

RESEARCH

Open Access



# Usability enhancement model for unstructured text in big data

Kiran Adnan<sup>1\*</sup>, Rehan Akbar<sup>2,3\*</sup> and Khor Siak Wang<sup>1</sup>

\*Correspondence:  
kiranadnan@utar.edu;  
rehankb@yahoo.com

<sup>1</sup> Faculty of Information & Communication Technology, Universiti Tunku Abdul Rahman, Kampar, Malaysia

<sup>2</sup> School of Computing and Information Sciences, Florida International University, Miami, FL, USA

<sup>3</sup> Department of Computer and Information Sciences, Universiti Teknologi Petronas, Seri Iskandar, Malaysia

## Abstract

The task of insights extraction from unstructured text poses significant challenges for big data analytics because it contains subjective intentions, different contextual perspectives, and information about the surrounding real world. The technical and conceptual complexities of unstructured text degrade its usability for analytics. Unlike structured data, the existing literature lacks solutions to address the usability of unstructured text big data. A usability enhancement model has been developed to address this research gap, incorporating various usability dimensions, determinants, and rules as key components. This paper adopted Delphi technique to validate the usability enhancement model to ensure its correctness, confidentiality, and reliability. The primary goal of model validation is to assess the external validity and suitability of the model through domain experts and professionals. Therefore, the subject matter experts of industry and academia from different countries were invited to this Delphi, which provides more reliable and extensive opinions. A multistep iterative process of Knowledge Resource Nomination Worksheet (KRNW) has been adopted for expert identification and selection. Average Percent of Majority Opinions (APMO) method has been used to produce the cut-off rate to determine the consensus achievement. The consensus was not achieved after the first round of Delphi, whereas APMO cut-off rate was 70.9%. The model has been improved based on the opinions of 10 subject matter experts. After second round, the analysis has shown majority agreement for the revised model and consensus achievement for all improvements that validate the improved usability enhancement model. The final proposed model provides a systematic and structured approach to enhance the usability of unstructured text big data. The outcome of the research is significant for researchers and data analysts.

**Keywords:** Unstructured data usability, Pragmatic quality, The usefulness of unstructured text, Big data analytics

## Introduction

The complexity of unstructured text on a technical and conceptual level makes effective data preparation essential for big data analytics. These complexities diminish the usability of unstructured text for big data analytics. The unstructured data usability problems are not effectively considered by the current approaches and techniques of big data analytics [1–4]. In other words, data delivery has been more focused, whereas the importance of data “use” has been considered less [5]. Most existing prescriptive and

descriptive approaches for data usability only handle structured data [6, 7]. Further, the existing literature highlighted the significance of unstructured data usability during the preparation of data for analysis [2, 3]. To the best of our knowledge, the existing literature lacks structured and systematic solutions to enhance the usability of unstructured text.

Therefore, unstructured text usability has been addressed to fulfill the research gap, and a usability enhancement model has been developed using a systematic literature review (SLR) [6]. The SLR achieved the objective of investigating the unstructured text usability dimensions, addressing the usability problems with enhancement factors i.e. usability determinants, and formulating the usability rules based on the dimensions and determinants according to the problem at hand. The usability enhancement model comprising of the usability issues, key processes, usability improvement factors, and formulation of usability rules to address the usability of unstructured text in big data analytics is a significant contribution of this research. In order to address the usability of unstructured text in big data for better analytics, the proposed usability enhancement model considers “subjective intentions” as a major aspect and proposes a systematic approach.

In this paper, the usability enhancement model for unstructured text has been validated with the experts’ opinions using the Delphi technique. The validation of the model or framework refers to the evaluation procedure to judge fitness for the purpose [8]. It is an important way to measure the usefulness of a research outcome [9]. Further, the experts’ opinion is one of the famous validation methods in qualitative research [10, 11]. The Delphi technique has been considered the most suitable method for gathering opinions from Subject Matter Experts [12–14]. Therefore, the usability enhancement model for unstructured text, its components, and its working have been validated using the Delphi technique to ensure the correctness of the model components and increase its confidence and reliability. The Subject Matter Experts from different countries have participated in different rounds of Delphi. The usability enhancement model has been revised and improved in rounds until a consensus has been reached. This research contributes to the literature with a validated model and systematic approach to enhance the usability of unstructured text for big data. The validated usability enhancement model provides a structured approach to address the usability issues of unstructured text considering subjective intentions.

