TÉCNICO LISBOA

Delphi Method: A comparative analysis of existing Web-Delphi platforms

António Francisco Nascimento e Oliveira Semedo Fidalgo

Dissertation to obtain the Master of Science Degree in

Industrial Engineering and Management

Supervisors: Prof. Teresa Sofia Cipriano Gonçalves Rodrigues
Prof. Ana Catarina Lopes Vieira Godinho de Matos

Júri

Chairperson: Prof. Carlos António Bana e Costa
Supervisor: Prof. Ana Catarina Lopes Vieira Godinho de Matos
Members of the committee: Prof. Klára Dimitrovová

November 2021

Acknowledgments

Quero dedicar este trabalho em primeiro lugar à minha Avó Maria do Rosário que infelizmente já não está entre nós. Quero agradecer por todo o apoio que me deu em 24 anos da minha vida, pelos valores que me transmitiu e por tudo o que me ensinou. Quero agradecer aos meus pais também por todo apoio e ajuda que me deram para tornar este momento possível, não só financeiramente como também emocionalmente, e também aos meus irmãos pelo apoio incondicional.

Quero agradecer às minhas duas orientadoras de tese, professora Teresa Rodrigues e professora Ana Vieira, por todo o apoio prestado, por todo o conhecimento que me transmitiram e também pela sua disponibilidade ao longo destes meses. Foram um pilar fundamental na construção da minha dissertação e uma enorme ajuda. Foi um enorme prazer conhecer e trabalhar com elas e espero que continuem com o excelente trabalho que têm vindo a exercer.

Um obrigado também a todos os restantes intervenientes que tornaram esta dissertação possível, não só os entrevistados como também todos os autores que citei ao longo do meu trabalho.

Quero também agradecer ao Instituto Superior Técnico, por todos os valores e conhecimento que me transmitiu, pelas amizades que me proporcionou e por todos os bons e maus momentos que atravessei ao longo do meu percurso académico.

Por último, quero agradecer a todos os meus amigos, colegas e em especial à minha namorada por terem sido uma força e motivação para à conclusão de uma das etapas mais importantes na minha vida.

Declaration

I declare that this document is an original work of my own authorship and that it fullfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Declaração

Declaro que o presente documento é um trabalho original da minha autoria e que cumpre todos os requisitos do Código de Conduta e Boas Práticas da Universidade de Lisboa.

Abstract

In the last 20 years, there has been a growth of web platforms designed exclusively to implement the Delphi Method. Companies have developed these platforms benefiting from the evolution around web surveys, whose technological features helped overcome limitations inherent to the Delphi method's implementation. The Web-Delphi platforms available in the market differ among them regarding the features available for users, influencing how a Delphi process is implemented. Given the lack of tools to compare these platforms, the process of choosing one to implement a Delphi study becomes complex. Thus, this dissertation aims to perform a comparative analysis on the operation of these platforms and explore how their features can improve the implementation of the method. To meet this objective, the WDP Features Framework was created to allow for a comparative analysis between the different platforms available. The Framework is divided into different areas of concern, and each one has a list of features. Based on the Framework, a multi-criteria decision analysis model was also proposed to evaluate these platforms and help users choose an adequate platform. It was possible to understand which features were considered more important by the different companies and which are less used. Furthermore, it was possible to observe the development of each platform in each area of concern, considering the number of additional features present in them. With this work, it is now possible to provide a future user with objective knowledge of how these platforms work and thus facilitate a future decision to choose one.

Keywords: Delphi Method; Web-Delphi; Web-Delphi Platforms; MACBETH; MCDA; Hierarchical Additive Model.

Resumo

Nos últimos 20 anos, tem havido um crescimento de plataformas web concebidas exclusivamente para implementar o Método Delphi. As empresas desenvolveram estas plataformas beneficiando da evolução em torno das plataformas de inquérito web, cujas características tecnológicas ajudaram a ultrapassar as limitações do Método Delphi. As plataformas Web-Delphi disponíveis no mercado diferem entre si quanto às funcionalidades, influenciando a forma como um processo Delphi é implementado. Dada a falta de ferramentas para comparar estas plataformas, o processo de escolha de uma para implementar um estudo Delphi torna-se complexo. Assim, esta dissertação visa realizar uma análise comparativa sobre o funcionamento destas plataformas e explorar como as suas características podem melhorar a implementação do método. Para cumprir este objectivo, foi criado o WDP Features Framework para permitir uma análise comparativa entre as diferentes plataformas disponíveis. A Framework está dividida em diferentes áreas de preocupação, e cada uma delas tem uma lista de características. Com base na Framework, foi também proposto um modelo de análise de decisão multi-critério para avaliar estas plataformas e ajudar os utilizadores a escolher uma plataforma adequada. Foi possível compreender que características foram consideradas mais importantes pelas diferentes empresas e quais são menos utilizadas. Além disso, foi possível observar o desenvolvimento de cada plataforma em cada área de preocupação, considerando o número de características adicionais presentes nas mesmas. Com este trabalho, é agora possível proporcionar a um futuro utilizador um conhecimento objectivo do funcionamento destas plataformas e facilitar assim uma decisão futura de escolha de uma.

Palavras-chave: Método Delphi; Web-Delphi; Plataformas Web-Delphi; MACBETH; MCDA; Modelo Aditivo Hierárquico.

Table of Contents

List of Abbreviations

AHP Analytic hierarchy process

DDA Delphi Decision Aid

DM Decision Maker

DM Delphi Manager

EL ExpertLens

Evamix Evaluation matrix

GFIS Global Futures Intelligence System

LD Live Delphi

MACBETH Measuring Attractiveness by a Categorical Based Evaluation Technique

MCDA Multi-criteria Decision Analysis

SD Standard Deviation

TC Tech Cast

TD Traditional Delphi

WDP Web-Delphi Platforms

List of Figures

List of Tables

1. Introduction

This section consists of an overview of the context that led to the realization of this dissertation; it includes the characterization of the problem, the definition of the objectives, and the structure of the document.

This document focuses on studying Web-Delphi platforms, i. e., platforms that were created exclusively for the implementation of online Delphi processes. The present work aims to understand the evolution and development of these platforms, which in turn have made the practice of the Delphi Method easier and more efficient, not only from the perspective of the process moderator but also from the participants' perspective.

1.1. Dissertation Contextualization

Described as a technique created to "obtain the most reliable opinion consensus of a group of experts by subjecting them to a series of questionnaires in-depth interspersed with controlled opinion feedback" (Dalkey & Helmer, 1963, p.7), Delphi Method is a social communication technique that helps structure group communication and at the same time helps inhibit problems associated with traditional group decision-making processes. The distinctive characteristics of the method: *anonymity*, *multiple iterations*, *controlled feedback,* and *statistic group response* (Skulmoski, Hartman & Krahn, 2007) are the reason why Delphi is so worthy and distinct among other face-to-face decision-making techniques and methods such as Focus Group Technique and the Nominal Group Technique (Andrews & Allen, 2002; Day & Bobeva, 2005). These characteristics allow a reduction of the incidence of unwanted psychological effects in the participants (dominant people, inhibition, constraint, etc.), offer better consideration due to repetition, allow the selection and provision of pertinent feedback to participants, allow the delivery of statistical results that provide a better summary of the group opinion, offer simple execution, and flexible methodology (Linstone & Turoff, 1975; Landeta, 2006). Over the years, several studies and variants of the original method have been developed (McKenna, 1994), to increase the efficiency of its survey procedure and also to overcome its weaknesses and defects, such as participants' panel instability and time consumption (Gordon & Pease, 2006). It started as a paper-based method but given the digital revolution between 1950 and 1970 and the introduction of the World Wide Web concept in 1990, it was possible to start thinking on a new perspective of the Delphi method: its implementation in digital format.

At that time, only traditional mail had been used to tackle participants' distance, but in 1997, the first online Delphi study was conducted by E-mail and was successfully done (Bourque et al., 2002). This study served as an example and as a light that the implementation of the method

using the web was possible. It was then possible to think further and use web-surveys platforms to conduct Delphi studies. Currently, there are plenty of web-survey templates such as Google forms (www.docs.google.com/forms), SurveyMonkey (www.surveymonkey.com), SoGoSurvey (www.sogosurvey.com), SurveyGizmo (www.surveygizmo.com), and many others. Unfortunately, they are not as effective when it comes to implementing the Delphi method, since none of these tools were specifically designed to implement a Delphi, it becomes difficult to conduct a process and respect all the key characteristics associated with the method. Given its many advantages, over the last 20 years, several online platforms with specific web-surveys for Delphi have been designed and created to mitigate its limitations and maximize its advantages, by making the method more flexible and accessible without losing its original essence and validity. For example, Welphi (www.welphi.com) and Mesydel (www.mesydel.com) are two cases of platforms that were designed specifically to implement Delphi processes. At the same time, the use of these platforms with specific web-surveys designed for Delphi not only tackles participants distance, but also turned to be quicker and cheaper than classic paper and pencil processes, thus more efficient, economic, and sustainable (Fricker & Schonlau, 2002; Donohoe & Needham, 2009). From now on, these software-based platforms with specific web-surveys for Delphi will be mentioned as *Web-Delphi platforms (WDP)*.

1.1.1. Problem Definition

Currently, there are several WDP available online, which in turn raises the question: “Which one is more suitable to implement a Delphi process?”. Although they were all designed with the same purpose, we cannot ignore the differences between them. Such differences can influence how Delphi is implemented and lead to different scenarios, such as lower dropout rates, simpler and more intuitive processes, better security and protection of user data, and other consequences. After performing a literature review focused on this subject, it was concluded that there is a significant lack of articles that address how WDPs work and their advantages in implementing Delphi processes.

To address the aforementioned issue, the present dissertation proposes and applies a methodological framework capable of analyzing the presence of features that facilitate and optimize the implementation of the Delphi Method on these platforms. In this way, it will then be possible to perform a clear and objective comparative analysis between the different WDPs available. However, the application of this framework is not enough to decide which platform is better and which criteria are more important to consider when choosing a WDP. Therefore, the respective framework will also be used to propose the construction of a multi-criteria model according to the MACBETH approach, which can evaluate and score the performance of the respective WDPs.

1.2. Dissertation Objective

The main objective is to perform a comparative analysis between the different available WDPs and explore how their features can advance the use of the Delphi method. These platforms have provided solutions to many Delphi method issues and limitations, but also left others to be resolved and even raised new ones. To achieve this goal the following objectives were set:

1. Understand the impact of the transition from paper-based Delphi to web format, as well as its advantages and disadvantages.
2. Understand the state of the art of Web-Delphi platforms and identify the Delphi method issues that were tackled by these platforms and the ones that are still open to research.
3. Construct a framework capable of analyzing and comparing Web-Delphi platforms.
4. Apply the framework to compare these platforms and identify guidelines for improving the Web-Delphi processes and respective platforms.
5. Understand the importance of multi-criteria models in decision-making processes and their advantages with a focus on the MACBETH approach.
6. Based on the framework outlined in objective 3, propose a multi-criteria decision analysis model to evaluate WDPs and consequently help decision-makers choosing a Web-Delphi platform to implement Web-Delphi processes.

1.3. Dissertation Structure

This dissertation project is divided into seven sections. The first is the present section which consists of the contextualization of the dissertation, the motivation, and the objectives to be achieved in its realization. The second section consists of a literature review where research was done on the importance of having participatory processes with structured communication with the main focus on the Delphi method. A detailed description of the Delphi method was made from its origin to its current state as well as an analysis of the impact of its transition from paper-based to the web with an emphasis on its benefits and drawbacks. The research was also done on the use of multi-criteria models with a particular focus on the MACBETH approach. The third section describes the methodology that will be applied for the development of the dissertation. The fourth section consists of the application of the proposed methodology, and the fifth section presents its results. The sixth section discusses the methodology used and the results obtained. The last section summarizes the most relevant information that was gathered with the realization of the dissertation and also some proposals for future work.

2. Literature Review

2.1. Importance of Participatory Decision Making

Concerns of trust, equity, and representativeness have been raised about the outcomes of decision processes that depend exclusively on formal evaluation techniques and where analysts are in total control of decision support. To get around these issues and guarantee the active engagement of stakeholders, the need to develop deliberative processes such as citizen juries, consensus conferences, or discussion groups has increased significantly (Antunes, Santos & Videira, 2006). These deliberative processes are usually examples of participatory models for decision making.

Before we proceed, it is necessary to establish the difference between two approaches - Participative Groups and Conventional Groups. The main difference between these two is that in Participative Groups, all the participants participate in the decision making, unlike Conventional Groups, where the individuals with the fastest thinking and greater ease of communication are the ones who have the most weight and influence in the final decision (Kaner, 2014).

Heisler and others stated that participatory decision-making is fundamental to achieve better results and social acceptance of decisions. Also, to achieve a good modelling practice, it is recommended to use a structured decision-making framework where the participants' engagement is mainly accomplished by good communication and trust existence. (Heisler et al., 2002). In addition to facilitating the implementation of decisions, participatory models also help facilitate communication between the individuals involved in the process (Hashim, Alam & Siraj, 2009). To ensure a feasible implementation of these models, these individuals involved in the process must be experts in the subject matter. However, the use of experts has to be well deliberated.

Participants can often be unaware of the subjective influences to which they may be subject. Whether they can win or lose with a decision, their values and state of mind are factors that can influence estimates. Therefore, it is important to ensure that decision-makers do not depend directly on participants to reach a decision, but instead use an appropriate decision support tool (Burgman et al., 2011) to deal with the participants' judgments.

To guarantee a good modelling practice, it’s important to guarantee the continuous involvement of the originally engaged people in the beginning. Addison highlighted the importance of stakeholders, experts, and decision-makers' involvement in all phases of the modelling process of participatory decision making. Together with other authors, he argued that this ongoing involvement in model building contributes to a clearer collective understanding of a problem, improving the knowledge base and social acceptance of decisions (Addison et al., 2013).

Participation has been a powerful tool that brings significant advantages in decision-making. Advantages such as the fact that they put together 'different perspectives' that guarantees a greater variety of interests and experience that leads to different viewpoints of the problem at stake; 'greater pool of knowledge', usually a group of people has more information and experience to support a decision than just one individual; 'greater comprehension', those who intervened in the discussion and exchanged ideas tend to understand the logic behind the final decision; ‘increased acceptance and motivation’, The fact that they are working mostly and practically as a team means that the participants do not see the decisions as a personal decision, but rather as a group decision. Making the decision collective makes the participants more motivated; ‘training ground’, The participation of individuals with less experience in these participatory models serves as a way for them to learn and understand the dynamics of the subject in the matter (Hashim, Alam & Siraj, 2010).

In some cases, it is highly essential to use subjective opinion in forecasting methods. Formal forecasting methods - such as *Extrapolation*, *Leading indicators*, *Causal Models,* and *Probabilistic Methods* - are based on objective data processing. However, the need to rely on the opinion and to face the issue in a subjective manner can be justified by the following conditions:

1. Inexistence of historical data
2. Impact of external factors that make data about the past irrelevant - for instance, changes in public opinion
3. When there are subjective issues where ethical and moral considerations must be considered (Martino,1993)

Therefore, for these cases, it is necessary to have a technique capable of organizing and structure group communication to gather and organize all subjective opinions to reach a group decision. The technique we will focus on is called Delphi method, which is based on the criterion that information from judgments is indispensable (Rowe, Wright & Bolger, 1991). These judgments are obtained from the participants responses in Delphi method surveys.

Evidence shows that Delphi method leads typically to stronger decisions when compared with unstructured interacting groups and static groups (Rowe & Wright, 1999), and Riggs stated that Delphi procedures for long-range forecasting are superior to traditional conference methods (Riggs, 1983). However, several comparative studies have been made over the years, and it is quite difficult to compare Delphi method to other techniques, in terms of efficacy. Rowe and Wright concluded that most of the studies performed at that time to compare the Delphi method with other techniques were not sufficiently compelling since there was no apparent attempt to describe the relevant characteristics of the task and the characteristics of the panels. The efficacy and validity of a Delphi process depend heavily on these variables, and therefore they must be well-considered and analyzed before comparing this method with other techniques (Rowe & Wright, 1999).

Later in 2006, Landeta stated that Delphi method continues to be applied and is a valid instrument to forecast and support decision making. He concluded that comparisons with classical and statistical groups that use direct interaction showed that Delphi method it's better, however, when comparing with other techniques that share its terms of action and scope, there were no decisive conclusions to prove that Delphi method is superior or inferior to these techniques (Landeta, 2006).

Nevertheless, it is highly recommended to use Delphi method when the following reasons are present (Linstone & Turoff, 1975; Hanafin et al., 2007; Donohoe & Needham, 2009):

- When the problem cannot be solved through analytical techniques but can be addressed based on subjective judgments.
- When it is not possible or convenient to schedule face-to-face meetings due to issues related to the geographic location of each participant, scheduling, travel costs, and insufficient funds available to organize these meetings.
- When the number of participants is quite large, making it difficult for everyone to attend a face-to-face meeting
- When the problem addressed is dominated by ethical, political, legal, and social dilemmas

In short, the Delphi method is a social communication method indicated to reach a group consensus on complex problems where subjective opinion is indispensable. At the same time, a viable alternative when physical distance between participants predominates (Dalkey & Helmer, 1963; Linstone & Turoff, 1975).

2.2. Delphi History

Initially, Delphi method was developed as a technique, but years later it was reconfigured as a scientific method (Dayé, 2018). The technique was originally created at RAND Corporation at Santa Mónica to be a forecasting tool for dealing with future situations, through the statistical analysis of individual opinions from a vast number of specialists (Lindeman, 1975). It is then a procedure to help the participants obtain more reliable predictions than they might get from a traditional group meeting (Rowe & Wright, 2001), allowing the assessment of participants groups judgments at a distance, without requiring their physical presence (Snyder-Halpern, Thompson & Schaffer, 2000).

The first use of the Delphi process was in 1948, by a group of researchers from the RAND Corporation, to forecast horse race outcomes. The experience was promising but was nevertheless subject to criticism due to the presence of some defects (Lindeman, 1975) Later in the 1950s, Dr. Olaf Helmer, Norman Dalkey, and Ted Gordon studied and developed the Delphi

technique also at RAND Corporation (Hsu & Sandford, 2007; Linstone & Turoff, 1975-2011; Okoli & Pawlowski, 2004). It was only in 1953 that Dalkey and Helmer made the first significant use of the Delphi technique. They adapted the technique to deliberate a defense scheme, asking for the opinion of 7 experts in atomic war. Yet, the work was only published in 1962 due to confidentiality and security issues (Lindeman, 1975).

However, at that time, Delphi method was a very poorly researched technique with an insignificantly small number of published studies, but between 1969 and 1974 the number of Delphi studies increased significantly. The technique and its application were always in constant evolution, in terms of where and how it is applied. (Linstone & Turoff, 1975).

Fifty years after its creation, the Delphi method has become a widely used and recognized instrument, which assists in decision-making and makes predictions (Landeta, 2006). Today, the Delphi process has been applied in numerous fields of study and used across multiple disciplines such as health care (Humphrey-Murto & de Wit, 2018; Hart et al., 2020), policy determination, program and event planning (Hsu & Sandford, 2007; Chen et al., 2020), government planning (Linstone & Turoff, 2002), supply chain management (Ogden et al., 2005) (Khan, Haleem & Khan, 2020; Durugbo et al., 2020), environmental management (Clark et al., 2020; Zhang & Xi, 2020), education (Pucher et al., 2020; Johnson, 2020), marketing planning (Kumar, Sindhwani & Trivedi, 2020; Thomas, Paolo & Patrick, 2020) and many other fields. Besides its flexibility, Delphi method has also been used as a complement to face-to-face meetings. An example is given by e Costa et al. (2019) in which the knowledge acquired in a Delphi process was used to complement and support a subsequent face-to-face conference (e Costa et al., 2019).

Before proceeding to the next section, it is important to note that most of the papers refer to the use of expert opinion when referring to Delphi method, however, since this method can be applied by any group of people (expert or not) who want to make a decision, I will call each group member as a participant or stakeholder. But of course, if it is a more scientific or political decision such as the issue of wearing a mask during the COVID-19 pandemic, for example, it is advisable to seek the opinion of health experts, who in turn are still participants or stakeholders as well.

2.3. Delphi Definition and Description

Originally, the main objective of the Delphi method was to reach a reliable consensus among the opinions of a group. This consensus is worked out and achieved through individual and repeated questioning of a series of questionnaires interspersed over a certain period of time where the questions are all centered around a problem and are answered without direct confrontation between the participants (Dalkey & Helmer, 1963).

However, in 1975, Linstone and Turoff stated that “Delphi may be characterized as a method for structuring a group communication process so that the process is effective in allowing a group of

individuals, as a whole, to deal with a complex problem" (Linstone & Turoff, 1975, p.3). Therefore, Delphi method was no longer seen as a method in which a consensus had to be reached but rather as a method that helped to structure group communication. Delphi processes can bring clarity to an issue, increase stakeholder's understanding of the respective views of other participants, and explore different opinions (Devaney & Henchion, 2018).

One of the main keys to ensure an effective Delphi process is the competence of the moderator (i.e. investigator, facilitator, or researcher). The moderator is responsible for conducting the project and must have the competence to deal with the problem at hand. The moderator collects, interprets, and synthesizes the judgments given in each response, often having to deal with highly subjective responses. Therefore, it is necessary to have a qualified moderator capable of handling structured communication between participants and also capable of interpreting and analyzing the data correctly and impartially (Grime & Wright, 2014; Avella, 2016).

Like any other method, the Delphi method has a set of specific characteristics that need to be followed. *Anonymity*, *Iteration* with *Controlled Feedback*, and *Statistical Analysis of Group Response* are the main characteristics that differentiate the Delphi method from other participatory models for decision making (Dalkey & Helmer, 1963; Martino, 1993; Landeta, 2006). Through these characteristics, the Delphi method becomes a technique with greater effectiveness and relevance than other traditional techniques.

Anonymity – Usually responders are never anonymous to the moderator (Okoli & Pawlowski, 2004), however, the participation is done anonymously among the members, which allows free expression. It avoids the possibility of one panel member being influenced by another. Decisions are evaluated by the content of the response and not by the reputation of the member who proposed it (Skulmoski, Hartman & Krahn, 2007). Another advantage of anonymity is that members can change their minds in the following rounds without anyone knowing they. Without anonymity, members might hesitate to change their opinion in view of the responses given in the previous round, wanting to defend a stand once taken (Martino, 1993). Nevertheless, there are frequent cases where panel members know each other, however, they are unable to assign responses to each other (Gerrish & Lacey, 2010).

Multiple Iterations – Gives panel members an opportunity to rethink their perspective towards the progress of group work (Skulmoski, Hartman & Krahn, 2007). Throughout the iterations, experts are expected to become more perceptive in expressing their opinion (Hsu & Sandford, 2007).

Controlled feedback – This control ensures that the group only focuses on the original goal and does not get lost in the personal goals of panel members (Martino, 1993). Only the information inherent to the problem is selected by the moderator, thus contributing to better efficiency of the process and noise reduction (Hsu & Sandford, 2007).

Statistical Analysis of Group Response – Group responses are presented in statistics. All the opinions of the group are included in the forecast, unlike most conventional forecast methods that only present the majority opinion (Martino, 1993).

Besides the need to respect these characteristics in a Delphi process, it is also important to know certain principles that make the practice of Delphi method more efficient and effective. Rowe and Wright (2001) developed the principles that should be applied to conduct structured groups and obtain reliable judgments:

- Use heterogeneous participants with proper knowledge.
- Provide median panel estimate or average in Delphi feedback plus the reasons of all participants for their estimates.
- Rounds should be continued until the responses show stability (Usually 3 or 4 structured rounds are sufficient).
- Aggregate and weight the participants' estimates equally to get the final forecast.
- Questions should be appropriately framed and requested clearly and succinctly, avoiding the use of emotional terms and irrelevant information. Burgman (2011) stated that In this way, we can avoid possible uncertainties and misinterpretations of language that can lead to a distortion of judgments (Burgman et al., 2011).
- When possible instead of probabilities, provide participants with estimates of uncertainty as frequencies.
- When eliciting estimates of probabilities, use coherence checks (Rowe & Wright, 2001).

