin.ufpe.br, jeffersonjpa@gmail.com
@licence:
"""

# !pip install bnlearn pgmpy

# Load library
from pgmpy.factors.discrete import TabularCPD
import bnlearn as bn

print(bn.__version__)
```