### **Usability enhancement model**

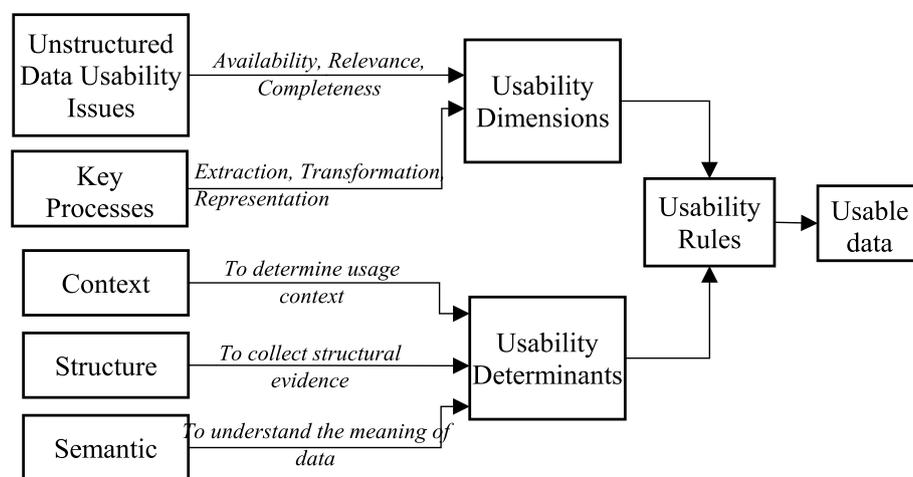
Data usability is a subjective dimension based on the user’s assessment to measure data usefulness [15]. It varies between different stakeholders’ requirements due to the various tasks and different interpretations of data values [16]. In this regard, topic reconceptualization for unstructured data usability has been used to identify three important aspects of unstructured data usability: addressing users’ needs, handling unstructured data issues, and the ability to make unstructured data usable [6]. Most of the literature is related to structured data usability. As structured and unstructured data are different in schema and structure, the usability dimensions are not usable in the same way [17]. However, the prominent existing literature has not adequately investigated these two aspects of data usability, i.e., addressing users’ needs and handling unstructured data issues. Likewise, the existing literature lacks unstructured text data usability

enhancement strategies or determinants. Such limitations of the existing literature highlight the need to identify the usability dimensions for unstructured text data and investigate the usability enhancement factors.

Therefore, a systematic literature review has been conducted to identify unstructured data usability dimensions and determinants in big data text analytics [6]. The purpose of conducting SLR was to investigate the usability issues of unstructured text data in the big data analytics process and identify facilitating factors to enhance the usability of unstructured text data. Further, a usability enhancement strategy has been developed, and a usability enhancement model has been proposed based on the findings of SLR, as shown in Fig. 1. Three major components of the usability enhancement model have been identified from the findings of the SLR, which are usability dimensions, usability determinants, and usability rules [6].

The first component of the usability enhancement model, i.e., usability dimensions, comprised of two subcomponents, “usability issues” and “key processes”. Various usability issues have been identified and divided into three major categories such as data availability, relevance, and completeness. The results of the SLR also reveal that three key processes in the pipeline for big data analytics are where the observed usability problems occur, such as data extraction, transformation, and representation. One usability issue at the particular key process forms a usability dimension for unstructured text. For example, data availability at data extraction creates a usability dimension. The first subcomponent inputs usability issues, while the second subcomponent provides input in the form of process selection.

Based on the findings of SLR, the second major component of the usability enhancement model, usability determinants, has been derived. Different facilitating factors and their variants have been identified and divided into three categories of usability determinants such as context, structure, and semantics, as shown in Fig. 1. The findings of the SLR have demonstrated that various types of contextual information address users’ requirements and understand data & user perspectives such as user context, data context, and domain knowledge. This contextual information in the data extraction process



**Fig. 1** Unstructured text usability enhancement model [6]

helps the user to identify useful information. Likewise, semantics has been identified as another usability determinant in this research that is helpful for the user to understand the meaning of text according to the surrounding scenario. The extraction of contextual, structural, and semantic information as usability determinants helps the user to make the unstructured text data usable. Therefore, these facilitating factors have been grouped as usability determinants and formulated as the second component of the model.

It has been analyzed from the findings of SLR that usability issues of unstructured text should be handled at the early stages of big data analytics to improve the efficiency of the analytics process. The identified facilitating factors are the determinants of usability to overcome the usability challenges of unstructured text. However, selecting appropriate usability determinants for the problem at hand is important. The relationships between usability dimensions and determinants have been assessed using the data retrieved from a few selected SLR studies. The synthesis identified the appropriate determinant(s) for each dimension. The findings on relationships between usability dimensions and determinants have been used to design the rules. These usability rules take inputs from the first component, i.e., usability dimensions, and the second component, i.e., usability determinants, and form the third component of the proposed model, as shown in Fig. 1.