Regarding the process, it starts by defining what is the problem that is going to be discussed. After defining the problem, it's necessary to know who are going to be the participants of the process. In case the use of experts is required, it is necessary to make a selection and recruitment of a proper group of experts. Recruitment is done individually and is based on statements made by each one, which in turn are gathered in documents, literature analysis, or even possible interviews (Holloway, 2012). After defining who will participate, the questionnaire is prepared. Is then distributed to the participants and the rounds process begins. In the following figure 1, a scheme of how a Traditional Delphi (TD) process is accomplished is presented:

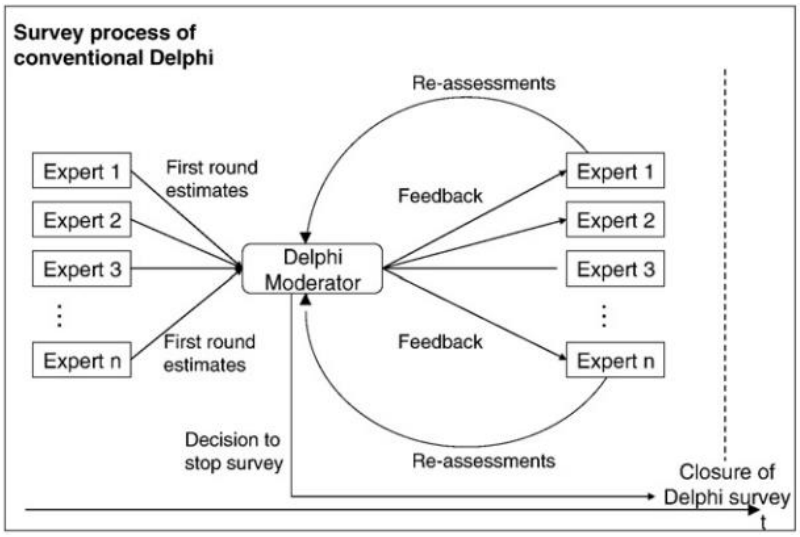

Figure 1 - Survey process of conventional Delphi (Gnatzy, Warth & Darkow 2011)

The survey starts with a first unstructured round in which a questionnaire is distributed to each panel member. Each questionnaire may have one or more questions on the subject that is being addressed, to which the participants have to respond. Questions usually include additional space where participants can present a justification for their responses (Landeta, 2006). After responding, the questionnaires are returned to the moderator who will analyze the statements made and arrange them in statistical format.

In the next round, the estimates calculated in the previous round are distributed to all participants. Participants have now the possibility to rethink and change their response in light of the previous round's results (Martino, 1972). The results are presented in statistical formats, such as measures of central tendency (e.g. median, mode, mean) and variation (e.g. interquartile range, range, standard deviation) (Flostrand, Pitt & Bridson, 2020). Each participant must defend their position in case their estimate is not aligned with the statistical results of the other estimates. The responses are again distributed to the moderator.

The procedure of the previous round is repeated, but with new statistical feedback. The moderator removes all comments that deviate from the main subject or comments that somehow identify the person, compromising anonymity. Participants returning to the questionnaire can again change their position or defend their previous position.

In short, this procedure must be repeated until panel members show no significant changes in their responses throughout the rounds (stability in responses) or if any previously selected stop criteria have been reached. Typically the statistical results are presented with the calculated mean and quartiles and the greater the difference between the upper and lower quartiles (interquartile interval), the greater the level of disagreement between the panel members. Finally, after

summarizing the participants' opinion on the problem in question, the moderator checks if consensus has been achieved or not (Martino, 1972).

Again, in Figure 1 the participants are referred to as experts but according to the context of this dissertation the members participating in any Delphi process will be only referred to as participants. It is also important to clarify that this description of the process is relative to TD. Delphi method has many variants, and there may be steps that are different from those described. For example, in the Real-time Delphi described by Gnatzy et al. (2011), the first round is skipped, and participants receive immediate feedback in their first statements.

2.3.1. Stop Criterion

One of the most important cautions to take in the execution of a Delphi process is to define which stop criteria is going to be used. We need to define when we will stop the process and summarize the final results. The following criteria is the most common according to a study from Diamond et al. and Dajani:

Number of rounds – Before the Delphi process starts, a maximum number of rounds is defined. When this number is reached, the process stops. In this case, the authors of the process must specify how the degree of agreement between the participants will be quantified (Diamond et al., 2014).

However, Dajani, Sincoff & Talley (1979) stated that stability of results is the main basis to declare the termination of a Delphi process:

Stability of results – If responses begin to stabilize and become more consistent (without significant changes) between successive rounds, the process stops (Dajani, Sincoff & Talley, 1979). Otherwise, if instability is still present, a new round is performed.

The majority of the Delphi studies evaluated by Diamond et al. (2014) were done by using a limited number of rounds as a stopping criterion without a formal criterion for consensus (Diamond et al., 2014). As previously mentioned, Delphi's main objective is not to reach consensus, but to structure group communication in an efficient way to deal with complex problems. However, regardless of the stop criteria used, an analysis of the summarized data from the last round is performed to see if consensus has been achieved. Consensus is difficult to measure, and it is often interpreted in different ways (Heiko, 2012; Humphrey-Murto & de Wit, 2018). The definition of consensus in a Delphi process has never been rigorously established (Mitchell, 1991), so its interpretation is often at the discretion of the team conducting the study. This absence of standards makes this part of the process quite debaTable (Heiko, 2012). Nevertheless, Dajani et al. stated that stability is the crucial key because consensus is worthless without stability in the responses. They created the following hierarchical stopping criteria for Delphi studies.

1. **Consensus:** When group unanimity is achieved by the same decision.
2. **Majority:** When there is consistency in more than 50% of the respondents, and if it is subsequently agreed among the minority, the study is terminated.
3. **Bipolarity:** When respondents are divided equally on an issue. When this phenomenon occurs it is necessary to understand and evaluate the stability of each half or there is a reformulation of the question and a new round is carried out.
4. **Plurality:** When a greater group of respondents reach an agreement, but the group is less than 50% of respondents. It is necessary to evaluate the nature of stability and consensus of this group of respondents and if stability is not stable, they should then rephrase the question and conduct a new round. In the case of stability, they can finish this question in specific and remove it from the questionnaire, or even reframe the question and include it in the next round.
5. **Disagreement:** When there is practically no agreement between the responses of the individuals keeping them unchanged regardless of the feedback from the other responses. In this case, if the disagreement is stable, either the question is reformulated or the process is given as terminated (Dajani, Sincoff & Talley, 1979).

If no maximum number of rounds has been stipulated, studies are usually guided by this hierarchy. However, as soon as one of these levels has been set, it is necessary to check the stability of responses to see whether the study should end or proceed to a next round. In conclusion, the stability of responses is one of the most important criteria to consider in a Delphi study, however, it is recommended to check not only for stability but also for consensus at the end of each round (Von der Gracht, 2012).

2.3.2. Delphi Weaknesses, Limitations and Challenges

The Delphi method is applied in several areas and some are quite distinct from others. It should be noted that there may be problems with its application depending on the area being studied, therefore the following limitations are inherent to the Delphi method in general and not in particular cases. Additionally, it is also important to note that some of the points mentioned below turn out to be challenges rather than weaknesses of the Delphi Method because they are related to the moderator's ability to implement Delphi processes.

Starting with the inherent weaknesses and limitations of the Delphi Method:

- **Participants panel instability** – The ideal way to conduct a Delphi study would be to maintain the same participants throughout the study without substitution. Unfortunately, one of the problems with the Delphi method is that there can be a significant dropout rate that compromises the effectiveness of the process, making it difficult to attain

convergence of opinions. Contractual agreements with participants can contribute to keeping them throughout the process (Gordon & Helmer, 1964) In case experts are used, this effect can be even more severe because they are usually busy individuals and a long process that is not practical and is poorly described/explained by the organizer does not contribute to their continued commitment (Webler et al., 1991).

- **Time consumption** – Possible interviews, design, distribution, collection, and analysis of questionnaires are activities that are typically time-consuming (Keeney, Hasson & McKenna, 2006). It was even reported by Duffield (1993), that it can usually take 8 weeks to complete a round of the process (Duffield, 1993). Too much time between rounds can be harmful, not only because it calls for participants to give up, but also because it causes a change of opinion since the more time goes by, the more knowledge is acquired by each participant, which can sometimes lead to a severe change of their response (Gordon & Helmer, 1964; De Loe, 1995; Day & Bobeva, 2005; Donohoe, Stellefson & Tennant, 2012). Also, if the process begins to be time-consuming, participants who have opinions far from the majority may tend to give up the process and not complete it to the end, or even agree with the majority just to make it end (Humphrey-Murto & de Wit, 2018). In these cases, moderators should begin the process by emphasizing to all participants the importance of maintaining their participation until the end (Addison et al., 2013; Humphrey-Murto & de Wit, 2018; Devaney & Henchion, 2018).
- **Poor definition of consensus** – The definition of consensus is one of the major problems in Delphi studies. Agreement among participants is a measure that has not been fully defined (since the origin of the method) and therefore there are many different interpretations of consensus in Delphi studies (Boulkedid et al. 2011; Heiko, 2012; Humphrey-Murto & de Wit, 2018). Diamond et al. conducted a research and evaluation of 98 Delphi studies and they concluded that the interpretation of consensus was highly varied and most of the times were poorly explained (Diamond et al. 2014). To mitigate this problem, consensus should always be defined before starting the process and it should also be very well explained to all participants (Humphrey-Murto & de Wit, 2018). It is also important to distinguish that consensus is not the same as the level of agreement, but is rather a measure used to define the level of agreement in a Delphi process (Heiko, 2012).
- **Consensus through an undue average** – The median as a measure of the group opinion and the quartile as a measure of disagreement between opinions cannot be taken as absolute truths. Opinions that are far from the majority opinion are often lost throughout the process, however, many of them may be correct if they are properly justified to the panel. It is, therefore, necessary to insist on the justifications of minority opinions so that they are not rejected in vain, only in an attempt to reduce noise (Gordon & Helmer, 1964).
- **Self-defeating and self-fulfilling prophecies** – There is a possibility that there are responses that somehow manipulate their own prediction. Following the example given by Gordon and Helmer, if at the time a participant said that the US was closer to the Moon

than Russia, the prediction would be self-fulfilling, but if an expert said that Russia was closer to the Moon than the US, the effort would probably become greater on the part of the Americans to overcome the "advantage of the Russians". The forecast would therefore influence the US effort. This effect brings the possibility of manipulating the forecasts in some way, damaging the results (Gordon & Helmer, 1964).

- **Competence and Applicability –** Just like most methods, the Delphi method is not suitable for all types of problems. However, it offers strong flexibility when it comes to problems that mix social, ethical, and moral values with scientific evidence, unlike other conventional strategies (Webler et al., 1991).

The following points relate to the challenges a moderator faces in implementing Delphi processes:

- **Misinterpretation of results** – Subjective opinion is predominant in the responses given by panel members. One of Delphi's challenging problems is the way the moderator interprets such responses, as it can lead to an incorrect summary of results. The moderator is in a certain way decisive and has an important influence on the results (Webler et al., 1991; Grime & Wright, 2014).
- **Ambiguous and vague questions** – Overly ambiguous questions can lead to different interpretations by participants. It is important to reduce this ambiguity as much as possible to ensure that all participants respond to the same intended purpose of the question (Gordon & Helmer, 1964; Burgman et al., 2011).

In addition to the present weaknesses and challenges, it is also necessary to consider some limitations regarding the use of experts in Delphi studies:

- **Selection of experts** – The problem consists of identifying the experts and selecting who participates. Different experts have different perspectives, and as a consequence, no matter how large the sample, the selection of the different perspectives should be systematic. Cultural bias is also an aspect that must be considered in the selection. The questions can be interpreted differently, and one way to mitigate this problem and cancel the biases is to ensure diversity among the panel members (Webler et al., 1991).
- **Participants' Competence** – The questions are asked around the same issue, however, the areas covered may vary from question to question. Although experts are selected carefully, we cannot ignore the fact that there may be experts who do not have the necessary competence to answer certain questions. This effect may be detrimental to the process as some responses from competent members may be diluted by responses from less competent members. One solution would be to encourage experts to leave answers blank when they feel insecure about their competence (Gordon & Helmer, 1964).

In general, the main limitations of the Delphi method are the time associated with its practice (mail system and other paper-based methods for data collection are slow), which in turn can lead to instability in the panel of participants, the high subjectivity in the responses of some unclear and

vague questions, which can lead to different interpretations and the lack of agreed standards such as the difficulty of defining consensus.

2.3.3. Delphi Types

Over the years, the Delphi method has constantly been a target of studies and has also been approached with new perspectives and slight modifications of the original RAND design. These new perspectives and modifications have originated several variants of the original method and their creation often depends on the study's objectives and aims (Linstone & Turoff, 1975). Hugh McKenna stated that it is important to bear in mind that too many modifications may compromise the rigor and validity of the TD approach (McKenna, 1994).

One of the main and most relevant variants that have been mitigating the previously mentioned weaknesses of TD is the e-Delphi (Donohoe, Stellefson & Tennant, 2012). According to Donohoe, Stellefson & Tennant (2012), e-Delphi is used when Delphi method is applied through the use of a web network, i. e. when a Delphi process is conducted for example by e-mail, by web-surveys, or even by specific web-surveys integrated into a web platform that was designed specifically for Delphi processes. However, going back to section 1.1. where a contextualization of the dissertation was made, it was assumed that platforms that were designed with specific web-surveys to implement Delphi processes are referred to in this context as Web-Delphi Platforms. For this reason and communication purposes, it was assumed in this dissertation that Web-Delphi refers only in particular to the use of Delphi method in these specific platforms called Web-Delphi Platforms while e-Delphi refers to a general term when a Delphi is conducted through any online format.

Having the concept Web-Delphi and e-Delphi well defined, research was done on the various Delphi variants that have emerged over the years. On Table 1 is presented an adapted Table from Gerrish and Lacey (2010) with a brief description of the TD and some of its most significant variations:

Table 1 - Delphi Variants Table adapted from (Gerris & Lacey, 2010, p.232)

Delphi Type	Description
Classical Delphi	The original purpose was to structure and organize group communication with the aim of reaching a group consensus by conducting 3 or more postal rounds (Gerrish & Lacey, 2010)
Decision Delphi	Similar to Classical Delphi but its aim is to prepare, assist, and make decisions rather than coming to consensus (Rauch, 1979).
Policy Delphi	The goal is to generate the strongest possible opposing views on a major political issue. In this case, the participants are informed referees and defenders. There are no experts (Turoff, 1970)
Real-Time Delphi	Feedback is given directly to the expert in real time as soon as he makes a judgement. There are no explicit rounds. So when making judgments, participants are immediately confronted with the aggregate results of all the estimates of the experts who participated previously (Aengenheyster et al., 2017)
Group Delphi	Anonymity is not applied. This variant was designed to obtain and consolidate expert opinions on a subject in short periods of time (Webler et al., 1991)
Modified Delphi	The main difference is the replacement of the first postal round with focus groups or face-to-face meetings. Usually it's used fewer than three postal rounds (McKenna, 1994)
Technological Delphi	Similar to Real-Time Delphi, the difference is that experts use technology like manual keyboards that allow experts to answer questions immediately and receive instant feedback. The technology allows the average and median to be calculated as the experts respond (Gerrish & Lacey, 2010)
e-Delphi	similar process to the classical Delphi but administered by web system techonologies such as email and online web-surveys (Chou, 2002).

The present Table 1 and also statements from Grime and Wright (2014) show how flexible Delphi method is and how it can be easily adapt to different group's needs (Grime & Wright, 2014). Although all these Delphi categories have a similar structure with several points in common, it is crucial to choose the most appropriate format to use, depending on what is intended to be concluded in the research (Novakowski & Wellar, 2008). Different designs can result in different goals. Some are very specific techniques while others result from combinations between parts of different designs. Each design has a different purpose, the choice depends on what is required in the problem addressed. To ensure a correct and effective application of the Delphi method, whatever the variant, the moderator must be duly qualified to deal with its complexity (Gerrish & Lacey, 2010). However, when using Delphi variants, we have to consider if the previously presented characteristics of the original method, are present.

2.4. Web-Surveys

Before moving on to the WDP subject, it is important to discuss the emergence of web-surveys and what advantages they brought in the implementation of Delphi processes.

Web-surveys are an online method of systematic data collection that appeared in the early '90s shortly after the web concept was introduced to the world. Its appearance made it much simpler to build a survey and made it even easier to collect data when comparing to traditional pen and paper and telephone surveys. However, several questions arise when using web-surveys. One of the main questions and also the one that concerns us most when we talk about Delphi studies is whether the use of a web-survey is suitable for the problem in question. To ensure a good practice it is advisable to have an enhanced knowledge of web-survey methodology.

The webform is a component of a web page that allows the user to have passive access and submit some data. After concluding the form, the submitted data is then transferred to the server. When referring to web-surveys, we are talking about a special type of web form called *web questionnaires* where a user submits the input (survey answers) and then the same input it's transferred to a researcher's server(Callegaro, Manfreda & Vehovar, 2015).

In the beginning, when the web-surveys emerged, they were static and quite basic with just questions organized in a single web page. Years later in 1994, web-surveys started to be more developed and became more interactive instead of static. The constant technological progress and development have also allowed easier access to web-surveys giving users greater flexibility when filling out a questionnaire. Initially, these surveys were only accessed through computers (desktop, laptops, and notebooks), but over the years, Internet access has become much more flexible with the emergence of smartphones and Tablets, thus increasing the use of these devices to access and answer web-surveys (Callegaro, Manfreda & Vehovar, 2015).

2.4.1. Web-Survey Advantages

Web-surveys brought a new direction to the practice of the Delphi method. Compared to the traditional pen and paper method, mail and telephone, web-surveys have advantages that have allowed to create WDP to overcome some of the weaknesses in the Delphi method. Compared to traditional survey methods, such as telephone and face-to-face surveys, web-surveys end up being cheaper to perform and faster (Fricker & Schonlau, 2002). Data from web-surveys is available in real-time and is easier to import into data analysis programs (Archer, 2003).

Here is a more detailed description of the advantages of using a web-survey gathered by Callegaro, Manfreda & Vehovar (2015):

Reduced costs – In a comparative study between the web-based and the paper-based survey methods, Greenlaw and Brown-Welly concluded that a web-survey response costs on average $0.64 (with previously email invitations), and responses done by email costs $4.8. The mixed option where the invitations are done by mail and the responses by a web-based survey costs $3.6 (Greenlaw & Brown-Welty, 2009). A similar study done through a cost-effectiveness analysis also concludes that web-based surveys result in lower costs, being 2.7 times more cost-effective than paper-based surveys. (Hardigan, Succar & Fleisher, 2011). Also compared to telephone surveys, the price per respondent in web-based surveys can be three to four times lower (Strabac & Aalberg, 2011).

Speed in processing and collecting data – The difference in the speed of data collection is noTable compared to traditional surveys (Mehta & Sivadas, 1995; Bachmann, Elfrink, & Vazzana, 1996; Callegaro, Manfreda & Vehovar, 2015). Heiervang and Goodman concluded in a study that the time expended for full web-based interviews was a quarter of the time expended for face-to-face interviews (Heiervang & Goodman, 2009). Additionally, Aoki and Elasmar have concluded that on average a mail survey takes 50 days while via web-survey it takes about 3 days on average (Aoki & Elasmar, 2000).

Ease of deployment – Nowadays it is not necessary to have notions of programming or even a vast computer literacy to be able to develop a web-survey. Callegarro, Manfreda, and Vehovar stated that ease of implementation is an absolute advantage of web-surveys when comparing to face-to-face, mail, and telephone surveys that involve a much more complex process (Callegaro, Manfreda & Vehovar, 2015).

Computerization of the survey – The computerization of the questionnaire allows making the response process more efficient. Data can be validated and processed immediately, errors can be avoided, fields are filled in more easily with help tools and each question in the questionnaire can be programmed according to the previous answer. It also allows you to save the progress of responses, which is useful, in case there is a possible breakdown of the site or the computer goes instantly off (De Leeuw, Hox & Snijkers, 1995; Callegaro, Manfreda & Vehovar, 2015).

Multimedia capabilities – Multimedia capabilities that allow an innovative and extensive interface such as pictures, sounds, backgrounds, more innovative and appealing question display that allow an improved survey experience making them more appealing to participate (Callegaro, Manfreda & Vehovar, 2015).

Geographical and temporal flexibility – Web-surveys are seen as a 24/7 data collection where an individual can have access from any place in the world. Unlike face-to-face meetings and mail, the web-survey consists of a unique alternative to access all respondents who are geographically dispersed, thus saving travel time and costs as well as the time associated with the traditional mail process (Callegaro, Manfreda & Vehovar, 2015).

Respondents autonomy – Respondents have greater freedom to answer the questionnaire since they are able to reread the questions, answer the questionnaire at their own time without the pressure of the person asking the questions and even have time to consult documents to clarify doubts before answering. They can even rethink the answers and if necessary, change them before submitting the survey. There is no need to schedule appointments unlike telephone and face-to-face (Callegaro, Manfreda & Vehovar, 2015).

Another aspect that should be considered for increasing the speed of data collection and as well for decreasing dropout rates is the use of reminders and invitations by e-mail. Several studies highlight the importance of this practice (Kwak & Radler, 2002; Stanton & Rogelberg, 2001). These advantages contribute to the mitigation of some limitations associated with the Delphi method. The fact that it is possible to administer online surveys has made it conceivable to build platforms with specific web-surveys for Delphi (WDP).

2.4.2. Principles for Constructing Web-Surveys

Web-surveys must have a user-friendly design supported by different computers and other electronic devices ensuring maximized participation. The survey design has to be sufficiently user-friendly to match the way computers operate as well as the way people usually handle the surveys. Also, every kind of survey should enable different ways to complete the survey (traditional mail for example) to ensure maximization of participation (Dillman, Tortora & Bowker, 1998). In order to ensure good and effective web-surveys practice, Dillman, Tortora, and Bowker (1998) have developed the following principles:

- Provide an appealing and motivational design on the welcome screen of the questionnaire, with simple instructions to ensure the engagement of participants.
- The first question should be fully visible and easily perceived. The first question is the participant's first impression and therefore will define whether the survey is complicated or easy to complete.
- The presentation and format of the questions should be similar to those used in paper questionnaires.
- Provide instructions for all necessary computer actions in a clear and succinct way (Button instructions, explanatory examples, and others).
- The filing instructions should be accompanied by the question and not in a separate section outside the survey.
- The questions should allow participants to skip questions to answer the subsequent ones. Participants should also have the possibility to go back and change the answers.

- The questionnaires should be sufficiently perceptive to ensure that participants see all questions and know which ones are unanswered.
- Provide graphic symbols for monitoring the progress of the participants and to be able to see at what stage of the questionnaire they are.

Although these principles were developed 20 years ago, they are highly significant and applied nowadays and therefore these principles should be present in the WDP.

2.5. Web-Delphi Platforms

One of the biggest complications of TD is the time associated with its practice. The long duration is considered attrition and also a deterrent for Delphi research, not only for moderators but also for participants (Donohoe, Stellefson & Tennant, 2012). Traditional mail systems for conducting Delphi studies have the disadvantage of subjecting moderators to long waiting periods in gathering the participants' opinions and in turn, in building the questionnaires. Another problem is the possibility of panel members not returning or even completing the questionnaire, so several contributions could be valuable and not being accounted for (Chou, 2002).

However, today's technologies have brought new opportunities to the Delphi Method (Haynes & Shelton, 2018). The creation of WDP and its use allow organizing, controlling, and facilitating communications between the panel members and the moderator (Donohoe, Stellefson & Tennant, 2012). The use of online surveys in WDP to conduct a Delphi study minimizes the time needed to conduct it and maximizes its participation. The ease of connection through the Internet makes it much easier to select potential participants globally and, consequently, the communication between the moderator and the participant. The delivery of surveys, survey deadline reminders, data collection, and analysis have become much easier with the introduction of Internet-based technology. Time is then optimized and there is a drastic decrease in recruitment costs (Donohoe & Needham, 2009; Holloway 2012).

2.5.1. Web-Delphi Platforms Advantages

Looking back to the inherent advantages of web-surveys, we can then associate them with the advantages related to Delphi processes conducted on a WDP:

Geographical and temporal flexibility – Full-time work, family events, and commitments are obstacles that compromise the presence of a participant in a Delphi study. As long as a participant has access to the internet and the respective web address of the platform, can be anywhere in

the world and participate in the study. Accessibility is superior as it eliminates the possibility of having to travel to the research site (Reddy, 2006; Jones et al., 2008; Holloway 2012).

Anonymity within the group is promoted – No face to face contact, so there is no possibility of a panel member identifying another member (Reddy, 2006; Holloway 2012).