Before formulating usability rules for a particular scenario, reference data needs to be identified according to the usability dimension such as input data issues and target output. However, the usability rules can be modified or generated according to the problem and task. The steps to formulate usability rules have been defined as follows:

- 1) Select a usability dimension that deals with identifying the usability issue of unstructured text data and selecting key process.
- 2) Identify the reference data, which includes data issue and target.
- 3) Explore the appropriate variants of usability determinants from the main determinant.
- 4) Use the identified information to form the usability rules.

Based on the abovementioned steps, usability rules can be generated. For example, heterogeneity is the input issue for data availability at data collection from different sources. This indicates that the usability dimension will be 'data availability at data collection.' The identification of reference data includes data issues, i.e., heterogeneity and target can be getting structured intermediate form of data. Based on the mapping from SLR, the usability determinant can be 'context' with two variants, i.e., data context and user context. Therefore, one example rule for this scenario can be "The relationship between data and user context should be identified to improve the data availability".

Hence, the usability enhancement model for unstructured text data integrates the usability dimensions, determinants, and rules. The proposed model consists of three major components: usability dimensions that use usability issues and key processes as input, usability determinants that use inputs from context, structure, and semantics, and usability rules that use the outputs of the first two components as input. In this paper, the unstructured text usability enhancement model has been validated using experts' opinions. The subsequent section presents a detailed discussion of the model validation process of this research.

### Model validation

The Delphi technique has been considered a well-suited means of consensus-building to seek an opinion from a panel of experts [18]. It has been used widely in different research domains such as medical and healthcare, management sciences, social sciences, and many others [19–22]. It has also achieved popularity in information systems and software engineering research domains due to its relevance, support, suitability, and compliance with the nature of research studies conducted in these areas [23, 24]. It is an appropriate method to achieve the opinion of Subject Matter Experts and synthesize the experts' opinions into a usable product [25, 26]. There is no standard procedure or variation in procedures of the Delphi, but there are four common characteristics of a Delphi such as anonymity, the number of rounds, controlled feedback, and data summary & analysis to generate valid results [27–29]. Further, Delphi invites experts from different locations to participate in this research rather than in-person interviews from specific countries. So, it is a viable tool for learning from highly experienced practitioners in the least amount of time. The Delphi technique has been considered appropriate to validate the usability enhancement model because it allows the experts to judge and evaluate the proposed model, guides them to identify the issues in the proposed solution, and achieves consensus concerning problem resolution [8, 30–32]. These factors and properties of the Delphi technique make it an appropriate and suitable method for the present research work according to its format and nature. The primary goal of adopting Delphi is to validate the proposed model until a consensus achievement.

### Protocol design

At first, the Delphi protocol was designed to determine the expert selection criteria, sampling strategy, panel size, panel group, number of rounds, mode of interaction, initial questionnaire, data analysis method, and consensus measurement method. The experts have been selected based on these four fundamental requirements of knowledge, expertise, willingness, and communication skills [8, 33].

Following the guidelines, the Subject Matter Experts selection criteria for Delphi in this research have been defined as follows:

- A high degree of knowledge and experience in the subject matter: Participants have experience in academics or industry but preferably both, at least in one of the following domains such as big data analytics, data science, data/text analytics, information extraction, text mining, or unstructured data analytics.
- The participants have peer-reviewed journal publications in the relevant field. The research interests of the participants include the area of big data analytics or field experience that must be relevant to big data analytics.
- Participants have the capacity and willingness to participate in this Delphi.
- Participants have enough time to participate in this Delphi.
- Participants have effective communication skills.

Other than the four fundamental requirements, extensive theoretical knowledge and industry experience should also be considered [34]. Therefore, this research has

considered a heterogeneous group of panel experts. The preference has been given to the experts who have experience in both, i.e., academia and industry. Thus, the Subject Matter Experts in this Delphi belong to either academia or industry, or both.

The panel size of Subject Matter Experts is one of the major concerns in the Delphi technique. The size and constitution of the panel depend on the nature of the research and its dimensions [35]. There is no general agreement or consensus on the panel size for Delphi [33, 36]. Table 1 summarizes the sample sizes and panel groups used in different studies where the Delphi technique has been applied for model or framework validation. The potential sample size to validate the usability enhancement model has been selected based on the literature. The sample size of 8–15 has been considered a sufficient target sample size to conduct this Delphi. However, it has been taken care to include a maximum number of participants rather than restricting the sample size to a certain limit.

Purposive snowball sampling has been used to identify the relevant experts based on the selection criteria and to access more experts using the social networks of identified experts [37, 38]. Electronic mail has been considered the most commonly used mode of interaction in Delphi due to its quick turnaround time and expediency [27, 39]. Also, the Subject Matter Experts in this Delphi belong to different countries, therefore, emails are considered a more suitable interaction mode. This Delphi has used an open-ended structured questionnaire in the initial round. The main purpose of adopting broad initial questions was to seek more insights and collective intelligence of the research participants on the topic [14, 28]. The number of rounds in the Delphi technique is variable and depends on the purpose of the research. Reaching consensus is considered the stopping criterion of any Delphi study [40]. This Delphi has used the consensus achievement method to determine the number of rounds for model validation. At each round of this Delphi, data has been collected from the experts, analyzed, and then consensus has been measured. The interviewing method has been used for data collection at each round of Delphi. The structured interview approach has been considered a well-designed and managed approach that maintains the focus of the participants in the particular area of discussion [41, 42] using an objectively designed questionnaire.

The systematic process of directed content analysis has been used to transform the level of abstraction from low to high and transcribe the interview data in all rounds [43]. The directed content analysis method is a deductive approach and is suitable when the categories are predetermined, and questions are targeted to explore the participants' experience in a particular category [44, 45]. The outcome of directed content analysis provides supportive and non-supportive evidence for a theory [44, 45].

**Table 1** Panel sizes for model/framework validation in Delphi studies

Study	Panel size	Panel group
[8]	3	Homogeneous
[30]	5	Homogeneous
[31]	6	Homogeneous
[32]	11	Heterogeneous

The consensus measurement methods include mean/median ranking, a certain level of agreement, the Average Percent of Majority Opinions (APMO), interquartile range, coefficient of variation, and post-group consensus [46]. The selection of appropriate consensus methods depends on the research problem, analysis results, and stopping criteria. In Delphi, where a certain level of agreement or disagreement is received in a predetermined range, the consensus is assumed to be achieved, known as APMO. This Delphi technique has used APMO method to produce the cut-off rate to determine the consensus achievement [46–49]. The APMO cut-off rate as a consensus method has been considered appropriate in this research due to its predetermined range as the directed content analysis determines the supportive and non-supportive responses.

The following formula has been used to determine the APMO cut-off rate for each round of the Delphi:

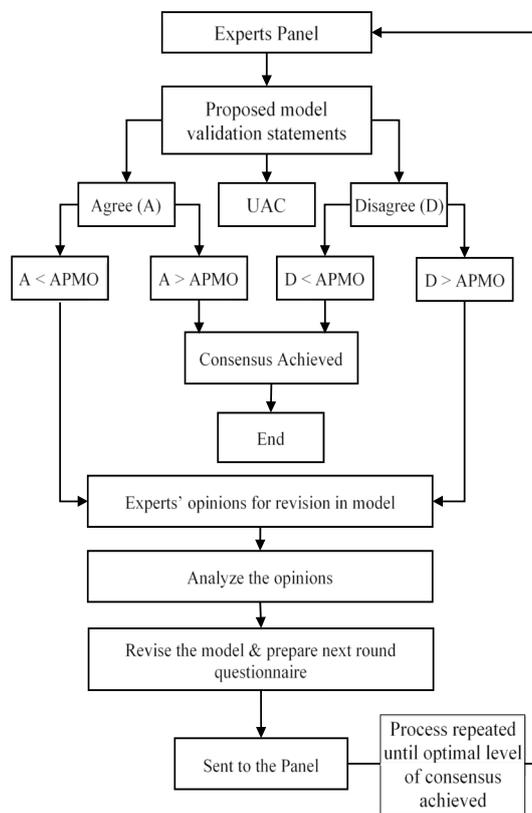
$$APMO = \frac{(Majority\ Agreements) + (Majority\ Disagreements)}{Total\ Opinions} \times 100$$

There are different schools of thought on measuring the majority opinions for agreements or disagreement [50]. Usually, a majority is defined as a percentage above 50% [46, 50]. However, the level of agreement as a measure for the APMO cut-off rate is defined according to the requirements of analysis and research study [46, 51]. However, a higher number of majority agreements and stability in experts' opinions are two important factors. Therefore, this Delphi used 80% for majority agreement and 20% for majority disagreements. This majority percentage has shown that the statements with more than 80% agreement score have majority agreed whereas the statements with more than 20% disagreement have considered majority disagreed.