Speed in processing and collecting data – The data are collected more easily and more quickly introduced into the software, unlike traditional methods that use paper and pen and telephone (Jones et al., 2008; Holloway 2012).

More cost-effective – Reduced travel costs, since participants do not need to be physically present at interviews and survey rounds (Holstein, 2002; Cantrell & Lupinacci, 2007; Holloway 2012) and also the cost-effectiveness demonstrated in the previous Web-Survey Advantages Section.

Efficient recruitment – The possibility of recruiting and interviewing participants via email or other online platforms has made the recruitment process quicker and easier, without limiting the choice of participants due to their geographical location (Holstein, 2002; Holloway 2012).

Promotes more reflective responses – There is more freedom for participants. They have more time to think and work on what they will answer as they can access the question more easily and the answer is sent immediately once they submit it (East et al., 2008; Holloway 2012).

Multimedia capabilities – Advanced platforms that allow the insertion of multimedia content such as explanatory videos and other informative content that can be very helpful to engage participants and also promote a better explanation of the process (Helms, Gardner & McInnes, 2017).

Embedded data along with adaptive questions – Through this data, it is possible to create a more interactive system of questions and in a certain way adapt them to the person who is answering them. In this way, the process becomes more efficient and easier in the user's view, increases the participants' response rate and decreases the time taken to finish the surveys (Helms, Gardner & McInnes, 2017).

By analyzing the previous points, we can conclude that the emergence of the WDP has contributed to the mitigation of the main limitations of Delphi and also to the maximization of its advantages, however, several issues mentioned above in section 2.3.2. remain unresolved such as the high subjectivity which makes it difficult to interpret results and lack of agreed standards. Therefore, it is a subject that should be well studied and developed to increase the efficiency and effectiveness of the Web-Delphi process.

2.5.2. Web-Delphi Platforms Shortcomings

Although the WDP has brought several advantages and also tackled issues related to the original Delphi Method, the platforms are also subject to other shortcomings:

Possible lower response rates than mail questionnaires – A recent study from Daikeler et al. (2019) showed that web-surveys, in general, have lower response rates than traditional methods. This can result from different causes such as users feeling that web questionnaires are less mandatory, errors in filling in and submitting the questionnaire, lack of support and understanding of the platform, and other causes (Daikeler, Bošnjak & Lozar Manfreda, 2019). Other earlier researches also support this idea (Duffy, 2002; Cantrell & Lupinacci, 2007; Boulkedid et al. 2011), however, people are becoming more and more familiar with the use of the web and therefore this scenario is constantly evolving.

Participants and even researchers may need support – Conducting a process through a digital platform can be a difficult task if the platform is complex and not very self-explanatory. The participants especially can have difficulties during the process, leading to time-consuming situations or even the dropout of the process (Duffy, 2002; Holloway, 2012).

Privacy threat with potential hackers – Confidentiality and privacy may be threatened due to possible hacker attacks due to lack of data protection. However, data protection systems and privacy policies are increasingly developed (Cantrell & Lupinacci, 2007; Allen & Roberts, 2010; Holloway, 2012).

No Access to Electronic Devices and/or internet – Nowadays it is quite unusual, but there may be participants who do not have access to the internet and/or electronic devices to answer the questionnaire (Jones et al., 2008). There must be an alternative to ensure the participation of these individuals.

Progress saving – Computer malfunctions, internet failures, and other problems can occur during the process. In these cases, it is possible that participants may have already answered a large part of the questionnaire and the progress may not have been saved (Toma & Picioreanu, 2016). Saving progress is vital to ensure good practice in completing the questionnaire. This feature makes the process easier and more practical in the eyes of the consumer.

When analyzing the respective WDPs, it is necessary to consider these drawbacks and understand how these issues can be mitigated or even solved to optimise these platforms' performance. With the development of technology, the evolution of features and the appearance of new ones have made it possible to combat these drawbacks, and therefore, it is important to register which are these features and if they are currently present in today's WDPs. Therefore, all these insights obtained in the literature review will serve as one of the bases for constructing the framework outlined in section 1.2, point 3, which allows recording which of these features that optimize and facilitate the implementation of Delphi processes WDPs are present in these

platforms. In this way, it will be possible to analyse the different WDPs found in the present dissertation.

However, the Framework will not be enough to distinguish which of the WDPs is better or which criteria are more relevant. Therefore, besides the fact that this framework allows comparing WDPs with each other by examining the presence or absence of features, it will also serve as a basis for building a multi-criteria model that will allow to evaluate each WDP. In this way, the multi-criteria model will allow to compare WPD and investigate further which one should be chosen for conducting online Delphi processes. The next section presents a review of Multi-criteria Decision Analysis content with particular focus on the MACBETH approach.

2.6. Multi-criteria Decision Analysis (MCDA)

Reviewing the objectives defined in section 1.2, one of the objectives of this dissertation is to propose a multi-criteria decision analysis model. This model has the purpose of enriching and consolidating the comparative analysis between the platforms made by the Framework and consequently helping possible decision-makers choosing a Web-Delphi platform to implement Web-Delphi processes. Therefore, it is also necessary to conduct a literature review regarding MCDA contents, to propose a consistent and feasible multi-criteria model.

2.6.1. MCDA Introduction

One of the most well-known definition for Multi-criteria Decision Analysis (MCDA) in the literature was presented by Belton and Stewart (2002, p.2) as “an umbrella term to describe a collection of formal approaches, which seek to take explicit account of multiple criteria in helping individuals or groups explore decisions that matter”. MCDA are methods that are used to support decision-makers in decision-making processes, through the use of multiple criteria and also through the use of subjective opinion of each decision-maker involved in the process (Ishizaka & Nemery, 2013). Therefore, in addition to offering better-supported techniques to compare project alternatives, MCDA also has the advantage of having the ability to incorporate project stakeholders’ views and opinions into the evaluation of alternatives (Linkov et al., 2005). As Belton & Stewart (2002) stated, MCDA seeks to integrate objective measurement with value judgment in decision-making processes while managing the subjectivity and making it explicit. It is important to note that multicriteria frameworks are only used as an aid to decision-making and there is no optimum solution or “right answer” as response to the problem at hand. MCDA processes are divided into three phases: Problem identification and structuring; model building and use; and the development of action plans (Belton & Stewart, 2002). The following Figure 2 describes the different steps that each phase includes:

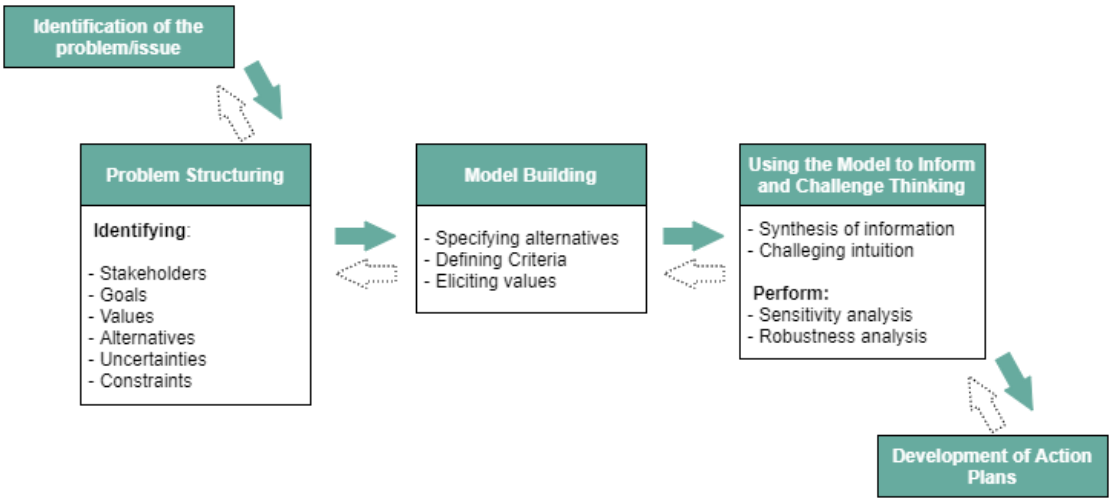

Figure 2 - MCDA process (Adapted from Belton and Stewart, 2002)

The use of MCDA methodology is important in decision-making processes since it helps decision-makers to learn and understand the problem faced, be more impartial, deal with subjective elements with transparency, not making decisions with insufficient information, and not focusing only on a small set of criteria or even on only one criterion (Lahdelma, Salminen & Hokkanen, 2000; Belton & Stewart, 2002). Usually, the construction and application of these models involve 4 actors: The **Decision Maker**, which can be a single one or a group of individuals; the **Analyst**, which provides methodological support for the decision process; the **Specialist**, which is a professional of the object of study and the **Client**, which is an intermediary between the Decision Maker and the Analyst (Frazão et al., (2018).

Currently, there are many different MCDA methods, for example, Analytic Hierarchy Process (AHP), Evaluation matrix (Evamix), ELECTRE III, MACBETH and many others. What all MCDA methods have in common is the fact that the decision-making process is based on the analysis of multiple criteria rather than just on the analysis of one dominant criterion (De Montiset al., 2000).These MCDA can be classified into 3 different types of models (Belton & Stewart, 2002):

- *Value measurement models*: consists of constructing numerical scales that allow each alternative to be scored with a numerical value. First these scores are given partially regarding each criterion, and then they are aggregated to generate an overall score for each alternative (example:multi-attribute utility theory and multi-attribute value theory).
- *Goal, aspiration or reference level models*: consists of seeking alternatives that are closer to achieving pre-established goals and aspirations (example: weighted and lexicographic goal programming).
- *Outranking models*: consists of comparing pairs of alternatives with respect to each criterion in order to establish levels of preference among the different alternatives (example: ELECTRE, PROMETHEE, REGIME).

The proposed model of the present dissertation is focused on Value Measurement Models following the MACBETH approach developed by Bana e Costa & Vansnick (1994). The reason for opting for a Value Measurement Model is because this paper intends to evaluate the platform on multiple criteria and to know not only its raking but also the distance in value of the alternatives. The MACBETH approach was chosen because it requires only qualitative judgments regarding differences in attractiveness to quantify the criteria, the areas of concern, and consequently the alternatives, unlike other methods that require quantitative judgments. The MACBETH methodology is described in the following sections.

2.6.3. MACBETH

MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique) is described as a “multicriteria decision analysis approach that requires only qualitative judgments about differences of value to help an individual or a group quantify the relative attractiveness of options” (Bana e Costa, De Corte & Vansnick, 2016, p.3). Since the goal of this section is to build an MCDA that aims to quantify the attractiveness of each alternative, in this case, of each WDP, MACBETH was the method chosen for this purpose.

MACBETH is a socio-technical process and, as the name suggests, a socio-technical process consists of a process involving two distinct components: a *social component* and a *technical component*. The social component of a socio-technical process consists of the adoption of social elements of decision-making techniques (Decision Conferences, Delphi Method, etc.) to incorporate the decision-maker's needs. It's essentially "how and when will the stakeholders and key players contribute in the MCDA model?". On the other hand, the technical component it’s essentially “What form of MCDA will be applied and how will it be implemented?”. This technical component consists of the integration of technical elements of multi-criteria value measurement, and in this case, the MACBETH methodology is proposed (Bana e Costa & Oliveira, 2012). This is a practice that offers multiple benefits in multi-criteria model constructions: improves internal communication between the different stakeholders; develops a shared understanding of the portfolio in question; and generates a mutual commitment and a sense of common purpose regarding which alternatives best meet the decision-maker objectives (Phillips & Bana e Costa, 2007).

In contrast to other multi-criteria methods, the generation of numerical scores for the alternatives in each criterion and the weighting of the criteria by MACBETH is done through the collection of qualitative judgments about the difference of attractiveness between two elements at a time (Bana, Meza & Oliveira, 2013; Dhouib, 2014; Pereira Dias & Fontes, 2019). The method follows the principle of converting qualitative information into quantitative information and this conversion is done through the use of a semantic scale with different verbal levels of attractiveness (Clivillé, Berrah & Mauris, 2007). The DM is asked to compare pairs of alternatives/options (two options

at a time) by choosing one of the seven qualitative categories of difference in attractiveness (null, very weak, weak, moderate, strong, very strong, extreme) as a value judgment about the difference of attractiveness he feels between one option and another (Bana e Costa & Vansnick, 1997, 2011).

To facilitate the implementation of the MACBETH method, Carlos Bana e Costa, Jean-Marie de Corte, and Jean-Claude Vansnick designed a software called M-MACBETH (available at www.m-macbeth.com) (Bana e Costa, De Corte & Vansnick, 2003). This software allows to generate scales, generate graphs and figures to support the analysis, detect inconsistencies in the judgments entered in the program, and also allows to perform sensitivity and robustness analysis regarding the models created. In these processes, two fundamental entities play an important role: the decision-maker (DM) who consists of the person who wants to make the decision regarding the defined problem, and the consultant (facilitator or decision analyst) who handles the implementation of the process in M-MACBETH. The implementation of MACBETH follows the principle that the problem and its solution belong solely to the decision-maker and not to the consultant, i.e., the consultant cannot have any influence on the decision and must remain neutral towards the decision maker's judgments (Bana, Meza & Oliveira, 2013).

MACBETH follows the most common stages of MCDA approaches (Bana, Meza & Oliveira, 2013):

- **Structuring** – Structuring the problem and evaluation elements.
- **Evaluation** – Development of the multi-criteria evaluation model.
- **Recommendations/Testing** – Testing to determine if the model is requisite, performance of robustness and sensitivity analysis and development of recommendations.

The following points describe in more detail the respective phases of a MACBETH process.

2.6.3.1. Structuring

It is necessary to structure the problem, which means defining which alternatives to analyze and which areas of concern and criteria are going to be used to evaluate the alternatives. A criterion is a tool that allows you to compare alternatives according to a particular point of view (Bouyssou, 1990) and an area of concern consists of a set of criteria. To make the criteria operational it is necessary to construct descriptors of performance that enable the evaluation of the performance of each alternative on each criterion. The descriptors of performance correspond to levels of impact of each of the defined criteria. In order to avoid model redundancy, the definition of areas of concern, criteria, and descriptors is a recursive learning process (figure 3).

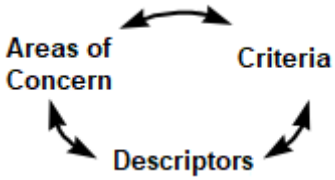

Figure 3 - Learning process (Adapted from Bana e Costa & Beinat, 2005)

Linking a descriptor to a criterion is basically choosing which of the characteristics of the alternatives are most appropriate for the problem in context. However, sometimes it is common to have different characteristics that are different descriptors for the same criterion thus implying a possibility of overweighting that criterion. In addition, the discussion of these alternative descriptors may reveal areas of concern that had not been considered before and which in turn may require restructuring the criteria and the remaining elements (Bana e Costa & Beinat, 2005).

The criteria represent the aspects that are going to be appraised and compared in the different alternatives and must have the following characteristics present (Bana e Costa, De Corte & Vansnick, 2005):

- **Intelligible:** the criterion should be clear and objective in a way that does not cause misunderstanding and possible miscommunication between the decision-maker and the facilitator.
- **Isolable**: the criterion is isolable if its preference is independent from the performance of the other criteria.
- **Consensual**: the criterion must be accepted by the entire group of decision-makers as a sufficiently significant criterion on the influence of the decision.
- **Operational:** the criterion is operational if it enables the construction of a descriptor of performance and the construction of a local preference scale associated with the performance levels of this criterion.

A criterion must fulfill this set of properties. On the other hand, if we are dealing with a family of criteria (area of concern), this set needs to satisfy the following conditions (Keeney, 1996; Belton & Stewart, 2002):

- **Consensual:** in order to facilitate the communication and provide a shared understanding between the decision-makers of the concepts used in the analysis.
- **Non-redundant** and **Concise:** in order to avoid double counting of possible consequences. Ensure that no more than one criterion measures the same factor.
- **Decomposable:** in order to allow a separate treatment of different goals in the analysis
- **Complete (exhaustive):** in order to capture all important aspects of the problem and keeping the level of detail to the minimum required.

Generally, methods based on the use of value functions, seek to identify a value tree that exposes the hierarchy of the criteria by using two different approaches (Belton & Stewart, 2002):

- **Top-down approach** described by Von Winterfeldt and Edwards (1986)
- **Bottom-up approach** described by Buede (1986)

The Top-down approach is a more objective approach where an expansion of the general objectives into more detailed concepts is made in order to better explain and clarify them. This expansion should stop when it is felt that the emerging criteria are measurable. Contrary to the previous approach, in the bottom-up approach details are initially elicited from a reflection on the weaknesses and strengths of the available alternatives. The result of this phase of the analysis should be a value tree or a family of criteria that reflects the values of the decision makers. In case the value tree that has been created or the criteria that have been defined do not respect the previously mentioned characteristics, it would be necessary to restructure the model (Bana e Costa e Beinat, 2005).

2.6.3.2. Evaluation

- **Construction of Value Scales**:

Once the previous steps are completed we can build the value functions that allow transforming the performance of each alternative into value. As already mentioned, following the MACBETH approach, the construction of the value scales are made from qualitative judgments, based on a predefined semantic scale, regarding the difference in attractiveness between the performance levels of the descriptors. These judgments are then entered into the M-MACBETH software where value scales are generated through linear programming.

- **Criteria weighting**:

After building and validating the value functions, the next step is to calculate the weights of the criteria and the weights of the areas of concern. First, it is necessary to consider the performance levels (good and neutral) of each criterion to calculate the weights. The definition of these levels contributes to the intelligibility of the criteria and their absence is seen as an mistake in the calculation of the criteria weights (Keeney, 1996). Similar to the previous step, MACBETH is a non-numerical method that allows the calculation of these weights through the use of qualitative judgement. Similarly to the value functions, qualitative judgments are made regarding the importance of improvement of the performance levels (swings) of each criterion between its reference levels, based on the semantic scale present in MACBETH. Then, through linear programming, the weights of each criterion are calculated and normalized.

- **Additive Model**:

Subsequently, it is possible to evolve the MACBETH process into a quantitative evaluation model. With the respective value scales suggested by the M-MACBETH software and their consequent validation, along with the weights of the criteria previously calculated, it is possible to calculate the overall score for each of the alternatives, through a weighted sum of the scores obtained by them in the multiple existing criteria. The simplest and most common way of adding value is through the additive model represented in the following equation 1 (Bana e Costa & Vansnick, 1997):

$$V(A) = \sum_{j=1}^{n} p_j \cdot v_j(A_j) \quad with \begin{cases} v_j(good_j) = 100 \\ v_j(neutral_j) = 0 \end{cases}$$

$$with \quad \sum_{j=1}^{n} p_j = 1 \; and \; p_j > 0 \; (j = 1, \dots, n) \qquad (1)$$

Where,

- $V(A)$ is the overall performance value of alternative A;
- A_i is the performance of alternative A no critério j $(j = 1, \dots, n)$;
- $v_j(A_j)$ is the partial performance value of the alternative A on criterion j $(j = 1, \dots, n)$;
- $v_j(good_j)$ and $v_j(neutral_j)$ are the "good" and the "neutral reference levels of performance on criterion j;
- p_j corresponds to the weighting coefficients or relative weights of the criteria, which allow the transformation of the partial value units, v_j, into overall value units V.

2.6.3.3. Recommendations/Testing

Finally, it is necessary to understand if the model is requisite considering the problem at hand. When the model's form and content are sufficient to provide satisfactorily uncontroversial answers to the problems that spurred its development, the model is considered "requisite" (Bana e Costa & Oliveira, 2012). It is also a recursive process since it may be necessary to rethink and adjust previous steps of the model to make the model "requisite".

Two distinct analyses can be performed by the M-MACBETH software: the *sensitivity analysis* that allows studying the impact of the overall score of the alternatives when the weight of a criterion is changed and the *robustness analysis* that allows analyzing the degree of dominance between pairs of alternatives according to variations in the scale and weight of a criterion (Bana e Costa et al., 2002)

In this way, it is possible to develop recommendations for the problem at hand by evaluating the respective global scores of each alternative and also based on the conclusions drawn from the two analyses described above, sensitivity analysis and robustness analysis.

2.7. Web-Delphi Platform Comparison Studies

During the development of the literature review, it was concluded that there is a significant absence of articles that perform a comparative analysis on WDP. Most of the comparative studies related to Delphi are studies that only compare the different types of Delphi. There is for example the article by Gnatzy et al. (2011) that provides a detailed comparison between Real-Time Delphi and the Conventional/Traditional Delphi Method, however, does not compare different WDPs.

On the other hand, the article that comes closest to the goal of the present dissertation is one by Aengenheyster et al. (2017) that performs a comparative analysis of 4 different online tools designed to implement Real-Time Delphi processes:

- Risk Assessment and Horizon Scanning (RAHS)
- eDelfoi
- Global Futures Intelligence System (GFIS)
- Calibrum Surveys - Surveylet

In turn, as previously mentioned, Real-time Delphi is not the same as the Traditional Delphi Method. Real-Time Delphi has the distinctive feature of providing immediate feedback to the participant from the first round of the process. While in the Traditional Delphi Method, feedback is only given from the second round of the process and is not immediate.

Returning to the proposed objectives of the present dissertation, is reviewed that the main objective is to perform a comparative analysis of all WDP that implement Delphi processes (regardless of the type of Delphi applied). Therefore, not only will be included platforms that implement Real-Time Delphi but also platforms that implement any other type of Delphi as long as it meets its 4 key features: Anonymity, Iteration, Controlled Feedback, and Statistical Analysis.

2.7 Literature Review Conclusions

The Delphi method has been growing a lot, and consequently, it is becoming widely used nowadays. Although its emergence has raised many doubts and criticism, its practice and application have evolved to fight many of its limitations.

We then began by analyzing its importance and origin, as well as its first uses. Delphi method was mainly used in matters of problems with high complexity, which required the intervention of

experts. Since they were complex problems in specific areas where subjective opinion was needed, experts were almost indispensable to ensure a more viable group decision. However, although the Delphi method is associated with the use of experts, it can be applied by any group of people who want to benefit from the advantages of this process to make a group decision.

After an extensive review of the original Delphi method, it was concluded that this method is quite flexible and easily adapted. The fact that there were different decisions with different needs led to the formation of new variants of the original method. Yet, it is necessary to deliberate well what is considered Delphi, since the method has specific characteristics that have to be respected (anonymity, controlled feedback, multiple iterations, and statistical analysis of group response). As soon as one of these variants jeopardizes one of these characteristics, it is necessary to consider its association with the Delphi method. However, the appearance of these variants allowed the reconfiguration and development of the method to mitigate a series of limitations associated with it.

The main limitations found that deteriorate the practice of this method are mainly associated with the time of its application. Long periods in the collection of responses and subsequent analysis ended up making the process inefficient and often caused other problems, such as instability in the panel of participants who often eventually gave up and left the process in the middle. The main way to conduct a Delphi process was through the mail, which in turn guaranteed anonymity among participants and was also a way to get answers from a distance. However, the associated time was also high.

In response to these limitations, the implementation of the Delphi method in digital format emerged. The evolution from paper-based Delphi method to online, contributed to the limitation of a series of problems, especially the time it takes to apply a Delphi process. At first, existing platforms were used, such as e-mail, survey platforms, and others, however, as none of these had been created to implement the Delphi method, they were not totally efficient. Over the last decade, several specific software has been developed for the implementation of Delphi processes and have been an asset to its efficiency. However, some questions are raised regarding the use of these platforms and there is a significant absence of studies that address these questions.

Each WDP is different and not all of them have the same characteristics and features. Depending on the features, some of them make the implementation of Delphi processes more efficient and easier for the moderator and the participants. The present dissertation is based on the goal of studying the emergence of these WDPs through a comparative study between them. This comparative study is achieved by applying a framework capable of registering the presence or absence of these features in the respective platforms. This approach allows the analysis to be objective and succinct, where the reader can easily understand what each WDP includes and what each one allows to do.

Additionally, an MCDA will also be proposed following the MACBETH approach that allows for the construction of value scales to numerically score each WDP. This model aims to help a

decision-maker or group of decision-makers in decision-making processes regarding the choice of a WDP to implement online Delphi processes. The particularity of this model is that its construction will be based on the previously mentioned framework.

The next section describes in detail the methodology applied in the present dissertation.

3. Methodology

3.1. Introduction

This section describes the methodology proposed to address the problem in study, based on the literature review that was exposed in section 2. Reviewing the contents of section 1.2., it is reminded that the present dissertation aims to perform a comparative analysis between the different available WDPs and explore how their features can advance the use of the Delphi method . The literature review allowed us to reach two important conclusions:

- There is a huge scarcity of scientific articles that address the topic of WDP
- There is poor knowledge about which WDPs exist at the moment and which ones are available for use

With these two conclusions in mindt, this study aims not only to report which WDPs are currently available for use, but also, what each platform offers in terms of features and functionality, taking into account their benefit in the implementation of Delphi. Additionally, since there is little knowledge about the WDPs, the following research question is also raised:

- Which WDP should a decision-maker choose to implement an online Delphi process?

Based on this question, this study also aims to help decision-makers choose which WDPs to use when implementing Delphi processes online.

3.2. Methodology Overview

Taking into account the objectives proposed in the previous section (section 3.1.), the following methodology represented in flowchart format in figure 4 was built:

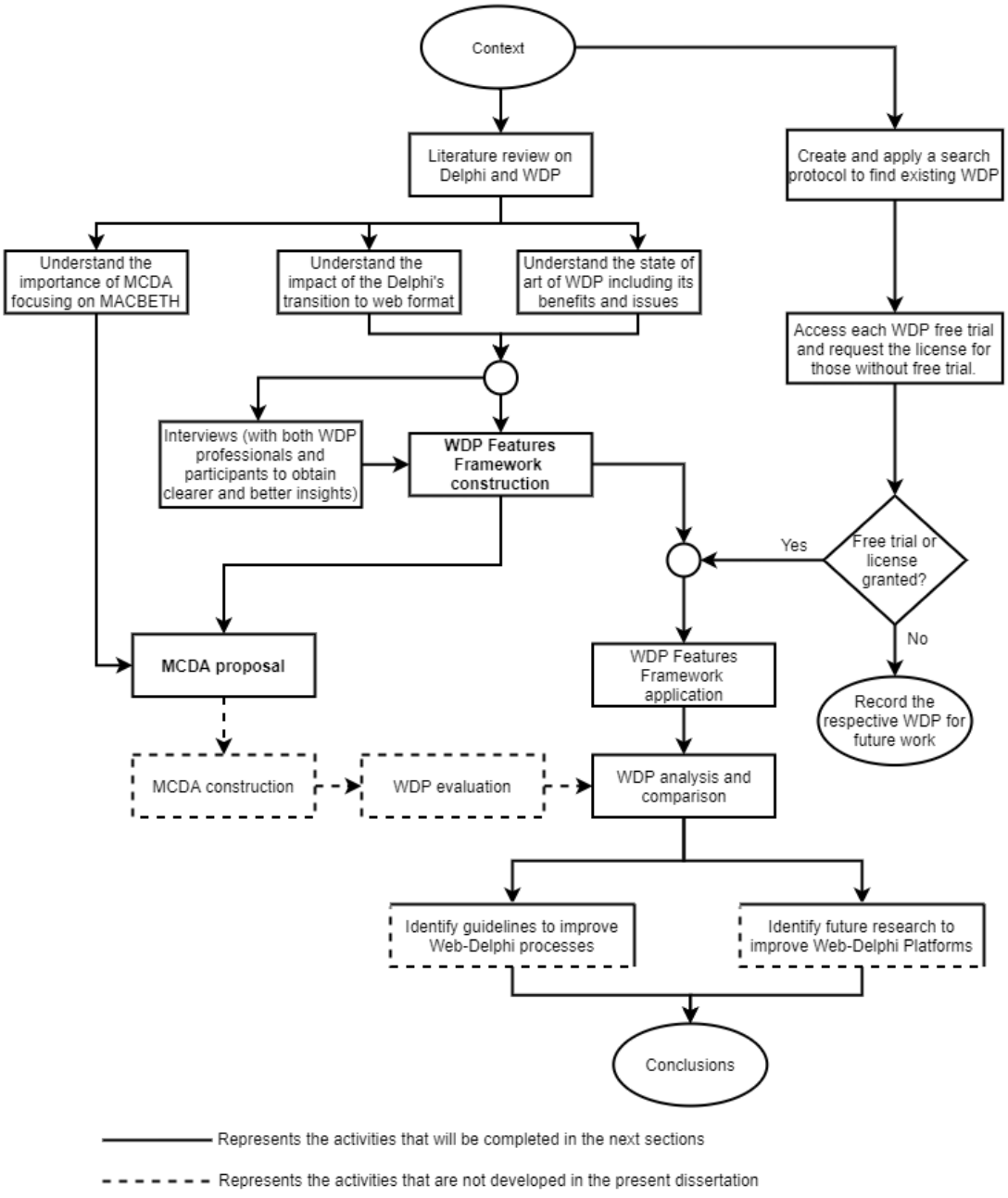

Figure 4 - Dissertation Methodology in Flowchart Format

This methodology is mainly focused on two different phases:

- Construction and application of a framework capable of performing comparative analysis on the different WDPs.
- Propose a multi-criteria model using the previous framework as its basis and that can evaluate the WDPs.

These two phases are represented in bold in figure 4 as "WDP Features Framework Construction" and "MACBETH model proposal" respectively.

3.3. WDP Features Framework construction

The WDP Features Framework is the framework that was created to develop the comparative analysis between the different WDPs. The framework is based on the evaluation of the presence of features in the respective platforms. It works as a checklist in which the presence or absence of the features in question is recorded in the respective WDPs. The use of this framework allows the analysis to be simple and objective and easily understood by the reader. For instance:

Table 2 - WDP Features Framework application example

Feature	Feature description	WDP.1
Qualitative data	Platform enables qualitative data collection	Yes

In this example, the criterion being evaluated is "Qualitative data" which corresponds to the fact that the platform has the ability to collect qualitative data. In this case, the platform that is being evaluated is WDP.1 (a fictitious platform that was created just for this example) and since it is filled with a "Yes", it means that the WDP.1 platform has the ability to collect qualitative data in its Delphi processes. In case any Platform that is being evaluated does not have this capability, then a "No" will be entered instead of a "Yes" in the evaluation of that platform regarding this Feature.

As shown in Figure 4, the construction of the WDP Features Framework is fed by two distinct activities: the literature review and the interviews. The literature review allowed not only to understand the impact of the transition of the Delphi Method to Web format but also the state of the art of it, pointing out its issues and benefits. However, in addition to the information gathered from the literature review, semi-structured interviews were also conducted in order to gain additional insights regarding the implementation of online Delphi processes and their participation. Thus, semi-structured interviews (McIntosh & Morse, 2015) were conducted not only with people who have implemented Delphi processes in WDPs as moderators but also with people who have participated in Web-Delphi processes as participants, to gain additional insights regarding WDPs.

Therefore, all the information that was gathered in the literature review (especially on WDP shortcomings) plus the insights gained from the interviews with the WDP professionals and participants, served as a basis for the framework construction. The framework is the structure for comparison between the differents WDPs that are studied in this dissertation.

3.3.1. Interviews

The semi-structured interview technique was used to conduct the interviews. This type of interview consists of the realization of pre-defined questions which in turn are adapted according to the answers of the interviewee. Conducting a semi-structured interview has the advantage of bringing a more real and detailed perception of the interviewee regarding the subject (McIntosh & Morse, 2015).

A total of 4 Interviews were handled in two different approaches:

- Two interviews were conducted with people who had already participated in Delphi processes conducted in a WDP, to have a better understanding of the process from a participant perspective.
- Two interviews were also conducted with administrators/moderators who have implemented Delphi processes in WDP to have a better understanding of the process from a back-office perspective.

The questions asked in the interviews were based on the literature review of the present dissertation and also based on additional research from survey platforms and WDPs websites information. The selection of interviewees was made from contacts provided by Professor Teresa Rodrigues, who works with one of the platforms studied in the present dissertation (Welphi).

3.3.1.1. Data Documentation

In order to filter out the most important information, a data documentation process was developed based on Flick (2019). Flick supports that in conducting oral interviews there are three main steps: **Recording the data**, **Editing the data**, and **constructing a "new reality"** using the acquired data.

Recording the data: In this case, since the interviews were conducted orally, it was necessary to record them and then transcribe them. Due to the Covid-19 pandemic, the interviews were all conducted through the Zoom platform, which in turn has a feature that allows the recording of the interviews. Before starting the interviews, each interviewee was requested to authorize the audio recording of the interview to facilitate and optimize the data collection process. Since the interviews were recorded, it made the transcription process much easier and more efficient so that all relevant information is properly registered. Before beginning the interviews, was also told to each interviewee that all information taken from the interviews would be presented in the dissertation anonymously, which means that the identity of all interviewees will be anonymous to the reader. The confidentiality that anonymity ensures, allows the interviewee not to be associated

with their answers, thus mitigating possible constraints that may compromise and influence the interviewee's opinion.

Editing the data: After the interviews were recorded, it was necessary to transcribe them. Since this is an oral interview, sometimes the interviewees digress a bit in their answers and therefore the transcript of each interview was edited in order to filter out only the important information. Sometimes the interviewees answered questions that had not been asked yet, and as such, it was necessary to manage the interviewees' answers according to what was asked in each question, i.e., some answers were moved to later questions and there were questions that were omitted and others that were adapted as the interviews proceeded.

Constructing a "new reality": Once the data is transcribed and edited, it is necessary to summarize the data to make the qualitative analysis more objective and succinct. To do this, the questions were divided and grouped into different categories and within each category, all the important information from the interviews was summarized and paraphrased. Then the respective conclusions were drawn for each category, which served as support in the construction of the WDP Features Framework.

Regarding the "**Participant Perspective**", the questions were grouped into 8 different categories to facilitate the course of the interview:

- **Category AP**: Participant background.
- **Category BP**: Communication between the moderator and the participant during the process.
- **Category CP**: Registration and Log-in.
- **Category DP**: Learning and training.
- **Category EP**: Platform User-Friendliness.
- **Category FP**: Feedback delivery and interpretation.
- **Category GP**: Anonymity during the process.
- **Category HP**: Navigation.

As for the "**Moderator's Perspective**", the questions were grouped into 9 different categories:

- **Category AM**: Moderator background.
- **Category BM**: Choose participants and invitation.
- **Category CM**: Questionnaire construction.
- **Category DM**: Learning and testing.

- **Category EM**: Rounds and stopping criteria.
- **Category FM**: Feedback delivery and interpretation.
- **Category GM**: Communication between the moderator and the participants.
- **Category HM**: Anonymity during the process.
- **Category IM**: Data collection, interpretation, exhibition, delivery, and exportation.

Having these categories well defined facilitates the course of the conversation, making it simpler to document the interview. Once the literature review and interviews were completed, the WDP Features Framework was constructed.

3.3.2. Develop a Search Protocol

To apply the WDP Features Framework and compare all the available WDPs it is first necessary to search and identify which WDPs are currently available for evaluation. For this reason, a search protocol has been created to find all currently available WDPs. The following steps have been set in the construction of the protocol:

1. Define the search question.
2. Identify key concepts in the search question.
3. Identify synonyms of the search question keywords and identify other keywords, phrases, or terms.
4. Identify the string to use on each research place.
5. Identify the place where the research will be conducted.
6. Identify Rejection Criteria.

Define the search Question

A Search Question has to be clear and objective in order to understand its purpose easily and quickly. As such, the following question has been registered as the Search Question for this search protocol:

Table 3 - Search Question

Search Question
Which Delphi online platforms are currently available?

Identify key concepts in the Search Question

Since the Search Question is already registered, the most relevant Keywords have to be retrieved from it. The following words were registered as Keywords:

Table 4 - Search Question Keywords

Search Question Keywords
Delphi
Online
Platforms

Identify synonyms of the search question keywords and identify other keywords, phrases, or terms

Once we have the Keywords registered, it is necessary to identify their synonyms and alternative terms. This step is very important since not all documents use the same words for the same meanings. The following Table represents the alternative terms and synonyms found for the Keywords in question:

Table 5 - Identified synonyms of the search question as well as other keywords, phrases or terms

Keywords	Synonyms	
Delphi	**OR**	Delphi Method
		Delphi Process
		Delphi Technique
		Delphi Survey
AND		
Online	**OR**	Internet*
		Web*
		Computer*
		Network*
AND		
Platforms	**OR**	Program
		System

		Tool
		Software

Identify the string to use on each research place

Based on the previous Table, it was possible to create a String of words that facilitates the application of the search protocol in the different research places. The String created and used was the following:

Table 6 - Research String used

String
"Delphi" OR "Delphi Method" OR "Delphi Process" OR "Delphi Technique" OR "Delphi Survey" AND "Online" OR "Internet*" OR "Web*" OR "Computer*" OR "Network*" AND "Platform" OR "Program" OR "System" OR "Tool" OR "Software"

Identify the place where the research will be conducted

In order to conduct the research effectively and efficiently, it is necessary to conduct it in different databases and search engines. As such, the following research places were used to conduct this research:

Table 7 - Research Places used in the WDPs research

Research Places
Google
Google Scholar
Web Science
Scopus

Identify Rejection Criteria

Initially, it was assumed in the present dissertation that a platform is considered as WDP if it were exclusively designed with specific questionnaires to implement Delphi processes. Therefore, we will only consider the currently available platforms that comply with this criterion.

These are the criteria that a platform must have to be included in this study:

- The platform was exclusively designed to implement Delphi processes.
- The platform is currently available and accessible to the public.

- The platform is able to fulfill all the key features of the Delphi method during a Delphi process.
- The platform software is designed and optimized for use on a computer (although in some cases it may also be optimized for use on Tablets and smartphones).

3.3.3. License Requests

Once the platforms were selected, it was necessary to apply for an access license on the platforms that were not free and also on those that did not have a free trial. For platforms that did not meet these two standards, the following alternative methods of analysis were used:

- User guides were requested to provide information regarding the functioning and features of the platforms.
- Video calls were made with members of the platforms' team and e-mails were exchanged to clarify doubts about the functionalities of the platforms.
- Video tutorials were requested.

The platforms on which it was not possible to apply the previous analysis methods, thus making it impossible to apply them to the WDP Features Framework, have been recorded and proposed for future work.

3.3.4. Section Conclusions

The knowledge gathered from the literature review and the insights from the interviews were used to build the WDP Features Framework. A search protocol was constructed to find and select the eligible platforms for analysis and free licenses were also requested to run some of the selected WDPs.

3.4. Multi-criteria Evaluation Model (MCDA) Proposal

3.4.1. Introduction

Recalling the research question that was raised in section 3.3.2, one of the purposes of building this model is to help possible Decision Makers choosing a WDP to implement Web-Delphi processes. Since choosing a platform that implements a very specific and particular method (such as the Delphi Method) is a rather complex process where several rules, concerns, and criteria

need to be considered, a multi-criteria decision aid approach was chosen to tackle this issue. This MCDA not only serves to help in choosing a WDP, but also helps to evaluate and compare each WDP. This section presents the proposed MCDA model following the MACBETH methodology that allows the calculation of the overall value associated with each WDP, thus facilitating the process of choosing one of them.

Based on the documents from Bana e Costa et. al (2008), Bana e Costa et. al (2010), and Bana e Costa, Carnero & Oliveira (2012), the following methodology was adapted to apply the MACBETH model to the underlying problem of the present dissertation.

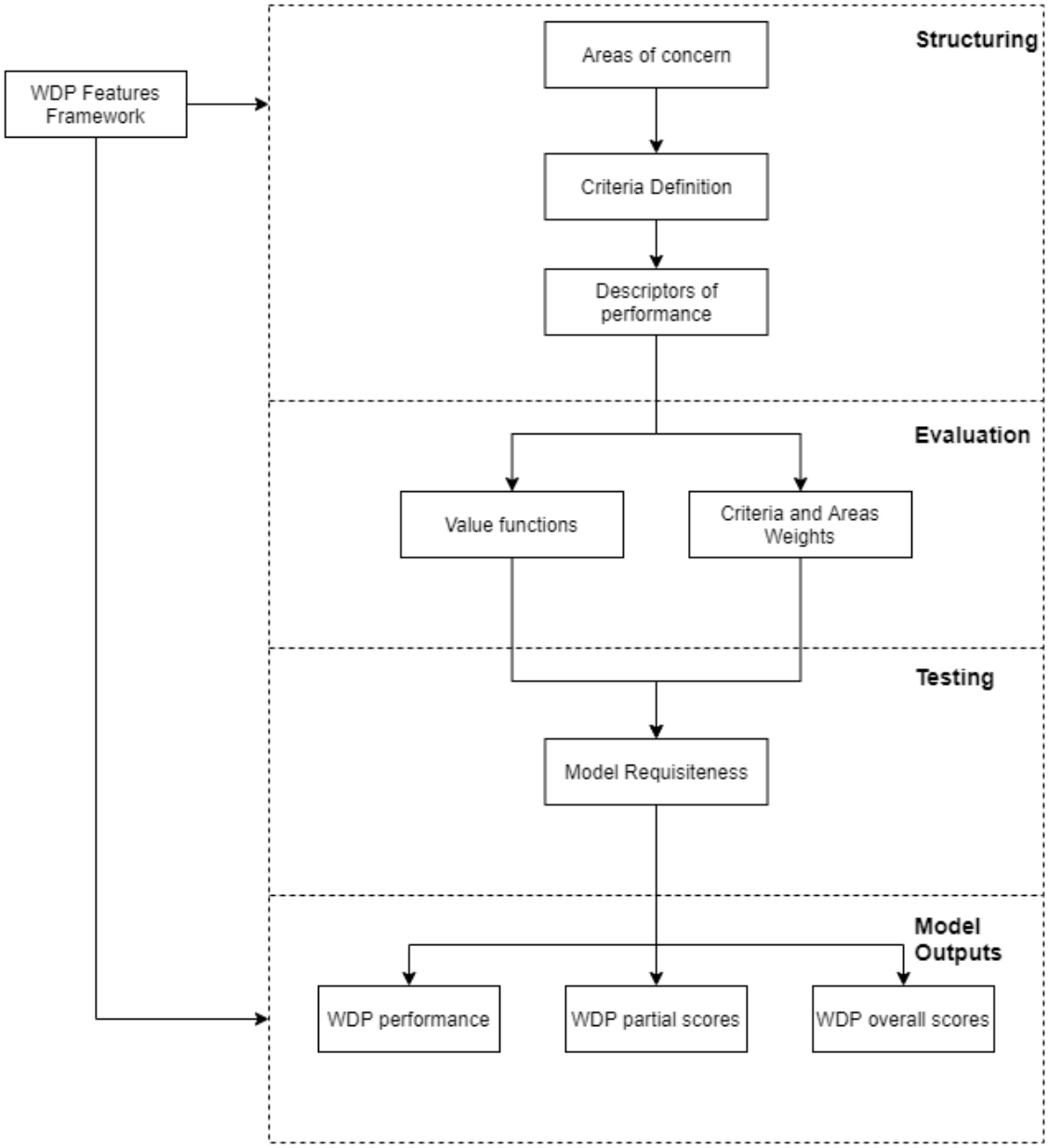

Figure 5 - MCDA proposed methodology

Once the WDP Features Framework is built, it is possible to proceed with the construction of the proposed multi-criteria model. As we can see in figure 5 the model building process is divided into four different parts: Structuring, Evaluation, Testing, and Model Outputs. In each of the parts, different activities are performed.

The construction of this model is achieved through a socio-technical process where a social component and a technical component are present. Since the model will only be proposed, this dissertation focuses more on the technical component than on the social component. Starting with **Structuring**, this is where the areas of concern and the respective evaluation criteria are defined. In addition, it is also in this part where the descriptors of performance for each evaluation criterion are built, thus allowing these criteria to become operational. In the **Evaluation** phase, we have the construction of value scales and the relative determination of the weights of each evaluation criterion and each area of concern.

Once the previous steps are completed and the model validated as requisite, it is then possible to evaluate each WDP performance and also to calculate the **partial value score** of each WDP in a specific area of concern and the **overall score** of each WDP, using the **simple additive model** and the **hierarchical additive value model,** respectively (Bana e Costa, Carnero & Oliveira, 2012).

$$v_h(x_{1_h}, \dots, x_{j_h}, \dots, x_{n_h}) = \sum_{i=1}^{n} k_{j_h} v_{j_h}(x_{j_h}) \ with \begin{cases} v_{j_h}(good_{j_h}) = 100 \\ v_{j_h}(neutral_{j_h}) = 0 \end{cases} \qquad (1)$$

$$v(WDP) = \sum_{h} \left(k_h \sum_{j_h} k_{j_h} v_{j_h}(x_{j_h}) \right) \qquad (2)$$

Where:

- $h = 1, \dots, m$ designates the areas of concern;
- $j_h = 1_h, \dots, n_h$ designates the criteria of area h;
- X_{j_h} corresponds to the descriptor of performances of criterion j_h and $v_{j_h}: X_{j_h} \rightarrow \mathbb{R}\ ou\ \Re$ designates the respective value scale;
- x_{j_h} designates the performance of alternative x on criterion j_h;
- $v_{j_h}(x_{j_h})$ designates the partial value score of the alternative x on the criterion j_h;
- $(x_{1_h}, \dots, x_{j_h}, \dots, x_{n_h})$ designates the performance of alternative x on area of concern h;
- $(v_{1_h}(x_{1_h}), \dots, v_{j_h}(x_{j_h}), \dots, v_{n_h}(x_{n_h}))$ designates the value profiles of alternative x;

- $good_{j_h}$ and $neutral_{j_h}$ are the “good” and the “neutral reference levels of performance on criterion j_h;
- k_{j_h}designates the weight of criterion j_h;
- k_h designates the weight of area h;
- $\sum_{j_h} k_{j_h} = 1$ and $k_{j_h} > 0$;
- $\sum_h k_h = 1$ and $k_h > 0$.

The Equation (1) allows to calculate the score of each WDP on each area of concern separately while the Equation (2) allows the DM to understand which WDP has the best overall score considering all the criteria and areas of concern involved in the process. Thus, it is possible not only to analyze each platform regarding each criterion and each area of concern, but also to globally evaluate each platform regarding all areas and criteria.

Next, the **Testing** phase is where it is understood whether or not the model is sufficient to answer the questions that led to its development. In other words, this is where it is determined whether the model is considered "requisite" or not. *Sensitivity analysis* and *robustness analysis* are also performed at this stage in order to test the requisiteness of the model.

Finally, the respective model results displayed in the **Model Outputs** are obtained. These results will support the development of recommendations to evaluate each WDP in more depth. With these results it will be possible to evaluate each platform regarding each criterion, regarding each area of concern, and finally regarding all criteria and areas overall. It allows to know which platforms are more developed in each criterion and in each area of concern.

The next subsections provide a detail description of the different steps that are taken to construct the multicriteria model based on the steps presented in the previous figure 5 and based on the WDP Features Framework.

3.4.2. Problem and Model Structuring

Once the Decision Maker is identified and the problem is defined, we can proceed to the problem structuring which consists in the definition of the different alternatives and the consequent analysis of the characteristics of each alternative in order to understand which are the relevant characteristics to the decision maker. In this case, the alternatives consist of the WDPs that are included in the WDP Features Framework, and the characteristics consist of their features.

The first phase in structuring the model consists of identifying the areas of concern as well as the evaluation criteria to be used in the model. As shown in figure 3, the particularity of this model's

construction is that the definition of the areas of concern and the evaluation criteria can be based on the areas of concern and features present in the WDP Features Framework. It is suggested to directly use the areas of concern defined in the WDP Features Framework (*Anonymity*, *Iteration*, *Controlled Feedback*, *Statistical Aggregation*, *Implementation Features,* and *Participant User-Friendliness*) as the areas of concern in the proposed MCDA. It is possible not only to define the areas of concern but also the evaluation criteria by using the features and their descriptions that are present in the Framework as a basis. This way we can build a value tree using the WDP Features Framework.

Given the previous suggestion of using the WDP Features Framework to define the areas of concern and the respective criteria, it is recommended to use the **bottom-up approach** since the areas and their features are already defined in the Framework.

3.4.2.1. Descriptors of Performance

Before constructing the value functions, it is also necessary to define the descriptors of performance of each criterion. The definition of these descriptors is necessary to make the criteria operational to allow the evaluation of the attractiveness of the alternatives (Bana e Costa e Beinat, 2005). The descriptor levels can be quantitative or qualitative (or mixed) and sometimes can be conceived (visual presentations, verbal descriptions, etc.).

The descriptors of performance correspond to levels of impact of each of the defined criteria and these impact levels allow the evaluation of the attractiveness of the performance of the alternatives in each defined criterion. To construct the descriptors of performance is necessary to define the reference levels “Neutral” and “Good” of each criterion. The DM is asked what he considers to be “Good” and what he considers to be “Neutral” in each criterion and then these levels will serve as reference levels to construct the value scale for each criterion, where the “Neutral” level is assigned with a score of 0 and the “Good” level is assigned with a score of 100. Again, the WDP Features Framework can serve as a basis in the construction of these descriptors. A criterion can contain several entries from the Framework and since each feature is accompanied by a description, it becomes easier to create the respective descriptors of performance for each criterion.

3.4.3. Model Evaluation

3.4.3.1. Value functions

Value functions consist of functions that allow transforming performance into value, thus making it easier to perceive the differences in the attractiveness of the levels defined in each descriptor.

The construction of these value functions can be done from numerical techniques or non-numerical techniques and the proposed model follows the **MACBETH** approach which is a non-numerical technique. Following the MACBETH approach, the construction of a value scale for each of the criteria is based on the reference levels “Neutral” and “Good” of each one. Having these reference levels well defined in each criterion, the DM is then asked to rank the WDP and the respective reference levels in order of their attractiveness on each criterion. Next, using the semantic scale “null”, "very weak", "weak", "moderate", "strong", "very strong" or "extreme" the DM elicits qualitative judgments regarding the difference in attractiveness between the performance levels of the descriptors, in pairs, at the same time. If consistency is checked by the software M-MACBETH in the elicited comparisons, then through linear programming, the **qualitative judgments** of the decision-maker are **converted** into **numerical values** where a score of 100 is given to the most attractive alternative and a score of 0 is given to the least attractive alternative. Once the value scale is built, the DM has to validate it, otherwise, it is necessary to make alterations and repeat these steps (Pereira Dias & Fontes, 2019).

3.4.3.2. Criteria and Areas Weights

Once again, in this model, the non-numeric MACBETH method is used to calculate the weights of the criteria and the weights of the respective areas of concern. Following the same methodology used in Bana and Costa, Carnero & Oliveira (2012), it is suggested that the weights of criteria and areas are weighted hierarchically. One of the advantages of following the MACBETH methodology is that the M-MACBETH software has a version that allows to assign the weights hierarchically (Costa, Lopes & Oliveira, 2014). To calculate criteria weights, decision makers are asked to compare pairs of criteria with respect to the improvement from "neutral" to "good" for each criterion (swings) based on the semantics previously mentioned (very weak, weak, moderate, strong, very strong, extreme). After performing consistency checks for all pairwise comparisons, through linear programming the weights of each criterion are then calculated, with 0 corresponding to the weight of the neutral option and 100 the sum of the weights of all the criteria. This calculation is done using the M-MACBETH software and results in the normalized weighting coefficients for each criterion. As long as they do not affect the consistency of the judgments provided, these coefficients can still be adjusted by the decision-maker. For instance, if a criterion A has a weight of (0.22) and a criterion B has a weight of (0.11), it means that the DM value the improvement of criterion A from "neutral" to "good" twice as much as the improvement of criterion B from "neutral" to "good" (It must be emphasized that the statement found in the previous example is only relative to the reference levels that have been defined for this case, i.e., the fact that the swing of criterion A is twice as attractive as the swing of criterion B is relative to these reference levels that have been pre-defined). Once the criteria weights have been calculated, the process proceeds to calculating the weights of the areas of concern. This requires selecting the criterion with the highest score in each area and applying the weighting

process described earlier. New scale constants are then generated and used to scale the initial weights of the criteria. Finally, these weights are introduced within each area in order to derive the weights across areas (Bana and Costa, Carnero & Oliveira 2012).

3.4.4. Model Requisiteness

Once the model is built, it is necessary to understand whether the model meets its predefined objectives. The model will only serve as an aid for a decision-maker or group of decision-makers to think and learn as a group about the issues related to the problem and in light of these thoughts be able to choose a WDP. There is no optimal solution or "right answer". However, the model needs to be requisite in a way that is sufficient to solve the problems at hand.

A model is a tool that allows capturing the views of the decision-makers and it can be used as a means to investigate the influence of changes in perspective or data ambiguity (Phillips & Bana e Costa, 2007). Therefore, to complement the analysis, it is possible to perform two distinct analyses using M-MACBETH, the sensitivity analysis, and the robustness analysis.

- Starting with the **sensitivity analysis**, this analysis allows the DM to study the impact on the overall scores of the alternatives when the weight of a criterion is changed. In short, the software allows you to change the weight of a criterion between 0 and 100% (inclusive) and see the overall score of each WDP given this change (Bana e Costa & Chagas, 2004).
- The **robustness analysis**, on the other hand, consists of an analysis that allows evaluating the dominance relationships between the different WDP. Basically, the robustness analysis allows us to observe the dominance relationship between two WDP, where only one of the following 3 can be found in each pair: dominance, additive dominance, or non-dominance. The software M-MACBETH can generate pictorial information regarding these relationships where *dominance* = red triangle, additive dominance = green plus, and non-dominance = question mark. Additionally, through robustness analysis, it is possible to determine whether the outcome of the best alternative changes when varying the scale and weight of each criterion under a predefined percentage (Pereira, Dias & Fontes, 2019).

3.4.5. Section Conclusions

Besides allowing a comparative analysis between WDPs in terms of the presence of features, the WDP Features Framework also enables the construction of an MCDA to evaluate the performance of each WDP. It was concluded that a multicriteria model capable of calculating the

overall value associated with each WDP performance can be built and used to evaluate each platform and support the decision process of choosing a WDP to implement eventual Web-Delphi processes. In this case, the MACBETH approach was proposed, which can be implemented from the M-MACBETH software described in section 2.6.3. of the literature review. This software not only facilitates the execution of the steps mentioned in figure 5 but also allows performing sensitivity and robustness analyses that strongly contribute to the decision process.

4. Application of the Methodology

First, in the present section, the conclusions obtained in the semi-structured interviews will be presented, which in turn were used as insights in the construction of the WDP Features Framework. After the construction of the Framework, the search protocol built in section 3.3.2. was applied, which allowed finding and selecting the WDPs that were included in the comparative analysis.

Regarding the multi-criteria model that was exposed in section 3.4. it is reminded that it is only a proposal that can be used in future work, and therefore, it will not be built in the present dissertation. The information presented below are exclusively concerning the construction and application of the WDP Features Framework.

4.1. Framework Construction

4.1.1. Interviews conclusions

After the semi-structured interviews were conducted and documented, the most important conclusions were drawn in order to serve as a basis for building the WDP Features Framework. To ensure the anonymity of each interviewee in the present dissertation, participants are represented as, Participant X (PX) and Participant (PY), and moderators as Moderator X (MX) and Moderator Y (MY). Due to the limits imposed on the number of pages of the dissertation, the editing, and analysis of the transcribed data from the interviews have been submitted in section A (Appendix), and only the conclusions of each category of questions will be presented in the body of the dissertation.

Starting from the participant's perspective, the following Table 8 represents the conclusions for each category that were drawn in the interviews:

Table 8 - Conclusions from the participant's perspective interviews

Category	Description	Conclusions
AP	Participant background	**PX**: Participant with extensive experience in Web-Delphi, both as participant and moderator. This participant started working with Web-Delphi in 2015 and has been a participant in at least two studies.
		PY: Participation in only one Web-Delphi process, but nevertheless studied and also designed a Web-Delphi process for a course in college.
BP	Communication between the moderator and the participant during the process	-Communication should be done by e-mail. -Live-chat is a value feature to support participants during the process. -Reminders are important to ensure participant's engagement.
CP	Registration and Log-in	-The registration process and the log-in should be simple enough to not fatigue the participant
DP	Learning and training	-It is important to have an instruction page that explains in a clear way, the procedure of the process in question. -Having a short tutorial that interactively explains the process was also considered important. However, it is a bit difficult to put into practice because the questionnaire designs vary from study to study, and this implies having to make a custom tutorial for each study.
EP	Platform User-Friendliness	-The platform should be Tablet-friendly and especially mobile-friendly. -It is very important to have an auto-saving system for the answers and for all the progress made in the questionnaire. -It is important that once the quiz is not completely finished, the platform indicates which questions are left to complete and what percentage of the quiz the participant is currently on. -Each question should contain a short instruction to ensure that it is answered properly. -A progress bar while completing the questionnaire is very important for participants to know what stage of the questionnaire they are at.
FP	Feedback delivery and interpretation	-There are three easy ways of receiving the feedback: Written in the email itself, on a document attached to the e-mail and along with each question. -Feedback along with each question is valued. -Visual outputs such as bar graphs, pie charts, graphics, etc. are important to interpret better the results. -Answers from the previous round already typed in the subsequent round is also valued.
GP	Anonymity during the process	-Anonymity can be easily guaranteed if a specific ID is assigned to each participant as soon as they register on the platform. -There must be a proper control by the moderator in the comment sections to prevent possible threats to anonymity.
HP	Navigation	-It is important to have the possibility to build a questionnaire on a single page, but also to have the option to be able to split the questionnaire into different pages. -The moderator has to consider whether or not it makes sense for the questionnaire to be on one page or split into different pages.

As previously stated, interviews were also conducted with moderators who have implemented Delphi processes in WDPs. In the following Table 9, the respective conclusions for each category from the moderator's perspective, which were drawn throughout the interviews, are represented.

Table 9 - Conclusions from the moderator's perspective interviews

Category	Description	Conclusions
AM	Moderator background	**MX**: Has been working in the Web-Delphi area since 2016 by being involved in several Web-Delphi processes. Was involved as a moderator in only one process but supported other processes not only in the design but also in the analysis of results.
		MY: Only moderated one Web-Delphi process
BM	Choose participants and invitation	-Choose heterogeneous participants who have interest in the process outcome. -Post the project in social networks and investigate who is publishing scientific articles related to the process subject to collect data of possible participants and invite them.
CM	Questionnaire construction	-Moderator has to consider that the platform may not be Tablet-friendly or mobile-friendly. It is advisable to test first and see if it is necessary to give a warning in the convocation regarding this issue. -It is important to have both options: possibility to create pages from scratch and the possibility of using pre-defined themes to construct the questionnaire. -Background customization, insertion of logos, images, and instructions along with questions is important. -Possibility of breaking the questionnaire into separated pages is also valued. -Adding a comment section along with each question is a good practice.
DM	Learning and testing	-Video tutorials are important to learn how the platform works. -Is important to do a pre-test to see if everything is working as expected.
EM	Rounds and stopping criteria	-The stopping criteria should be decided before the process starts, however, it can be changed according to the outcome of each round. -Platforms should have a feature where moderators can manage who has responded, who is responding and who has not responded, in order to manage the process efficiently.
FM	Feedback delivery and interpretation	-The fact that the platform has an intuitive reminder feature where moderators can know the current status of each participant's activity is valued. -Spam box is a possible threat to participant engagement: One way to mitigate this is to send e-mails from the moderator's personal address. -One way to improve engagement is to call each participant to ensure their participation in the process. -Questionnaires should be done on participants' native language, and if possible, should be available an option where the participant could choose the language of the questionnaire.
GM	Communication between the moderator and the participants	-Communication should be done by e-mail. -Live-chat is a value feature to support participants during the process. -Moderators should explain well the process in the beginning to avoid possible doubts.
HM	Anonymity during the process	-Participants should not have way of talking with each other. -Every participant should have a unique ID to ensure anonymity and to track their responses.

IM	Data collection, interpretation, exhibition, delivery, and exportation	-There must be a way to ensure that each participant only have a submission per round. -Platform should have an integrated feature for data exportation to perform detailed analysis. -The Platform should perform an automatic analysis of the results and display them in graphical form, such as the distribution of responses in bar charts and not only in numerical format. -Moderators should send a detailed report with the feedback to ensure that each participant understands the aim of the process and the respective feedback.

With the literature review concluded and the interviews conducted and documented with the respective conclusions, the WDP Features Framework was then constructed to perform the comparison between the different platforms.

4.1.2. WDP Features Framework Construction

The framework was built and first divided based on the previously mentioned key criteria of the Delphi Method: Anonymity, Iteration, Controlled Feedback, and Statistical Aggregation. Thus, based on the findings of the interviews and the literature review, a set of features/characteristics were created and divided among those four areas of concern (Delphi Key features). Since not all of the features that have been registered fit within these four areas, two more were created making a total of six areas of concern:

- Anonymity
- Iteration
- Controlled Feedback
- Statistical Aggregation
- Implementation Features
- Participant User-Friendliness

Below is a detailed description of what each area of concern encompasses:

- **Anonymity**

This area is particularly important since anonymity must be guaranteed in every Delphi process. Here it is not only analyzed whether the platform guarantees anonymity between the participants, but also whether the platform guarantees anonymity between the participants and the facilitator.

- **Iteration**

Nominated as iteration this area covers the features that make iteration efficient and effective in a Delphi process conducted in a WDP. Included in this area are features such as the way the questionnaire is shared with the participants, the way the participants are alerted by the facilitator, the monitoring of the process and management of the participants, and others.

- **Controlled Feedback**

It allows you to reinforce anonymity, decrease possible noise from non-relevant participant comments by delivering only the information relevant to the Delphi study in question. This area covers how the feedback is presented to the participants, whether or not their comments can be edited, what kind of data can be extracted in the questionnaire (quantitative and qualitative), whether feedback can be delivered in real-time, and other aspects.

- **Statistical Aggregation**

Statistical Aggregation is an area that covers how the collected data is generated and analyzed. Included are features such as data export, the platform's data analysis capabilities, and whether these analyses are done automatically or manually.

- **Implementation Features**

This area covers the features related to the implementation of a Delphi process in a platform. In other words, the features that are included in this area are relative to the construction and customization of the questionnaire, where not only the ease of questionnaire construction is addressed, but also the possibilities of customizing the questionnaire.

- **Participant User-Friendliness**

This area is crucial to ensure greater adherence of participants to Delphi processes. As already mentioned, the dropout of participants during the Delphi processes is one of the most characteristic problems pointed out to the Delphi method, and as such, it was necessary to create this area. As the name suggests, this area covers features relating to the ease of handling of the platform by the participant. This includes features such as how participants register on the platform, how instructions are entered into the questionnaires, whether the questionnaire has a progress bar where participants can see at what stage of the questionnaire they are at, how possible doubts can be clarified, and so on.

The WDP Features Framework is represented in the following Table 10. The framework includes the six areas of concern mentioned above, the features related to each area, and an objective and succinct description of each feature in order to be clearly understood by the reader.

Note: The features that are represented with an "*" at the end are features that were considered valuable by the interviewees.

Table 10 - WDP Features Framework

Features	Description
Anonymity	
Anonymity moderator	Participants' anonymity can be ensured between the participants and the moderator
Anonymity participants	Participants' anonymity is ensured within the participants
Iteration	
Process monitoring*	Platform is able to track each participant responses (e.g., grouping who did not respond, who did respond and who is responding at that moment)
Round replication	Moderator can replicate the previous round to create a new one (instead of creating another from scratch)
Unlimited rounds	Platform do not have limit of rounds per Web-delphi process (it is unlimited)
E-mail tracking	The moderator receives delivery reports of the e-mails he sends
Built-in E-mail system*	Platform has a built-in e-mail system
Personalized e-mails*	E-mails can be automatically personalized to each participant (e.g., "Dear 'Participant Name' ")
Reminders*	Platform has a built-in reminders option (automated specifically for reminders)
Link	Questionnaire can be shared by link (generated by the platform)
QR code	Questionnaire can be shared by QR code generated by the platform
Apps Share	Questionnaires can be shared through other apps via a share button from the platform that was specifically created for that purpose (Facebook, Twitter, LinkedIn, etc.)
Controlled feedback	
Quantitative data	Platform enables quantitative data collection
Qualitative data	Platform enables qualitative data collection
Qualitative data editing	Platform enables the editing of qualitative data for feedback control purposes (e.g., edit comments)
Visual presentation*	Statistical group response data can be presented in the form of charts in the platform itself
Real-time Feedback	Feedback can be delivered in real-time
Feedback along with the questions*	Platform has an integrated feature that allows the feedback to be displayed along with each question
Date Filter	Time interval for analysis can be set by choosing a start date and an end date. The collected data is only from that period even if the round has a longer duration
Individual respondents	Particular individual answers can be selected for individual analysis (Not only particular groups of stakeholders but also particular individual singular responses)
Grouping users	Platform allows the aggrupation of users in different groups of stakeholders (Allowing analysis per group of stakeholders)
Comment boxes*	Platform allows to add boxes along with each question for comments with multiline
Previous answers*	Answers from the previous round are already typed in the subsequent round
Statistical Aggregation	
Data generation*	Statistical group response data is automatically generated
Data analysis tool	Platform has an integrated feature that allows advanced data analysis besides the generation of statistical group response data (e.g., consensus measurement, stability of responses)
Data exportation*	Raw data can be exported to external statistical analysis tools

Implementation features	
Predefined themes*	Platform have predefined themes that help you to create questionnaires quickly
Page replication	Users can replicate a page to create a new equal one
Question replication	Users can replicate a question to create a new equal one
Auto-saving construction*	In the questionnaire construction the progress is automatically saved whenever there are any changes
Same page*	Different questions type and chapters of the questionnaire can be put in a same page
Page break*	Questions and chapters of the questionnaire can be separated on different pages
Welcome/End pages*	Platform allows to create a personalized welcome/end page in the questionnaire (e.g., instructions page)
Multiple languages*	Allows the creation of questionnaires with multiple languages (the respondents can choose their preferred language)
Questionnaire customization*	Allows to customize the questionnaire freely (e.g., colors, font, and font size)
Attach pictures*	Each page and/or questions can be attached to a picture or logo
Type of questions	Platform offers a wide range of type of questions (dropdowns, radio buttons, response scales, text boxes etc.)
Input validation	Users can set which type of input is valid for an open question (e.g., email/date/time/month/week/color/phone)
Set required fields	Particular questions can be set as required fields
Set other choices	Participants can have the ability to type a different answer from the ones available (in multiple choice for example)
Learning*	Platform offers a learning and support system with different learning alternatives (e.g., video tutorials, examples, demonstrations, user guides, etc.)
Questionnaire preview	The platform has a feature that allows you to preview the created questionnaire pages
Pretest*	Platform offers a pretest function
Import a list of participants	Platform allows the importation of a list of participants (csv files)
Participant user-friendliness	
Registration*	Participants' registration is mandatory
Registration/log in easiness*	The registration process is simple and non-exhaustive. Users only need to submit their e-mail and a password as mandatory fields
Instructions along with the questions*	Questions can have instructions along with each question
Progress bar/indicator*	Questionnaires have a progress bar/indicator where the participants can see at what stage of the questionnaire they are in
Auto-saving answers*	The platform automatically auto-saves participants' progress during the round
Device adaptation*	Site is adapted for each participant depending on the device they are using
Live Chat*	Live chat available to clarify doubts and solve possible issues in real-time (with human intervention)
Chat bot	Live chat available to clarify doubts and solve possible issues in real-time (without human intervention)
GDPR	Platform complies with EU General Data Protection Regulation
Submission warnings*	Platform can give warnings regarding how many questions that are left to be answer and what percentage of the questionnaire the participant is currently on

.

Since the framework is already designed and built, it is necessary to proceed with the application of the search protocol in order to find all the WDPs eligible for application in the WDP Features Framework.

4.1.3. Web-Delphi Platforms Encountered

After applying the previously presented search protocol string in the respective research places, the following platforms were found:

- Delphi Decision Aid by Aegir St Louis Integration (DDA) (http://armstrong.wharton.upenn.edu/);
- DelphiManager (DM) (www.comet-initiative.org/delphimanager/index.html);
- DEMOS – Delphi Meditation Online System (www.tuhh.de/tbg/Demos/information_delphi_method.html);
- eDelphi (www.edelphi.org/);
- ExpertLens (EL) (www.rand.org/pubs/tools/expertlens.html);
- Global Futures Intelligence System (GFIS) (www.millennium-project.org/rtd-general/);
- HalnyX Delphi Platform (https://4cf.pl/en/project/halnyx-delphi-platform/); (Real-Time)
- Live Delphi (LD) (https://play.google.com/store/apps/details?id=fi.metatavu.LiveDelphiApp&hl=pt&gl=US) ; (Real-Time mobile application)
- Mesydel (www.mesydel.com);
- Surveylet by Calibrum (www.calibrum.com);
- Tech Cast (TC) (www.techcast.org);
- UpBoard (upboard.io/delphi-method-online-tools-web-templates/);
- Welphi (www.welphi.com/en/Home.html);

A total of 13 platforms were identified as possible WDP. Nevertheless, the rejection criteria were applied to identify which WDP meet the requirements previously defined on the rejection criteria and a total of six were eligible for evaluation. Table 11 represents which platforms were found in each previously predefined research places.

Table 11 - WDPs found in each research place

	Google	Google Scholar	Web Science	Scopus
DDA	x	-	-	-
DM	x	x	x	x
DEMOS	x	x	-	-
eDelphi	x	x	x	x
EL	x	x	x	x
GFIS	x	x	-	-
HalnyX	x	-	-	-
LD	x	-	-	-
Mesydel	x	x	-	-
Surveylet	x	x	-	-
TC	x	-	-	-
upBoard	x	-	-	-
Welphi	x	x	x	x

Note: Fields represented with "x" mean that the platform was found in the research place corresponding to the column in question.

4.1.4. Web-Delphi Platform Selection

Even though the platforms found are platforms that can implement Delphi processes, it is still necessary to select which ones meet the remaining requirements presented in the rejection criteria, in order to make their evaluation possible.

As already mentioned, after conducting the research, was concluded that many of the platforms are declining projects that have been left abandoned or even canceled. As such, one of the most challenging tasks was to contact the people who are in charge of these projects and subsequently request the use of their platforms.

Out of a total of 13 platforms, only six were eligible for evaluation:

- Delphi Manager
- eDelphi
- HalnyX*
- Mesydel

- Surveylet by Calibrum
- Welphi

*****HalnyX is a platform that was exclusively and uniquely created to implement Real-Time Delphi processes. Some of the features present in the Framework are only indicated for platforms that implement multi-round Delphi processes and therefore they cannot be evaluated in a platform that only implements Real-Time Delphi processes. However, HalnyX platform will be included in the Framework and in those specific features there will be a note indicating the impossibility of evaluation represented with N/A.

Below is presented a brief overview of each WDP that were eligible for evaluation.

Table 12 - Delphi Manager platform overview

Delphi Manager	
Released date	2012
Languages	English
Country of origin	England
Developed by	COMET
Price for user	€200/process
Free trial	No
Site URL	https://www.comet-initiative.org/delphimanager/

Of the WDP studied in the present dissertation, *Delphi Manager* may be the most distinctive. The questionnaires are created from an Excel template that the moderator fills out and gives to the COMET team. The questionnaire is returned and by using the Setup Wizard the moderator is able to change the questionnaire, send emails to participants, aggregate the data for each round, and download the data in CSV format. Unfortunately, the platform does not have a free trial, and access to it was not guaranteed either. All data was collected from email conversations with COMET members, through the information available on the website, and through the platform's brochures.

Table 13 – eDelphi platform overview

eDelphi	
Released date	1998
Languages	Finnish, English
Country of origin	Finland
Developed by	Hannu Linturi, Osmo Kuusi, Jari Kaivo-oja, open source community

Price	Basic €0; Plus €90/3months, €150/6months, €240/12months; Premium €180/3months, €300/6months, €480/12months; Premium ORG €450/3months, €750/6months and €1200/12months
Free trial	Free access to eDelphi Basic Version
Site URL	http://www.edelphi.org/

eDelphi is a Platform under constant development and was specifically designed for qualitative use of the Delphi Method. The interface of the platform is included in the site itself, all the user has to do is register, secure one of the available licenses, and the user can start building the questionnaire. Although the platform has a free trial for the eDelphi Basic version, I was granted the more advanced eDelphi Premium ORG license for free by one of the eDelphi team members. All results obtained are relative to the features that are present in this license.

Table 14 – HalnyX platform overview

HalnyX	
Released date	2015
Languages	Polish, English
Country of origin	Poland
Developed by	Polish Society for Futures Studies, and later improved by 4CF Strategic Foresight
Price	Not specified
Free trial	No
Site URL	https://4cf.pl/en/

HalnyX is a platform that only implements Real-Time Delphi processes. One of the main characteristics of this platform is that the implementation of a Delphi study is not independent, that is, the 4CF team follows the Delphi processes of the platform as consultants. It is a very distinctive platform because questionnaires can only be done on a single page. HalnyX also did not have a free trial, but a demonstration of how the platform works was made through a video call with one of the 4CF members. All the information gathered about the functioning of this platform was through this demonstration, through email conversations, and through the platform's brochures.

Table 15 - SurveyLet by Calibrum platform overview

Surveylet by Calibrum	
Released date	2000
Languages	All languages
Country of origin	United States of America
Developed by	Calibrum (originally developed by Jozsef Nagy, founder of Calibrum)
Price	Multi-level pricing starting from $12/response to $2/response depending on the number of responses you purchase; a "response" is a panelist's response to the entire survey
Free trial	No
Site URL	https://calibrum.com/

Surveylet by Calibrum is the best known of the evaluated platforms. It is a platform that despite not offering a free trial, offers a demo. Free access to the platform was not guaranteed, but the respective demo was made allowing to draw all the conclusions and respective doubts. This platform had also already been evaluated in a similar study by Aengenheyster (2017).

Table 16 - Mesydel platform overview

Mesydel	
Released date	2010
Languages	French, English
Country of origin	France
Developed by	University of Liege
Price	3 to 30€ per process
Free trial	No
Site URL	https://mesydel.com/

Mesydel is also a platform that I could not secure access to and that does not have a free trial. However, I was able to get a demonstration of it via video call and they also sent me documents related to the functioning of the platform. All the information gathered about the functioning of this platform was also through this demonstration, through email conversations, and through the platform's brochures.

Table 17 - Welphi platform overview

Welphi	
Released date	2018
Languages	Portuguese, English
Country of origin	Portugal
Developed by	Decision Eyes (Company)
Price	Welphi Basic €40/month; Welphi Plus €100/month; Welphi Premium €500/month
Free trial	Free 14 day trial
Site URL	https://www.welphi.com/

Finally, the Welphi Platform was one of the platforms that I had free access to. Through the 14 days of the free trial, it was possible to draw the necessary conclusions to its functioning. Some doubts were also clarified via e-mail concerning some functionalities that had not been possible to evaluate in the free trial. It was also used information that is available on the website.

The remaining seven platforms were excluded for the following reasons:

Delphi Decision Aid (DDA) by Aegir St. Louis Integration: DDA was available during the search protocol search, but it was later deactivated and can no longer be accessed. No support e-mail is available on the platform page, and it was also not possible to find another way to contact the platform's technical staff in a subsequent search.

DEMOS – Delphi Meditation Online System: DEMOS is not currently available for use, and no responses were received after sending multiple emails to the support email addresses that were available on the platform's website.

ExpertLens: This Platform is active and available at the moment, but it was necessary to pay a high cost to access it. It was not also possible to evaluate this platform due to insufficient information available from it and because the request for free access was denied.

Global Future Intelligence System (GFIS): GFIS gave evidence that it was active at the moment, but no responses were received from the email addresses available on the platform's website. The license to access this platform can only be granted by e-mail.

Live Delphi (LD): Although this dissertation focuses only on WDP created for use on computers, it was also found a mobile application to implement Delphi processes. This application was exclusively designed for smartphone use and cannot be used on computers. It is available for download, but it is currently disabled.

UpBoard: Contrarily to the other platforms, UpBoard is a platform that provides several templates to build not only questionnaires but also other analysis tools (such as a SWOT analysis matrix). Although on the official site there is a section that emphasizes the Delphi method, it was

concluded after a meeting with a member from UpBoard that the platform is not optimized nor automated to implement Delphi processes.

Tech Cast (TC): The software is under construction, so it is not currently available for use, but it is on record for future work.

Therefore, only six platforms that were found will be evaluated.

5. Results

This section presents all the results obtained after applying the WDP Features Framework that was built and applied in the previous section.

5.1. WDP Features Framework Application Results

Once we have the WDP Features Framework built and the WDPs selected, it was possible to proceed with its application. The following Table 18 shows the results obtained in the WDP Features Framework of each WDP:

Table 18 - WDP Features Framework application in each selected WDP

Features	Welphi	eDelphi	Mesydel	Calibrum	D.M.	HalnyX
Anonymity						
Anonymity moderator	Yes	Yes	Yes	Yes	Yes	Yes
Anonymity participants	Yes	Yes	Yes	Yes	Yes	Yes
Iteration						
Process monitoring*	Yes	Yes	Yes	Yes	Yes	Yes
Round replication	Yes	Yes	Yes	Yes	Yes	N/A
Unimited rounds	Yes	Yes	Yes	Yes	No	N/A
E-mail tracking	No	No	Yes	No	Yes	No(4)
Built-in E-mail system*	Yes	Yes	Yes	Yes	Yes	No(4)
Personalized e-mails*	Yes	No	Yes	Yes	Yes	No(4)
Reminders*	Yes	No	Yes	Yes	No	No(4)
Link	Yes	Yes	Yes	Yes	Yes	Yes
QR code	No	Yes	Yes	No	No	No
Apps Share	No	No	No	No	No	Yes
Controlled feedback						
Quantitative data	Yes	Yes	Yes	Yes	Yes	Yes
Qualitative data	Yes	Yes	Yes	Yes	Yes	Yes
Qualitative data editing	Yes	Yes	Yes	Yes	Yes	Yes
Visual presentation*	Yes	Yes	Yes	Yes	Yes	Yes
Real-time Feedback	No	Yes	No	Yes	No	Yes
Feedback along with the questions*	Yes	Yes	No	Yes	Yes	Yes
Date Filter	No	No	Yes	No	No	No

Individual respondents	Yes	No	Yes	No	Yes	Yes
Grouping users	Yes	Yes	Yes	Yes	Yes	Yes
Comment boxes*	Yes	Yes	Yes	Yes	Yes	Yes
Previous answers*	Yes	Yes	No	Yes	Yes	Yes
Statistical Aggregation						
Data generation*	Yes	Yes	Yes	Yes	Yes	Yes
Data analysis tool	No	No	Yes	Yes	No	No
Data exportation*	Yes	Yes	Yes	Yes	Yes	No
Implementation features						
Predefined themes*	No	No	No	No	No	No
Page replication	No	Yes	Yes	Yes	No	No
Question replication	No	Yes	No	Yes	No	Yes
Auto-saving construction*	Yes	Yes	Yes	Yes	No	No
Same page*	No	No	Yes	Yes	No	Yes
Page break*	Yes	Yes	Yes	Yes	Yes	No
Welcome/End pages*	Yes	Yes	Yes	Yes	Yes	Yes
Multiple languages*	No	No	Yes	Yes	Yes	No
Questionnaire customization*	No	Yes	Yes	Yes	No	Yes
Attach pictures*	No	Yes	Yes	Yes	Yes	Yes
Type of questions	Yes	Yes	Yes	Yes	No (6)	Yes
Input validation	No	No	No	Yes	Yes	Yes
Set required fields	Yes (1)	No	Yes	Yes	Yes	Yes
Set other choices	Yes	No	Yes	Yes	No	Yes
Learning*	Yes	Yes	Yes	Yes	Yes	Yes
Questionnaire preview	Yes	No	Yes	Yes	Yes	Yes
Pretest*	No	No	Yes	Yes	Yes	No
Import a list of participants	Yes	Yes	Yes	Yes	No	No
Participant user-friendliness						
Registration*	Yes	Yes	Yes	Yes	Yes	Yes
Registration/log in easiness*	Yes	Yes	Yes	Yes	Yes	Yes
Instructions along with the questions*	Yes	Yes	Yes	Yes	Yes	Yes
Progress bar/indicator*	Yes	Yes	Yes	Yes	Yes	Yes
Auto-saving answers*	Yes	No (1)	Yes	No (2)	No (1)	No (1)
Device adaptation*	No	No	Yes	Yes	No	Yes
Live Chat*	No	No	No	No	No	No
Chat bot	Yes	No	No	No	No	No
GDPR	Yes	Yes	Yes	Yes	Yes	No (5)
Submission warnings*	No	No	Yes	No (3)	Yes	Yes

As we can see in the WDP Features Framework some fields represent exceptions in each WDP. To see what each field implies, the following Table 19 was created to help the reader understand the results:

Table 19 - WDP Features Framework reading support

Yes	The feature in question is present in the WDP of the respective column.
No	The feature in question is not present in the WDP of the respective column.
(Feature)*	Features that were valued by the interviewees according to the interviews' conclusions

N/A	This feature is not possible to evaluate on this WDP, because this is a platform that only implements one round Real-Time Delphi processes.
Yes(1)	Only a few question types can have this characteristic
No(1)	When you click 'next page' the platform saves the answers from the previous page but does not save the answers from the page you were responding.
No(2)	Only for radio buttons responses. The remaining ones operate according to No(1)
No(3)	No direct warning, but missed questions and current progress is available for panelists to view at any time
No(4)	Facilitator do not have the power to send e-mails through the platform manually, but the platform sends automatic e-mails to the participants
No (5)	The platform does not comply with GDPR, but the facilitator can do this by himself (using a checkbox to comply with GDPR before the survey starts for example)
No(6)	The platform only has one type of answer available to use on the survey, but it is possible to add free text buttons, dropdowns, checkboxes, radio buttons and date buttons to the registration form

By analyzing the results obtained after applying the WDP Features Framework, the following conclusions can be drawn:

- What features are present in each platform
- What percentage of features each platform fulfills in each area of concern
- Which features are least present in platforms and which features are most present
- Which platform delivers the least features in each area of concern and which platform delivers the most features

The following Table 20 translates the percentage of features that each WDP covers in each area:

Table 20 - Compliance with the features of each WDP in each area of concern

Welphi	eDelphi	Mesydel	Calibrum	D.M.	HalnyX
Anonymity					
100%	**100%**	**100%**	**100%**	**100%**	**100%**
Iteration					
70%	60%	**90%**	70%	60%	30%
Controlled Feedback					
82%	82%	73%	82%	82%	**91%**
Statistical Aggregation					
67%	67%	**100%**	**100%**	67%	33%
Implementation Features					
50%	56%	83%	**94%**	50%	61%
Participant User-Friendliness					
70%	50%	**80%**	60%	60%	60%

By analyzing the previous Table 20, we can see which WDPs cover the most features and which cover the fewest features in each area. Looking at the rows we can see bold numbers and underlined numbers. The bold numbers correspond to the highest percentage in each area of concern, that is, the highest percentage of features that a WDP fulfills in each area of concern. Contrarily, the underlined numbers correspond to the lowest percentages in each row of the Table, i.e. the WDP with the lowest percentages of fulfilled features in each area of concern.

To study the variability of the results of Table 20, the following Table 21 was constructed. This Table allows us to study the heterogeneity of the sample in order to enrich the description of the results obtained in Table 20.

Table 21 - Results Measures of Variability

	Anonymity	**Iteration**	**Controlled Feedback**	**Statistical Aggregation**	**Implementation Features**	**Participant User-Friendliness**
Min	1,00	0,30	0,73	0,30	0,50	0,50
Max	1,00	0,90	0,91	1,00	0,94	0,80
Range	0	0,60	0,18	0,70	0,44	0,30
Mean	1	0,63	0,82	0,72	0,66	0,63
Variance	0	0,03	0,003	0,05	0,03	0,009
Standard Deviation	0	0,18	0,05	0,23	0,17	0,09
Coefficient of variation	0	28%	6%	32%	26%	15%

Where:

- **Min:** Represents the minimum value of the dataset;
- **Max:** Represents the maximum value of the dataset;
- **Range:** Represents the difference between Max and Min;
- **Mean:** Represents the sum of the dataset values divided by the number of dataset values. This measure is important to understand if the development of each area is good or poor on average;
- **Variance:** Measures the dispersion of the dataset values around the mean value and is equal to the sum of squared differences between the dataset values and the dataset mean, divided by the total number of dataset values;
- **Standard Deviation (SD):** Describes the standard amount that scores deviate from the mean. A low SD indicates that the data values tend to be close to the mean while a high SD indicates that the data values are spread out over a large range. Is equal to the square root of the variance;
- **Coefficient of Variance:** It is equal to the standard deviation divided by the mean, i.e., it is the standard deviation relative to the mean. The advantages to use this formula as comparison between two different data sets is the fact that does not have a unit of measurement, being universal across datasets. In this way it is possible to analyze the heterogeneity of the observations and to better understand the results obtained.

On the other hand, to study the presence of each feature in the WDPs, two Tables were built from the results obtained in the application of the WDP Features Framework. The following Table 22 represents which features are **most** present in WDPs, in other words, which features the platforms have most in common. In this way, we can understand and register which features are most in-demand and explored on a platform that implements Web-Delphi processes.

Table 22 - Most present features in each area of concern

Area of Concern	**Features**	**Presence in percentage (%)**
Anonymity	Anonymity moderator	100%
	Anonymity participants	100%
Iteration	Process monitoring	100%
	Round Replication	83%
	Link	100%
Controlled Feedback	Quantitative data	100%
	Qualitative data	100%
	Qualitative data editing	100%
	Visual presentation	100%
	Grouping users	100%
	Comment boxes	100%
	Feedback along with the questions	83%
	Previous answers	83%
Statistical aggregation	Data generation	100%
	Data exportation	83%
Implementation Features	Page break	83%
	Welcome/End pages	100%
	Type of questions	83%
	Set require fields	83%
	Learning	100%
	Questionnaire preview	83%
Participant User-Friendliness	Registration	100%
	Registration/log in easiness	100%
	Instructions along with questions	100%
	Progress bar/indicator	100%
	GDPR	83%

Contrarily, the following Table 23 represents which of the features included in the WDP Features Framework are less present in WDPs. Thus, we can see which features were considered important based on the literature review and interviews' conclusions but that somehow are poorly explored and used in WDPs.

Table 23 - Least present features in each area of concern

Area of Concern	Features	Presence in percentage (%)
Anonymity	N/A	N/A
Iteration	E-mail tracking	33%
	QR code	33%
Controlled Feedback	Date Filter	17%
Statistical aggregation	Data analysis tool	33%
Implementation Features	Predefined themes	0%
Participant User-Friendliness	Auto-saving answers	33%
	Live chat	0%
	Chat bot	17%

The following text provides a general overview of the results obtained in each area of concern after applying the WDP Features Framework:

- **Anonymity**

Regarding this area of concern, each WDP complies 100% with the established features in this area. Since Anonymity is an indispensable condition in the implementation of the Delphi Method, it is normal that all platforms have this area well worked and developed. It is the only area where all platforms comply 100%.

In this case, any of the platforms allows guaranteeing the anonymity of a participant not only towards the other participants but also towards the moderator of the process.

Most present features in this area: *Anonymity moderator* and *Anonymity participants.*

Least present features in this area: N/A.

- **Iteration**

Regarding Iteration, all WDP meet a minimum of 60% of Features except HalnyX which meets only 30%. Since HalnyX only implements Delphi processes in Real-Time, it is normal that this area of concern is not so developed in this WDP.

In terms of features, all WDPs have the ability to track the responses of each participant, so it is possible to track the responses and possible changes in the opinion of each participant. Of the features presented, there are three that are poorly explored by the WDPs in question. Email tracking and the ability to generate QR codes to share the questionnaire are only present in two WDPs. The feature that is less explored is the sharing of the questionnaires in apps (Facebook, Twitter, LinkedIn, etc.) via a share button included in the platform. The only platform that has this last feature is HalnyX.

Most present features in this area: *Process monitoring*, *Round replication,* and *Link.*

Least present features in this area: *E-mail tracking* and *QR code*.

- **Controlled Feedback**

All platforms meet 82% of the features represented in the Controlled Feedback area, with the exception of Mesydel which meets only 73%, and HalnyX which meets 91%.

All platforms have the ability to collect quantitative and qualitative data, edit qualitative data, present statistical aggregation data in graph format on the platform itself, group participants into different Stakeholder groups, and add comment boxes next to the questions. Feedback along with the questions and the ability to have answers from previous rounds automatically entered into the subsequent round are features that are present on all platforms except Mesydel. On the other hand, the feature that is less explored in this range of WDPs is the Date Filter feature and is only present in Mesydel.

Most present features in this area: *Quantitative data, Qualitative data, Qualitative data editing, Visual presentation, Grouping users, Comments boxes, Feedback along with the questions* and *Previous answers.*

Least present feature in this area: *Date filter*.

- **Statistical Aggregation**

Calibrum and Mesydel meet this area of concern 100%, while the others meet 67% (66.67%), with the exception of HalnyX which meets only 33% (33.33%) of the features.

Regarding the features, all platforms are able to automatically generate statistical group response data and all platforms except HalnyX are able to export raw data to external statistical analysis tools. Only Mesydel and Calibrum have a built-in feature that allows advanced data analysis within the platform itself, such as measuring consensus and studying the stability of responses across rounds.

Most present features in this area: *Data generation* and *Data exportation.*

Least present feature in this area: *Data analysis tool*.

- **Implementation Features**

For this area of concern, through the analysis of the WDP Features Framework, it is concluded that the Calibrum platform is the most developed, fulfilling 94% of the features. In second place

is Mesydel which complies 83%, HalnyX 61%, eDelphi 56%, and lastly, Welphi and DelphiManager which complies only 50%.

This area of concern is the one that concentrates the largest number of features. The feature that is common to all platforms is "Learning", meaning that all platforms offer a learning and support system with different learning alternatives (e.g., video tutorials, examples, demonstrations, user guides, etc.). On the other hand, none of the platforms offer predefined themes that help you create the surveys faster, instead, you basically have to create them from scratch.

It is also important to mention that DelphiManager is a platform that requires the intervention of the back-office team when constructing the questionnaire and that the Moderator's intervention at this stage is slightly limited. Basically, the Moderator fills out a template in Excel format with the respective questions and sends it to the DelphiManager team to build the questionnaire, and only afterward the Moderator can have intervention through the Setup Wizard. For this reason, DelphiManager has little development in this area of concern since most of the features are aimed at the questionnaire construction (construction done by the user/moderator).

Most present features in this area: *Page break*, *Welcome/End pages*, *Type of questions*, *Set required fields*, *Learning*, *Questionnaire preview*.

Least present features in this area: *Predefined themes*.

- **Participant User-Friendliness**

Regarding Participant User-Friendliness, Mesydel is the platform that delivers the most features, about 80%. In contrast, the platform that delivers the least features in this area is eDelphi with only 50%.

There is a high level of similarity between the platforms in this area. All platforms meet the following features: Mandatory and simple registration of participants, the ability to enter instructions alongside the questions, and the existence of a progress bar throughout the questionnaire. Only the Welphi and Mesydel Platforms comply 100% with automatic auto-saving of participants' answers, while the others have exceptions regarding this feature. In the case of eDelphi, DelphiManager, and HalnyX, the auto-saving is only done if the participant clicks on "Next page" to advance to the next page (only the answers on the previous pages are saved) and in the case of Calibrum it also only saves the answers if the participant clicks on "Next page" with the exception of the radio buttons which once answered are automatically saved even if the participant closes the survey tab. Another important point to draw out is that none of the platforms provides a Live Chat (a feature that was valued in the interviews) and that only the Welphi platform provides a Chatbot in order to help the platform users with doubts and frequently asked questions. As for compliance with the EU General Data Protection Regulation (GDPR), all platforms are compliant except HalnyX, which in turn advises that, if necessary, the moderator can do this

manually and independently of the platform by inserting a checkbox regarding GDPR compliance before the survey starts. Regarding submission warnings, only Mesydel, DelphiManager and HalnyX comply with this feature. Although Calibrum does not comply with this feature, each panelist can check the missed questions and their current progress (but without receiving a direct warning from the platform).

Most present features in this area: *Registration*, *Registration/log-in easiness*, *Instructions along with questions*, *Progress bar/indicator*, *GDPR*.

Least present features in this area: *Auto-saving answers*, *Live chat,* and *Chatbot*.

6. Discussion

In the present section, a discussion is held regarding the methodology used to achieve the objectives outlined in the present dissertation. After, a discussion of the results is also carried out in order to draw conclusions related to the objectives of the present dissertation.

6.1. On the methodology

By reviewing and analyzing the results obtained, it can be stated that the methodology applied in the present dissertation met the objectives previously proposed in section 1.2. However, it is also necessary to discuss some limitations that emerged during the course of the work and how they were mitigated in order to not jeopardize the objectives previously outlined.

The literature review was very important for the development of this dissertation. The review was mainly focused on the Delphi Method, thus allowing us to collect information and study its state of the art. Not only did it allow us to understand its characteristics, advantages, and disadvantages, but also, its transition from the traditional paper-based method to a digital method. Unfortunately, the implementation of the Delphi Method in Web-Delphi platforms is a subject that is poorly studied and explored nowadays, but on the other hand, it is widely used in the implementation of current Delphi processes given its advantages. One of the limitations in conducting the literature review was dealing with the lack of articles on this subject, and therefore, in order to consolidate the knowledge obtained in the literature review, semi-structured interviews were also conducted (with people who have worked with WDPs) to gather new insights and additional perspectives. The semi-structured interview format was adopted in these interviews because it facilitated communication between the interviewees regarding issues that were not foreseen and that arose during each interview. In this way, it was possible to build a set of areas of concern regarding the implementation of online Delphi processes, each composed of a group of features, which originated the WDP Features Framework. Although the Framework coincides with the insights from the literature review and the conclusions obtained in the interviews, we must

not forget that only four interviews were conducted and some possibly additional insights could have emerged if there had been a larger number of interviews. Nevertheless, the WDP Features Framework proved to be a solid framework that addresses the most important areas and characteristics of the Delphi Method and its online implementation.

In order to find the currently available WDPs, the search protocol developed in this dissertation also proved to be effective. However, it is difficult to know if all existing WDPs were found due to the large dimension of the internet and even with an exhaustive search it is possible that there are platforms that were not reached by applying this search protocol.

After applying the search protocol the platforms were analyzed using the WDP Features Framework. One of the major limitations faced throughout the work was securing free access licenses to the WDPs that did not have free access. As such, it was not possible to analyze all the WDPs available, which compromised part of the defined objectives. Some of the platforms could only be analyzed using user guides, Demos, and video calls with the technical staff of the respective platforms, which proved to be a limitation in their analysis.

Although the framework's application allows for a solid analysis of each platform's performance, it was also proposed to build an MCDA according to the MACBETH approach that allows for the evaluation of each platform's performance in quantitative values. The advantage of applying an MCDA to these platforms is that the scores given to each platform make it easier to compare them. As already stated, the particularity of this model is that it is proposed based on the WDP Features Framework, thus facilitating its construction. It is suggested that the areas of concern used in the Framework be used directly as areas of concern for the model, and that the respective features be used as the basis for creating the criteria that concern the model. One of the problems faced in the present work is that due to the space limits imposed in the writing of the dissertation, this model was neither built nor applied, but nevertheless, it was proposed for future work in order to not only reinforce the objectives outlined in this dissertation but also to help potential decision-makers in choosing a Web-Delphi Platform for the implementation of online Delphi processes. Given the situation, this dissertation focuses more on the technical component of this process and not on the social component.

6.2. On the results

The discussion of the results is based on the results obtained in section 5. By analyzing Table 21 it is possible to discuss the development of the different areas of concern in the WDPs. The adopted form was to discuss the variability of the results obtained in Table 20 that represent the compliance with the features of each WDP in each area of concern.

Starting with **Anonymity**, the average obtained is 1. Consequently, the range and the coefficient of variation are zero due to the fact that all platforms meet all the features mentioned in this area. This result was expected because this is an area that presents only two features, one of which (the anonymity among participants) is mandatory in the practice of the Delphi method. It is then the most developed area, but nevertheless, it is the area that represents the least number of features.

As for **Iteration**, the amplitude of this area is the second highest (0.60) in Table 21, meaning that there is a large difference in the number of features between the platform that fulfils more features and the platform that fulfils fewer features in this area. The average is also the lowest in the Table (0.63) and the coefficient of variation is the second highest in the Table (28%), meaning that it is the second area where the platforms differ most in terms of feature fulfillment and should be developed. It is recommended that the HalnyX platform further develop and invest in the features in this area since it comprises only 30% of the features in this area.

Contrarily, **Controlled Feedback** has the second highest average (0.82). The coefficient of variation is only 6%, which translates into a small variability of results in this area of concern, which translates an area that is well developed and present in the platforms. It should also be noted that the minimum value of this area is 0.73 (73%) and represents the second highest value of the Min values represented in the Table.

Regarding **Statistical Aggregation**, the range is the highest in Table 21. The average is high (0.72), being the third highest in the table, but nevertheless, we are talking about an area of concern that groups only three features. The coefficient of variation is the highest (32%), being thus an area that should be further developed and present across some platforms, namely on HalnyX that only meets 33% of the features related to this area. Once again, it is recommended that the HalnyX platform further develop and invest in the features in this area.

Analyzing the results concerning the **Implementation Features** area, we conclude that it is an area with a coefficient of variance (26%), very close to the coefficient of variance of the Iteration area (28%), but with a smaller amplitude and a higher average (0.66). Thus, it is an area a little more explored than Iteration but that also requires greater development and investment of features by some WDPs.

Finally, **Participant User-Friendliness** is an area that just like the iteration presents the lowest average (0.63) but on the other hand has a much lower coefficient of variation (15%), which means that although the features of this area are on average less present in the platforms, the platforms are more homogeneous in terms of features presenting similar degrees of development. Moving on from the previous analysis we now move on to an analysis of the respective features that are present in the data in Table 23. Although the features represented in Table 23 are poorly

explored, given the sample of WDPs used in this dissertation, these features bring many benefits in Web-Delphi practice. The following points translate some of the benefits that these features can bring if used:

E-mail Tracking: One of the main problems in using these WDPs, is that although they have a built-in e-mail system that allows facilitating the communication between the moderator of the process and the participants, this built-in e-mail system can also bring some obstacles to the Delphi process. According to the conclusions drawn in the interviews, one of the problems concerning the use of this system is that the emails sent can be moved to the participant's spam box. The Spam Box is an e-mail inbox that does not use notifications and consequently does not warn the receiver regarding the arrival of an e-mail. Thus, the moderator may be left in doubt and may think that the person has simply ignored the request to participate. As we know, the dropout rate of participants in Delphi processes is one of the big problems associated with the Delphi method and so there have to be ways to mitigate this. Using email tracking allows the moderator to see if the emails have actually been received and read by the participants and to better manage participants and their emails.

QR code: Nowadays people are increasingly using their cell phones to answer online surveys because of their ease of access and convenience. One of the technologies that allow the use of the cell phone camera to access websites is called a QR code. Besides that it is a good practice to have alternative ways to access the survey (in case one is not working, for example), the QR code can be very beneficial in other ways. Let's imagine that a Web-Delphi study was being prepared and that the study consisted of using people who frequented the facilities of the Instituto Superior Técnico as participants. One way to get people to participate would be distributing flyers around the buildings with an explanation of the study and a QR code in case people were interested in taking part and contributing to the study.

Date filter: This feature can be a very useful tool in the analysis of results for each round. This feature allows the moderator to establish a temporal period of analysis where the results obtained are only relative to that period. Let's imagine that a Delphi study regarding a moral and delicate subject like euthanasia is in progress, but that in the course of one Delphi round of the process a study that represents a new insight regarding this subject is published. Using this feature, the moderator is able to study the impact that this study had on the results and changes in the participants' opinions during that round.

Data analysis tool: Despite the fact that it is barely present in the mentioned WDPs, the Data analysis tool is probably one of the most impactful features in Web-Delphi practice. As we know, one of the main tasks that a moderator usually has in an online Delphi process is to export and process data in order to assess whether there is stability of responses and also to assess whether consensus has been reached among the group of participants. As platforms have evolved, they

have become increasingly autonomous and optimized for the implementation of Delphi processes. Mesydel and Calibrum are examples of WDPs that have this feature called "Data analysis tool". This feature allows automatic analysis of the data obtained in each round and calculates indicators such as "Consensus measurement" and "Group stability". The presence of this feature allows the process to become more automated, efficient, and faster, also reducing the moderator's workload. It also helps moderators who are not so familiar with the Delphi Method to understand the results obtained.

Predefined themes: This feature is not present in any of the WDPs but was highly valued according to the conclusions obtained in the interviews. Predefined themes can be a huge advantage when building a Web-Delphi process in a WDP. The presence of this feature not only allows reducing the time associated with the construction of the questionnaire but also helps the moderator to think about how he/she will implement his/her process in the questionnaire format. For example, a moderator who is unsure about which question type (dropdown, group button, radio button, etc.) to use to construct a particular question, the use of predefined themes can be a way to allow the moderator to visualize how the question would look in the survey without having to go through the effort of creating it first.

Auto-saving answers: This feature consists of the ability that a WDP has to automatically save the participants' answers. Imagine that a participant was answering a very long and exhausting questionnaire and for some reason, the electronic device he was using turned off and the answers had not been saved. This event could cause annoyance to the participant and even cause him to give up due to the time that the questionnaire is consuming. The way to overcome these events is to have a feature that allows you to save the participants' answers automatically without requiring any action from the participant. As shown in the results of the WDP Features Framework only two WDPs, Welphi and Mesydel, had this feature fully present. However, the remaining platforms all feature semi-automatic saving, meaning that when participants click to proceed to the next page of the quiz, the answers from the previous pages are all saved, but in the event that a participant answers a page of the quiz and accidentally closes that page, none of the answers from that page are saved. In addition to this semi-automatic saving, Calibrum has a special feature that consists of auto-saving only the answers given in radio buttons, while the other question types have only a semi-automatic saving.

Live chat and Chatbot: Due to their high similarity, these two features will be commented on together. Both valued during the interviews, Live chat, and Chatbot are features that can benefit not only the participant but also the moderator. The existence of a Chatbot or Live chat that allows the moderator to communicate with the technical staff of the platform is very beneficial because it can be a way to clarify possible doubts of the moderator regarding the operation of the platform in a faster way. On the other hand, a Live chat can also be beneficial to the participant in case it is a Live chat between the participant and the moderator that allows the participant to clarify

possible doubts with the moderator instead of having to go through the effort of sending an email. Therefore, Live chat and Chatbot are great tools to facilitate communication involved in online Delphi processes.

7. Conclusions and Future Work

The present dissertation performed a comparative analysis between the different WDPs that are currently available online. The comparative analysis was performed based on the application of the WDP Features Framework, which was also built in this dissertation. The objective of this dissertation was to perform a comparative analysis between the different available WDPs and explore how their features can advance the use of the Delphi method . Additionally, a multi-criteria model that adopts the MACBETH approach was also proposed in order to consolidate the comparative analysis between the platforms and help potential decision-makers choose a WDP to implement Delphi processes online. The main particularity of this model is that the WDP Features Framework will be the basis that will help build the proposed model, particularly in the definition of the areas of concern and the evaluation criteria, and also in the construction of the descriptors of performance. In this case, the model was only proposed, and its construction and application were not executed in the present dissertation.

After applying the Framework and analyzing the respective results obtained, it was possible to achieve the objectives previously outlined for this dissertation. It was concluded that although all these WDPs were designed for the same purpose (application of the Delphi Method), these platforms present some diversities among themselves. Although most of the features that were considered valuable in the implementation of online Delphi processes were present in most of the platforms, a portion of these features was not present in them. Therefore, the present work also explained the benefits and contribution that each of these features could have in implementing Delphi online processes even though they are poorly explored on these platforms. It was also possible to analyze the development of each concern area outlined in the Framework, and understand which ones need more development, which ones have a more homogeneous presence in the platforms, and which ones have a more heterogeneous presence, thus understanding which areas are more bet on today by the platforms and which ones are less bet on.

As future work, it is recommended to reapply the search protocol in order to find possible emerging WDPs that were not included in this dissertation and consequently apply the WDP Features Framework to these platforms. Since part of the found WDPs were not selected for analysis due to cost constraints, it would also be beneficial to apply the WDP Features Framework to these platforms in future work. Since the costs of using each WDP were not considered in this dissertation, it is also recommended to perform a cost-benefit analysis on them. Since it was also proposed a multi-criteria model adopting the MACBETH approach to help a possible decision-

maker choose a WDP for online Delphi study implementations, it is also recommended the future construction of this model and consequent application.

In conclusion, it is intended that this work be considered a useful tool in the analysis of Web-Delphi Platforms as well as a contribution to their development.

8. References

Abadie, M., & Waroquier, L. (2019). *Evaluating the Benefits of Conscious and Unconscious Thought in Complex Decision Making. Policy Insights from the Behavioral and Brain Sciences, 6(1), 72–78.*

Addison, P. F., Rumpff, L., Bau, S. S., Carey, J. M., Chee, Y. E., Jarrad, F. C., ... & Burgman, M. A. (2013). Practical solutions for making models indispensable in conservation decision-making. *Diversity and Distributions*, *19*(5-6), 490-502.

Aengenheyster, S., Cuhls, K., Gerhold, L., Heiskanen-Schüttler, M., Huck, J., & Muszynska, M. (2017). *Real-Time Delphi in practice — A comparative analysis of existing software-based tools. Technological Forecasting and Social Change, 118, 15–27.*

Allen, P. J., & Roberts, L. D. (2010). The Ethics of Outsourcing Online Survey Research. International Journal of Technoethics, 1(3), 35–48.

Andrews, C. G., & Allen, J. M. (2002). Utilization of technology-enhanced Delphi techniques. In *Workforce Education Forum* (Vol. 29, No. 1, pp. 1-15). Pennsylvania State University.

Antunes, P., Santos, R., & Videira, N. (2006). Participatory decision making for sustainable development—the use of mediated modeling techniques. *Land use policy*, *23*(1), 44-52.

Aoki, K.and Elsmar, M. (2000) Opportunities and Challenges of Conducting Web-surveys – Results From a Field Experiment. Proceedings of the Survey Research Methods Section, American Statistical Association, Alexandria, Virginia, AMSTAT. pp. 930–935.

Archer, T. M. (2003). Web-based surveys. Journal of Extension, 41(4), 1-5.

Avella, J. R. (2016). Delphi panels: Research design, procedures, advantages, and challenges. *International Journal of Doctoral Studies*, *11*(1), 305-321.

Bachmann, D., Elfrink, J., & Vazzana, G. (1996). Tracking the progress of e-mail vs. snail-mail. *Marketing research*, *8*(2), 30.

Bana e Costa, C. A., & Beinat, E. (2005). *Model-structuring in public decision-aiding* (No. LSEOR 05.79). Operational Research Group, Department of Management, London School of Economics and Political Science.

Bana e Costa, C. A., & Chagas, M. P. (2004). *A career choice problem: An example of how to use MACBETH to build a quantitative value model based on qualitative value judgments. European Journal of Operational Research, 153(2), 323–331.*

Bana e Costa, C. A., & Oliveira, M. D. (2012). *A multicriteria decision analysis model for faculty evaluation. Omega, 40(4), 424–436.*

BANA E COSTA, C. A., & VANSNICK, J. C. (1997). Applications of the MACBETH approach in the framework of an additive aggregation model. *Journal of Multi-Criteria Decision Analysis*, *6*(2), 107-114.

Bana e Costa, C. A., Carnero, M. C., & Oliveira, M. D. (2012). *A multi-criteria model for auditing a Predictive Maintenance Programme. European Journal of Operational Research, 217(2), 381–393.*

Bana e Costa, C. A., De Corte, J. M. & Vansnick, J.C. (2003). MACBETH. (Overview of MACBETH multicriteria decision analysis approach). International Journal of Information Technology and Decision Making. 11. 359-387.

Bana, C. A., De Corte, J. M., & Vansnick, J. C. (2005). On the mathematical foundation of MACBETH. In *Multiple criteria decision analysis: state of the art surveys* (pp. 409-437). Springer, New York, NY.

Bana e Costa, C. A., de Corte, J.-M., & Vansnick, J.-C. (2011). *MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique). Wiley Encyclopedia of Operations Research and Management Science.*

Bana e Costa, C. A., Lourenço, J. C., Chagas, M. P., & Bana e Costa, J. C. (2008). Development of reusable bid evaluation models for the Portuguese Electric Transmission Company. *Decision Analysis*, *5*(1), 22-42.

Bana e Costa, Carlos & De Corte, Jean-Marie & Vansnick, Jean-Claude. (2003). MACBETH. (Overview of MACBETH multicriteria decision analysis approach). International Journal of Information Technology and Decision Making. 11. 359-387.

Bana, C. A., Meza, L. A., & Oliveira, M. D. (2013). O método MACBETH e aplicação no Brasil. *Engevista*, *15*(1), 3-27.

Belton, V., & Stewart, T. (2002). *Multiple criteria decision analysis: an integrated approach*. Springer Science & Business Media

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., & Alberti, C. (2011). *Using and Reporting the Delphi Method for Selecting Healthcare Quality Indicators: A Systematic Review. PLoS ONE, 6(6), e20476.*

Bourque, P., Dupuis, R., Abran, A., Moore, J. W., Tripp, L., & Wolff, S. (2002). Fundamental principles of software engineering–a journey. *Journal of Systems and Software*, *62*(1), 59-70.

Bouyssou, D. (1990). Building criteria: a prerequisite for MCDA. In *Readings in multiple criteria decision aid* (pp. 58-80). Springer, Berlin, Heidelberg.

Buede, D. M. (1986). Structuring value attributes. Interfaces, 16:52-62

Burgman, M. A., McBride, M., Ashton, R., Speirs-Bridge, A., Flander, L., Wintle, B., ... & Twardy, C. (2011). Expert status and performance. *PLoS One*, *6*(7), e22998.

Callegaro, M., Manfreda, K., & Vehovar, V. (2015). *Web-survey Methodology*. Sage.

Cantrell, M. A., & Lupinacci, P. (2007). *Methodological issues in online data collection. Journal of Advanced Nursing, 60(5), 544–549.*

Chen, K., Ren, Z., Mu, S., Sun, T. Q., & Mu, R. (2020). Integrating the Delphi survey into scenario planning for China's renewable energy development strategy towards 2030. *Technological Forecasting and Social Change*, *158*, 120157.

Chou, C. (2002). Developing the e-Delphi system: A web-based forecasting tool for educational research. *British Journal of Educational Technology*, *33*(2), 233-236.

Clark, C. R., Heimlich, J. E., Ardoin, N. M., & Braus, J. (2020). Using a Delphi study to clarify the landscape and core outcomes in environmental education. *Environmental Education Research*, *26*(3), 381-399.

Clivillé, V., Berrah, L., & Mauris, G. (2007). *Quantitative expression and aggregation of performance measurements based on the MACBETH multi-criteria method. International Journal of Production Economics, 105(1), 171–189.*

Costa, C. A. B. E., Martins, P. A., Oliveira, M. D., Sernadas, A., & Soares, C. A. M. (2010, November). Faculty evaluation using multicriteria value measurement. In *Proceedings of the 12th WSEAS international conference on Mathematical and computational methods in science and engineering* (pp. 287-290).

Costa, C. A. B., Lopes, D. F., & Oliveira, M. D. (2014). Improving risk matrices using the MACBETH approach for multicriteria value measurement. *Parte: Colecao: http://hdl. handle. net/10316.2/33309*.

Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2019). *Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates. Journal of Survey Statistics and Methodology.*

Dajani, J. S., Sincoff, M. Z., & Talley, W. K. (1979). *Stability and agreement criteria for the termination of Delphi studies. Technological Forecasting and Social Change, 13(1), 83–90.*

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management science*, *9*(3), 458-467.

Day, J., & Bobeva, M. (2005). A generic toolkit for the successful management of Delphi studies. *The Electronic Journal of Business Research Methodology*, *3*(2), 103-116.

Dayé, C. (2018). How to train your oracle: The Delphi method and its turbulent youth in operations research and the policy sciences. *Social studies of science*, *48*(6), 846-868.

De Leeuw, E. D., Hox, J. J., & Snijkers, G. (1995). The effect of computer-assisted interviewing on data quality. A review. *Market Research Society. Journal.*, *37*(4), 1-19.

De Loe, R. C. (1995). *Exploring complex policy questions using the policy Delphi. Applied Geography, 15(1), 53–68.*

De Montis, A., De Toro, P., Droste-Franke, B., Omann, I., & Stagl, S. (2000, May). Criteria for quality assessment of MCDA methods. In *3rd Biennial Conference of the European Society for Ecological Economics, Vienna* (pp. 3-6).

Devaney, L., & Henchion, M. (2018). *Who is a Delphi "expert"? Reflections on a bioeconomy expert selection procedure from Ireland. Futures, 99, 45–55.*

Dhouib, D. (2014). *An extension of MACBETH method for a fuzzy environment to analyze alternatives in reverse logistics for automobile tire wastes. Omega, 42(1), 25–32.*

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of clinical epidemiology*, *67*(4), 401-409.

Dillman, D. A., Tortora, R. D., & Bowker, D. (1998, August). Principles for constructing web surveys. In *Joint Meetings of the American Statistical Association* (Vol. 64, pp. 1-16).

Dodgson, J. S., Spackman, M., Pearman, A., & Phillips, L. D. (2009). Multi-criteria analysis: a manual.

Donohoe, H. M., & Needham, R. D. (2009). Moving best practice forward: Delphi characteristics, advantages, potential problems, and solutions. *International Journal of Tourism Research*, *11*(5), 415-437.

Donohoe, H., Stellefson, M., & Tennant, B. (2012). Advantages and limitations of the e-Delphi technique: Implications for health education researchers. *American Journal of Health Education*, *43*(1), 38-46..

Duffield, C. (1993). *The Delphi technique: a comparison of results obtained using two expert panels. International Journal of Nursing Studies, 30(3), 227–237.*

Duffy, M. E. (2002). *Methodological Issues In Web-based Research. Journal of Nursing Scholarship, 34(1), 83–88.*

Durugbo, C. M., Al-Balushi, Z., Anouze, A., & Amoudi, O. (2020). Critical indices and model of uncertainty perception for regional supply chains: insights from a Delphi-based study. *Supply Chain Management: An International Journal*.

e Costa, C. A. B., De Corte, J. M., & Vansnick, J. C. (2016). On the mathematical foundations of MACBETH. In *Multiple criteria decision analysis* (pp. 421-463). Springer, New York, NY.

e Costa, C. A. B., Vieira, A. C., Nóbrega, M., Quintino, A., Oliveira, M. D., & e Costa, J. B. (2019). Collaborative Value Modelling in corporate contexts with MACBETH. *Procedia Computer Science*, *162*, 786-794.

East, L., Jackson, D., O'Brien, L., & Peters, K. (2008). *The benefits of computer-mediated communication in nursing research. Contemporary Nurse, 30(1), 83–88.*

Flostrand, A., Pitt, L., & Bridson, S. (2020). The Delphi technique in forecasting–A 42-year bibliographic analysis (1975–2017). *Technological Forecasting and Social Change*, *150*, 119773.

Frazão, T. D. C., Camilo, D. G. G., Cabral, E. L. S., & Souza, R. P. (2018). *Multicriteria decision analysis (MCDA) in health care: a systematic review of the main characteristics and methodological steps. BMC Medical Informatics and Decision Making, 18(1).*

Fricker, R. D., & Schonlau, M. (2002). Advantages and disadvantages of Internet research surveys: Evidence from the literature. *Field methods*, *14*(4), 347-367.

Gerrish, K., & Lacey, A. (2010). *The research process in nursing*. John Wiley & Sons.

Gnatzy, T., Warth, J., von der Gracht, H., & Darkow, I.-L. (2011). *Validating an innovative real-time Delphi approach - A methodological comparison between real-time and conventional Delphi studies. Technological Forecasting and Social Change, 78(9), 1681–1694.*

Gordon, T. J., & Helmer, O. (1964). *Report on a long-range forecasting study*. RAND CORP SANTA MONICA CA SANTA MONICA United States.

Gordon, T., & Pease, A. (2006). RT Delphi: An efficient, "round-less" almost real time Delphi method. Technological Forecasting and Social Change, 73(4), 321–333.

Graefe, A., & Armstrong, J. S. (2011). Comparing face-to-face meetings, nominal groups, Delphi and prediction markets on an estimation task. *International Journal of Forecasting*, *27*(1), 183-195.

Greenlaw, C., & Brown-Welty, S. (2009). A comparison of web-based and paper-based survey methods: testing assumptions of survey mode and response cost. *Evaluation review*, *33*(5), 464-480.

Grime, M. M., & Wright, G. (2014). D elphi Method. *Wiley statsref: Statistics reference online*, 1-6.

Hanafin, S., Brooks, A. M., Carroll, E., Fitzgerald, E., GaBhainn, S. N., & Sixsmith, J. (2007). Achieving consensus in developing a national set of child well-being indicators. Social Indicators Research, 80(1), 79-104.

Hardigan, P. C., Succar, C. T., & Fleisher, J. M. (2011). *An Analysis of Response Rate and Economic Costs Between Mail and Web-Based Surveys Among Practicing Dentists: A Randomized Trial. Journal of Community Health, 37(2), 383–394.*

Hart, L. M., Ferreira, K. B., Ambwani, S., Gibson, E. B., & Austin, S. B. (2020). Developing expert consensus on how to address weight stigma in public health research and practice: A Delphi study. *Stigma and Health*.

Hashim, F., Alam, G. M., & Siraj, S. (2009). Ensuring participatory based decision-making practice in Higher Education through E-management: A faculty initiative. *Recent Advances in e-Activities, Informationa Security and Privace/F. Hashim, G. Mahabubul, S. Siraj//WSEAS Press*, 280-285.

Hashim, F., Alam, G. M., & Siraj, S. (2010). Information and communication technology for participatory based decision-making-E-management for administrative efficiency in Higher Education. *International Journal of Physical Sciences*, *5*(4), 383-392.

Haynes, C. A., & Shelton, K. (2018). Delphi method in a digital age: Practical considerations for online Delphi studies. In *Handbook of Research on Innovative Techniques, Trends, and Analysis for Optimized Research Methods* (pp. 132-151). IGI Global.

Heiervang, E., & Goodman, R. (2009). Advantages and limitations of web-based surveys: evidence from a child mental health survey. Social Psychiatry and Psychiatric Epidemiology, 46(1), 69–76.

Heiko, A. V. D. G. (2012). Consensus measurement in Delphi studies: review and implications for future quality assurance. Technological forecasting and social change, 79(8), 1525-1536.

Heisler, M., Bouknight, R. R., Hayward, R. A., Smith, D. M., & Kerr, E. A. (2002). The relative importance of physician communication, participatory decision making, and patient understanding in diabetes self-management. *Journal of general internal medicine*, *17*(4), 243-252.

Helms, C., Gardner, A., & McInnes, E. (2017). *The use of advanced web-based survey design in Delphi research. Journal of Advanced Nursing, 73(12), 3168–3177.*

Hirokawa, R. Y., & Poole, M. S. (Eds.). (1996). *Communication and group decision making*. Sage Publications.

Holloway, K. (2012). *Doing the E-Delphi. CIN: Computers, Informatics, Nursing, 30(7), 347–350.*

Holstein, J. A. (2002). *Handbook of interview research: Context and method*. Sage.

Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical Assessment, Research, and Evaluation*, *12*(1), 10.

Humphrey-Murto, S., & de Wit, M. (2018). *The Delphi Method - More Research Please. Journal of Clinical Epidemiology.*

Ishizaka, A., & Nemery, P. (2013). *Multi-criteria decision analysis: methods and software*. John Wiley & Sons.

Johnson, J. P. (2020). *A Classical Delphi Study: Identifying the Essential Skills Educational Leadership Experts Perceive as Crucial for Novice Elementary School Principals to Support Students toward the Every Student Succeeds Act (ESSA) Goals* (Doctoral dissertation, Brandman University).

Jones, S., Murphy, F., Edwards, M., & James, J. (2008). Doing things differently: advantages and disadvantages of web questionnaires. *Nurse researcher*, *15*(4).

Kaner, S. (2014). *Facilitator's guide to participatory decision-making*. John Wiley & Sons.

Keeney, R. L. (1996). *Value-focused thinking*. Harvard University Press.

Keeney, R. L. and Raiffa, H., Decisions with Multiple Objectives: Preferences and Value Tradeoffs, New York: Wiley, 1976.

Keeney, S., Hasson, F., & McKenna, H. (2006). *Consulting the oracle: ten lessons from using the Delphi technique in nursing research. Journal of Advanced Nursing, 53(2), 205–212.*

Khan, S., Haleem, A., & Khan, M. I. (2020). Risk management in Halal supply chain: an integrated fuzzy Delphi and DEMATEL approach. *Journal of Modelling in Management*.

Kumar, A., Sindhwani, P., & Trivedi, S. K. (2020). Combined fuzzy theory and Delphi method for instruments designing framework of online marketing flexibility dimensions. *International Journal of Services Technology and Management*, *26*(4), 341-356.

Kwak, N., & Radler, B. (2002). A comparison between mail and web-surveys: Response pattern, respondent profile, and data quality. *Journal of official statistics*, *18*(2), 257.

Lahdelma, R., Salminen, P., & Hokkanen, J. (2000). Using multicriteria methods in environmental planning and management. *Environmental management*, *26*(6), 595-605.

Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological forecasting and social change*, *73*(5), 467-482.

Lindeman, C. A. (1975). Delphi survey of priorities in clinical nursing research. *Nursing research*, *24*(6), 434-441.

Linkov, I., Sahay, S., Kiker, G., Bridges, T., & Seager, T. P. (2005). Multi-criteria decision analysis: A framework for managing contaminated sediments. In *Strategic management of marine ecosystems* (pp. 271-297). Springer, Dordrecht.

Linstone, H. A., & Turoff, M. (2011). Delphi: A brief look backward and forward. *Technological forecasting and social change*, *78*(9), 1712-1719.

Linstone, H. A., & Turoff, M. (Eds.). (1975). *The delphi method* (pp. 3-12). Reading, MA: Addison-Wesley.

Loo, R. (2002). *The Delphi method: a powerful tool for strategic management. Policing: An International Journal of Police Strategies & Management, 25(4), 762–769.*
Martino, J. P. (1972). *Tools for looking ahead. IEEE Spectrum, 9(10), 32–40.*

Martino, J. P. (1993). *Technological forecasting for decision making*. McGraw-Hill, Inc..

McIntosh, M. J., & Morse, J. M. (2015). Situating and constructing diversity in semi-structured interviews. *Global qualitative nursing research*, *2*, 2333393615597674.

McKenna, H. P. (1994). The Delphi technique: a worthwhile research approach for nursing?. *Journal of advanced nursing*, *19*(6), 1221-1225.

Mehta, R., & Sivadas, E. (1995). Comparing response rates and response content in mail versus electronic mail surveys. *Market Research Society. Journal.*, *37*(4), 1-12.

Mitchell, V. W. (1991). *The delphi technique: an exposition and application. Technology Analysis & Strategic Management, 3(4), 333–358.*

Novakowski, N., & Wellar, B. (2008). Using the Delphi Technique in Normative Planning Research: Methodological Design Considerations. Environment and Planning A, 40(6), 1485–1500.

Ogden, J. A., Petersen, K. J., Carter, J. R., & Monczka, R. M. (2005). Supply management strategies for the future: a Delphi study. *Journal of Supply Chain Management*, *41*(3), 29-48.

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & management*, *42*(1), 15-29.

Pereira, M. T., Dias, E., & Fontes, D. B. (2019). A MCDA Model for olive oil supplier selection using MACBETH. *International journal for quality research*, *13*(4), 849-862.

Phillips, L. D., & e Costa, C. A. B. (2007). Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Annals of Operations Research*, *154*(1), 51-68.

Pucher, P. H., Peckham-Cooper, A., Fleming, C., Mohamed, W., Clements, J. M., Nally, D., ... & Mohan, H. M. (2020). Consensus recommendations on balancing educational opportunities and service provision in surgical training: Association of Surgeons in Training Delphi qualitative study. *International Journal of Surgery*.

Rauch, W. (1979). *The decision Delphi. Technological Forecasting and Social Change, 15(3), 159–169.*

Reddy, M. K., Fleming, M. T., Howells, N. L., Rabenhorst, M. M., Casselman, R., & Rosenbaum, A. (2006). *Effects of Method on Participants and Disclosure Rates in Research on Sensitive Topics. Violence and Victims, 21(4), 499–506.*

Riggs, W. E. (1983). The Delphi technique: An experimental evaluation. *Technological forecasting and social change*, *23*(1), 89-94.

Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International journal of forecasting*, *15*(4), 353-375.

Rowe, G., & Wright, G. (2001). Expert Opinions in Forecasting: The Role of the Delphi Technique. *Principles of Forecasting*, 125–144

Rowe, G., Wright, G., & Bolger, F. (1991). Delphi: A reevaluation of research and theory. *Technological forecasting and social change*, *39*(3), 235-251.

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research*, *6*(1), 1-21.

Snyder-Halpern, R., Thompson, C. B., & Schaffer, J. (2000). Comparison of mailed vs. Internet applications of the Delphi technique in clinical informatics research. In *Proceedings of the AMIA Symposium* (p. 809). American Medical Informatics Association.

Stanton, J. M., & Rogelberg, S. G. (2001). *Using Internet/Intranet Web Pages to Collect Organizational Research Data. Organizational Research Methods, 4(3), 200–217.*

Strabac, Z., & Aalberg, T. (2011). Measuring political knowledge in telephone and web-surveys: a cross-national comparison. *Social Science Computer Review*, *29*(2), 175-192.

Thomas, A., Paolo, C., & Patrick, L. (2020). The Future of Mobile Marketing in Italy: A Delphi Study. *Micro & Macro Marketing*, (1), 95-121.

Toma, C., & Picioreanu, I. (2016). The Delphi technique: methodological considerations and the need for reporting guidelines in medical journals. *International Journal of Public Health Research*, *4*(6), 47-59.

Turoff, M. (1970). The design of a policy Delphi. *Technological forecasting and social change*, *2*(2), 149-171.

Von der Gracht, H. A. (2012). *Consensus measurement in Delphi studies. Technological Forecasting and Social Change, 79(8), 1525–1536.*

von Winterfeldt, D. and Edwards, W. (1986). Decision Analysis and Behavioral Research. Cambridge University Press, Cambridge

von Winterfeldt, D. and Edwards, W., Decision Analysis and Behavioral Research, Cambridge: Cambridge University Press, 1986.

Webler, T., Levine, D., Rakel, H., & Renn, O. (1991). A novel approach to reducing uncertainty: the group Delphi. *Technological forecasting and social change*, *39*(3), 253-263.

Wibowo, S., & Deng, H. (2013). *Consensus-based decision support for multicriteria group decision making. Computers & Industrial Engineering, 66(4), 625–633.*

Zhang, W. Q., & Xi, Z. L. (2020). Application of Delphi method in screening of indexes for measuring soil pollution value evaluation. *Environmental Science and Pollution Research*, 1-11.

A. Appendix - Interviews

Participant Perspective:

Category AP: Participant background

In this category was collected information from the two interviewees, Participant X (PX) and Participant Y (PY), regarding their background in Web-Delphi.

PX: Participant with extensive experience in Web-Delphi, both as participant and moderator. This participant started working with Web-Delphi in 2015 and has been a participant in at least two studies.
PY: Participation in only one Web-Delphi process, but nevertheless studied and also designed a Web-Delphi process for a course in college.

Category BP: Communication between the moderator and the participant during the process

Here were asked questions in order to understand how the communication was done during the process and also to try to understand the strengths and weaknesses of this communication.

Both, PX and PY, received an invitation to participate in the process by e-mail and the rest of the communication was also done solely by e-mail. There was no other way to communicate with the moderator. The following Table 17 summarizes both participants' responses regarding the communication during the process:

Table 24 - Answers from Category BP

	Participant X	Participant Y
How did they know about the process?	E-mail sent from the moderator's personal e-mail address	E-mail forwarded by her thesis advisor (special case)

Official invitation	E-mail sent from the platform e-mail address	E-mail sent from the project team e-mail address
Doubts and clarification	The moderator has provided a specific e-mail address for doubts and there was also a user support e-mail from the platform available	The moderator has provided a specific e-mail address for doubts
Means of communication available	Only by e-mail	Only by e-mail

Both consider that E-mail is a great option, however, PX also values the possibility of having a live-chat where participants can communicate with the moderator during the process. It is important that this chat also be programmed. PY stated that the best way to ensure participation is to receive reminders while PX stated that the best way is to also receive reminders and e-mails but personalized to each participant.

Conclusions:

- Communication should be done by e-mail.
- Live-chat is a value feature to support participants during the process.
- Reminders are important to ensure participant's engagement.

Category CP: Registration and Log-in

In this category questions were asked regarding registration and log-in, to find out if any registration was necessary and also to understand if it was easy or difficult.

Both participants had to register on the platform and the data requested was only the e-mail address and a password. They both consider that the registration was easy and simple.

Conclusions:

- The registration process and the log-in should be simple enough to not fatigue the participant.

Category DP: Learning and training

Here the goal was to understand if any learning phases were done so that the participants know what they have to do during the process.

Regarding this topic, both participants only had access to one page of instructions, included at the beginning of the questionnaire, which explained in detail what was going to occur during the process and how to answer the questions. Was even reported that these instructions were done by the moderator.

PX: Noted that the written instructions were always sufficient, but nevertheless, PX would prefer to have a person explaining the steps of the process through a short interactive tutorial. PX says that this visual explanation is important.

PY: Considers that the written instructions given by the moderator were sufficient.

Conclusions:

- It is important to have an instruction page that explains in a clear way, the procedure of the process in question.
- Having a short tutorial that interactively explains the process was also considered important. However, it is a bit difficult to put into practice because the questionnaire designs vary from study to study, and this implies having to make a custom tutorial for each study.

Category EP: Platform User-Friendliness

In this category an analysis was performed regarding intuition and what the participants value in a platform from a user-friendliness perspective. The aspects discussed in this section were what types of electronic devices could be used to fill out the questionnaire, if there was auto-saving of the answers, if there was a progress bar so that the participant knew what stage of the questionnaire he was at, what type of questions were asked (qualitative, quantitative, radio buttons, dropdowns, etc.), and if the questions had instructions.

Table 25 - Answers from Category EP

	Participant X	Participant Y
Electronic Devices used	Computer, Smartphone, and Tablet	Computer
Auto-saving	Yes	Yes
Questions format	Multiple choice	Multiple choice
Instructions along with the questions	Yes	Yes
Progress bar	No	Yes

Both, PX and PY, only used the computer to respond the questionnaire, however, PX also experienced opening the survey on a Tablet and a smartphone and noticed that the Platform was not fully optimized for these two devices. The platform did not do an appropriate fitting to the Tablet and smartphone screens. PX pointed out that it is very important that the platforms are Tablet-friendly and especially mobile-friendly, because nowadays the smartphone is a much more convenient device to use and is becoming increasingly used in these form-filling occasions.

Regarding the auto-saving system, both participants stated that is very important to have this feature on a platform and PX also stated that the platform inform the participant of how much is left to conclude the questionnaire and which questions are missing, considering this really advantageous for a participant.

Both participants confirmed that each question contained an additional technical sentence about the method of answering the question in hand. They consider this important so that there is no doubt about what needs to be entered/ done in each field.

As for the progress bar, PY has vague recall that there was one at the top of the questionnaire while PX has not witnessed this feature in any of the processes that has participated in. Both participants find this feature advantageous since it helps give an idea about how much is left to finish the questionnaire.

Conclusions:

- The platform should be Tablet-friendly and especially mobile-friendly.
- It is very important to have an auto-saving system for the answers and for all the progress made in the questionnaire.
- It is important that once the quiz is not completely finished, the platform indicates which questions are left to complete and what percentage of the quiz the participant is currently on.
- Each question should contain a short instruction to ensure that it is answered properly.
- A progress bar while completing the questionnaire is very important for participants to know what stage of the questionnaire they are at.

Category FP: Feedback delivery and interpretation

In this section it was discussed and analyzed how feedback was received and interpreted.

Table 26 - Answers from Category FP

	Participant X	Participant Y
How was the feedback received?	By e-mail	By e-mail
Feedback was received in which format?	Written in the e-mail itself or in a pdf document attached to the e-mail	Had to click in a link on the e-mail and the feedback came in the application itself
Was the feedback received in real-time?	No	No

In which format was presented the data output?	Percentages and bar graphs	Bar graphs

Both participants stated that they had no problem in the interpretation of the feedback, everything was clear and simple. They always received the feedback by e-mail along with the convocation for the next round and was never received in real-time. They even stated that along with the questions there was a button where you can click and see a bar graphic with the results from the previous rounds or even the percentage of responses above each answer. PX and PY stated that visual outputs such as bar graphs, pie charts, graphics, etc., are really important to understand the feedback more easily and it is also valued that bar graphs and percentages are provided along with each question. Both also stated that their answers from the previous round were already typed in the subsequent round and they had the option to maintain that answer or to change if they wanted. They also considered this important.

Conclusions:

- There are three easy ways of receiving the feedback: Written in the email itself, on a document attached to the e-mail and along with each question.
- Feedback along with each question is valued.
- Visual outputs such as bar graphs, pie charts, graphics, etc. are important to interpret better the results.
- Answers from the previous round already typed in the subsequent round is also valued.

Category GP: Anonymity during the process

Since anonymity is one of the most important key features from Delphi, it was discussed if anonymity was ensured during the process.

Both participants said that anonymity was guaranteed during the process and that there was no way at all to know which person was behind each answer or comments presented on the feedback. Nevertheless, PY noticed in a comment the following phrase " As a health professional...", i.e., here it was known that the person who had made the comment was a health professional, however, it was not possible to know exactly which person commented that.

PX even stated that in a study was possible to enter, between the first and the second round, in a discussion room (sort of a forum) where the moderator and the participants could discuss, but all anonymously, i.e., each person with his or her platform ID.

Conclusion:

- Anonymity can be easily guaranteed if a specific ID is assigned to each participant as soon as they register on the platform.

- There must be a proper control by the moderator in the comment sections to prevent possible threats to anonymity.

Category HP: Navigation

Regarding this category the ease of navigating the platform during the process was discussed.

Both participants found that it was quite easy to navigate the platform by being able to even close and open the site, turn the computer off and on, and navigate other sites during the round.

It was also discussed whether the questions were all presented on just one page or if they were divided between several pages. PX noted that he has answered questionnaires where everything was on the same page and the questions were divided into different sections, but that has also answered a questionnaire that was divided over several pages. On the other hand, PY has only answered one questionnaire that was split over two pages. It was also discussed what each one's preference was regarding this aspect, and they both answered that it depends on the questionnaires. If it was a questionnaire where it made sense to consult the answers they had given in previous questions, then it was better to have everything on the same page, however, if the questionnaire is too long, it makes more sense to be divided into different pages. PX also added that if he used his smartphone to answer the questionnaire, he would prefer that the questions were all on the same page. With this, it is then concluded that, it is important to have both options which are to build a questionnaire on only one page and to build a questionnaire on different pages

Conclusions:

- It is important to have the possibility to build a questionnaire on a single page, but also to have the option to be able to split the questionnaire into different pages.
- The moderator has to consider whether or not it makes sense for the questionnaire to be on one page or split into different pages.

Moderator Perspective:

Category AM: Moderator background

In this category was collected information from the two interviewees, Moderator X (MX) and Moderator Y (MY), regarding their background in Web-Delphi.

MX: Has been working in the Web-Delphi area since 2016 by being involved in several Web-Delphi processes. Was involved as a moderator in only one process but supported other processes not only in the design but also in the analysis of results.

MY: Only moderated one Web-Delphi process but also has participated in two processes as a participant.

Category BM: Choose participants and invitation

The aim of this section was to understand how the participants are invited to the process. Table 20 shows how each moderator chose the people who participated in their processes and how they handled the invitations:

Table 27 - Answers from Category BM

	Moderator X	Moderator Y
How to choose participants?	Chose participants who had interest in the process' outcome (This reduces the drop-out rate)	Chose a very heterogenous group of participants due to its project subject and to have a wide range of different opinions from different stakeholders
How to collect the participants' contacts?	Posted in several social networks a warning that a study was going to be done and whoever was interested could sign up in the process by giving their contact in a google form. Also used internal networks of research partners to recruit more participants	Interviewed specific people to invite them to participate in the process and asked for their contact information. Also investigated who was publishing scientific articles related to the process subject area and collected their e-mails

Conclusions:

- Choose heterogeneous participants who have interest in the process outcome.
- Post the project in social networks and investigate who is publishing scientific articles related to the process subject to collect data of possible participants and invite them.

Category CM: Questionnaire construction

The construction of the questionnaires followed by its implementation in the platform was discussed in this section. Table 21 shows a summary of each moderator's responses regarding this matter:

Table 28 - Answers from Category CM

	Moderator X	Moderator Y
Electronic devices used to construct the questionnaire	Only computer	Only computer

Did the moderator use predefined themes to construct the questionnaire?	Yes, and also created pages from scratch to give instructions and make other announcements	No, created the questionnaire from scratch
Did the platform allow to customize the questionnaire?	Could only change words to bold, italic and underline. Could not insert logos and change the background.	Could only change words to bold, italic and underline. Could not insert logos and change the background.

MX and MY only used the computer to construct the questionnaire, however, both of them said that there are some problems with the screen fitting when using the phone or Tablet to respond the questionnaire. Both moderators value the existence of pre-defined themes. Although MY created the questionnaire from scratch, both of them agree that pre-defined themes make the questionnaire construction a lot easier and faster. They also value the fact of having the possibility to customize the questionnaire background and especially the possibility to insert logos and images into the questionnaire, since images can facilitate the understanding of questions. The necessity of having open text boxes for comments along with the questions was also stated on. These boxes are intended to collect comments from the participants. Both moderators consider this practice important in a Web-Delphi process. In addition to the comment box, they also valued the presence of instructions along with the questions so that participants are clear about the method of response, which in turn ensures that the data entered is valid. Finally, each moderator was also asked if there was any auto-saving system during the construction of the questionnaire. Both replied that there was no automatic system, only a manual one, but that it would be very important to have an auto-saving system at this stage of the process to ensure that all progress is properly saved even when there is some problem that is not possible to predict such as a sudden computer shutdown.

Conclusions:

- Moderator has to consider that the platform may not be Tablet-friendly or mobile-friendly. It is advisable to test first and see if it is necessary to give a warning in the convocation regarding this issue.
- It is important to have both options: possibility to create pages from scratch and the possibility of using pre-defined themes to construct the questionnaire.
- Background customization, insertion of logos, images, and instructions along with questions is important.
- Possibility of breaking the questionnaire into separated pages is also valued.
- Adding a comment section along with each question is a good practice.
- Auto-saving is important to mitigate possible progress losses.

Category DM: Learning and Testing

This category aims to understand how each moderator learned to handle the platform**.**

In this case MX and MY had the same learning method, both were inserted in teams of researchers who have experience in Web-Delphi and therefore they learned everything with the help of their colleagues. However, MY pointed out that the platform was very user-friendly, easy to handle, yet the platform also has video tutorials for learning. Both moderators, emphasized that video tutorials are the best way to learn for those who have never used the platform.

Regarding testing, MX stated that performing a pre-test before starting a Web-Delphi process is essential, not only to understand if the content presented works well and correctly, but also to understand if the platform is working according to what they want.

Conclusions:

- Video tutorials are important to learn how the platform works.
- Is important to do a pre-test to see if everything is working as expect.

Category EM: Rounds and stopping criteria

The aim of this category is to understand how the moderator manages and decides how many rounds are needed to conclude the process, how long will they take and also how is the conclusion of the process decided.

Both moderators advise conducting three rounds to study the stability of the responses, but, however, if the agreement between the participants' responses is significant enough (depending on the requirements demanded by the moderator) the process can be given as finished right after the second round. As for how long each round is active is dependent on each process. The time span of each round should not be large because it can tire the participants and also so that they are not temporally distanced from the answers they had given in the previous round. However, it also cannot be small because if it is a process involving a panel with a very large number of participants, it is likely to be more difficult to have everyone available to answer the questionnaire in the same short period of time. The best way to mitigate this problem is to manage the activity of the process, i.e., to see who has already responded, who has not yet responded and who is currently responding and manage the time of the round in this way.

Conclusions:

- The stopping criteria should be decided before the process starts, however, it can be changed according to the outcome of each round.

- Platforms should have a feature where moderators can manage who has responded, who is responding and who has not responded, in order to manage the process efficiently.

Section Category: Engagement and Reminders

Keeping the participants engagement is one of the biggest challenges, not only in Web-Delphi processes, but also in Delphi processes in general. In this section is discussed how this engagement is ensured.

MX and MY agreed that using reminders is the best way to keep participants engagement and MX also repeated that another good strategy is to choose participants who have a strong interest in the final outcome of the project. MY also added that one of the biggest problems that was related to the participant engagement was that the emails that were sent through the platform tended to be placed in the spam box. The solution to this problem was to send personal e-mails in which was included a warning regarding this spam box problem. Additionally, MY decided to personally call each participant by phone, to ensure their engagement in the process even further. Regarding the reminders, both moderators valued the performance of their respective platforms, saying that the platform had this functionality very integrated and intuitive. One of the advantages was being able to see who had already responded, who still had to respond and who was responding at that moment, thus making the management of reminders more efficient and effective. Also, by performing a pre-test, MY noted that the questionnaire should be held on the participants' native language to increase their engagement and to be easier for them to understand the questions.

Conclusions:

- The fact that the platform has an intuitive reminder feature where moderators can know the current status of each participant's activity is valued.
- Spam box is a possible threat to participant engagement: One way to mitigate this is to send e-mails from the moderator's personal address.
- One way to improve engagement is to call each participant to ensure their participation in the process.
- Questionnaires should be done on participants' native language, and if possible, should be available an option where the participant could choose the language of the questionnaire.

Category GM: Communication between the moderator and the participants

As in the interviews with the participants, here is discussed how the communication between the moderator and the participants is carried out, but this time from the moderator's perspective. Table 22 summarizes some responses regarding this section:

Table 29 - Answers from Category GM

	Moderator X	Moderator Y
How were the participants invited?	Invited by e-mail	Invited everyone by e-mail and to ensure everyone's participation also used phone calls
Doubts clarification	Provided a support e-mail for doubts and issues	Only by e-mail
Means of communication available	Only by e-mail	Only by e-mail but also phone calls in the invitation phase of the process
How is the process explained to the participants?	Prepared an instruction sheet (an initial instruction page) clear and objective so that people can complete the questionnaire correctly	In the inviting e-mail took the opportunity to explain how the process was going to be developed and made everything clear.

Both stated that evidently e-mail is a safe means to maintain the communication between the moderator and participant but both moderators also value the existence of a live-chat integrated in the platform where participants can communicate with the moderator during the process. MX even stressed that phone calls is not a good practice because “we are in a Delphi process where you are supposed to ensure anonymity as much as possible and be careful about data protection”. Besides that, phone calls are difficult to take place if there is a big number of participants involved in the process. Both moderators stated that the moderators should explain well how the process is going to be developed and how the platform works to avoid any doubts from the participants.

Conclusions:

- Communication should be done by e-mail.
- Live-chat is a value feature to support participants during the process.
- Moderators should explain well the process in the beginning to avoid possible doubts.

Category HM: Anonymity during the process

Again, since anonymity is one of the most important key features from Delphi, in this section was discussed with the moderators if anonymity was ensured during the process.

It was asked if it made sense to have a forum where participants could discuss among themselves their answers anonymously and both replied that this practice is not advisable because it could compromise anonymity to some extent, since the moderator in this case does not control the feedback. Both moderators said that anonymity was guaranteed in the processes they moderated and that there was no way for participants to communicate with each other. According to both

moderators, as soon as a participant registers on the platform he is assigned a unique ID, and all responses will be associated with that ID and not with the participant's name.

Conclusions:

- Participants should not have way of talking with each other.
- Every participant should have a unique ID to ensure anonymity and to track their responses.

Category IM: Data collection, interpretation, exhibition, delivery, and exportation

In this category was discussed how data is collected, exported, interpreted, and displayed.

Collection: It is possible to collect quantitative and qualitative data. Both stated that the participants could save and change their responses during each round, but as soon as the answers were submitted, the round ended automatically for that participant. This ensures that each participant submits his or her answers only once. The collection was in real-time, both had the option to see who already responded, who were responding at the moment and who did not respond.

Interpretation: Both moderators did not have difficulties in interpretating data, however, if there was any difficulty, MX stated that these difficulties are discussed among the research group. There were no difficulties because the questions were closed-ended, the only qualitative analysis needed was the comments.

Display: MX and MX stated that the platform did not generate any kind of graphics, just the distribution of responses. Participants had access to these distributions in the next round automatically and both moderators consider this feature important. However, the graphs could be generated manually by the moderator by exporting the data.

Delivery: In addition to the automatic feedback generated by the platform, each moderator also delivered a report via e-mail, where was presented a more detailed and more explanatory feedback with visual graphics, to ensure that the results were properly understood by all participants. MX and MY agree that this additional report with more detailed feedback is important, whether it is sent by e-mail, or presented on a home page in the questionnaire itself. Furthermore, MX also suggests that this detailed feedback could be presented throughout the questionnaire with the help of pop-up windows. Both sent feedback as soon as the next round started (feedback was never delivered in real-time).

Exportation: Both moderators stated that the platform only had the option to export data to Microsoft Excel, but MX would value it even more if the platform could generate the graphs in an integrated way without having to export the data to Excel. MX and MY agreed that this exportation feature is important, and MY stated that this exportation saves a lot of work.

Conclusions:

- There must be a way to ensure that each participant only have a submission per round.
- Platform should have an integrated feature for data exportation to perform detailed analysis.
- The Platform should perform an automatic analysis of the results and display them in graphical form, such as the distribution of responses in bar charts and not only in numerical format.
- Moderators should send a detailed report with the feedback to ensure that each participant understands the aim of the process and the respective feedback.