In the APMO method, the consensus is defined as a percentage higher than the APMO cut-off rate. Statements that could not reach this APMO cut-off rate have been included in the next round of Delphi. The process continues until all statements reach the APMO cut-off rate. The process of APMO cut-off rate calculation and implementation, followed in this research is shown in Fig. 2. The responses from the expert panel have been coded as agreement, disagreement, and UAC (Unable to Comment) to determine the cut-off rate in each round. When a statement attained a percentage of agreement higher than the APMO rate, it has been considered that the consensus is achieved for the particular statement. Otherwise, the model was improved according to the opinions, and the next round of questionnaires was prepared.

#### **Expert selection process—knowledge resource nomination worksheets (KRNW)**

The selection of experts is an important step in the Delphi technique as it directly affects the quality of results [36]. Therefore, the guidelines by [33, 35] have been followed to identify the relevant experts for expert identification and selection. The experts who met the selection criteria have been identified and invited to participate in this Delphi. In alignment with the expert identification process [35, 52, 53], this Delphi follows a multistep iterative process to identify the experts according to the criteria known as Knowledge Resource Nomination Worksheets (KRNW). The process of KRNW comprises four steps, as shown in Fig. 3.



**Fig. 2** Consensus achievement process

Step 1 Preparing KRNW	<ul style="list-style-type: none"> <li>• Mention high level required details</li> <li>• Mention experts selection details</li> </ul>
Step 2 Populating KRNW	<ul style="list-style-type: none"> <li>• Populate KRNW with names and emails</li> <li>• Rank the experts</li> </ul>
Step 3 Nominating Experts	<ul style="list-style-type: none"> <li>• Invite experts</li> <li>• Prepare list of agreed experts</li> </ul>
Step 4 Ranking Experts	<ul style="list-style-type: none"> <li>• Rank experts according to experience</li> </ul>
Step 5 Inviting experts	<ul style="list-style-type: none"> <li>• Send consent form to agreed experts</li> </ul>

**Fig. 3** KRNW steps and activities

**Prepare KRNW**

The primary purpose of KRNW is to make the expert identification and selection process systematic. KRNW sheets are helpful in categorizing and managing appropriate participants [35, 52, 53]. At the preparation stage, important high-level details have been mentioned without specifying an expert’s name.

### **Populate KRNW**

The experts have been identified from three different sources: social media search, universities/industries, and literature, as shown in Table 2. Therefore, the initial worksheet has been populated with the names of identified experts from both industry and academia. Each source provided a list of experts. At first, the personal list of contacts was explored to find and fill the KRNW with experts' names. Next, the experts who meet the selection criteria were identified through ResearchGate ([www.researchgate.net](http://www.researchgate.net)), Google Scholar ([www.scholar.google.com](http://www.scholar.google.com)), and LinkedIn ([www.linkedin.com](http://www.linkedin.com)). Similarly, experts from different universities and research centers have been searched to identify the experts. The experts have also been selected from different journals of related literature. A total of one hundred and thirteen experts have been identified from all three categories of sources. Among these three categories, 88 experts have been identified from social media, 18 from organizations, and 23 are listed from the third category, i.e., literature. The experts overlapping in all three categories are ranked first, and experts overlapping in two categories ranked second and remaining are listed similarly. Hence, a total of 12 experts have been identified as first rank, 21 in the second rank, and the remaining 57 were ranked third. However, this ranking aimed to identify the most relevant and suitable experts without affecting the participation criteria.

### **Nominate experts**

After identifying relevant experts, an invitation letter describing the research study has been sent to 90 experts. As this study followed purposive snowball sampling, the experts were also asked to nominate experts in their circle. Twelve experts agreed to participate in the identified list of experts, whereas eight experts were excused to accept participation invitations due to their busy schedules. Two more experts were nominated by experts who agreed to participate. However, no response has been received from the remaining 68 experts invited to participate. Eventually, a list of a total of fourteen experts was prepared.

### **Ranking experts**

The demographic information has been collected from the 14 experts. Later, two experts did not respond to the email or even reminders. Finally, the ranked list of 12 experts

**Table 2** Expert identification

Social media	Organizations	Literature
1. ResearchGate ( <a href="http://www.researchgate.com">www.researchgate.com</a> ) 2. Google Scholar ( <a href="http://www.scholar.google.com">www.scholar.google.com</a> ) 3. LinkedIn ( <a href="http://www.linkedin.com">www.linkedin.com</a> ) 4. Facebook ( <a href="http://www.facebook.com">www.facebook.com</a> )	Universiti Teknologi Petronas, Universiti Malaya, Universiti Sains Malaysia, Universiti Teknologi MARA, Taylor's University, Universiti Malaysia Terengganu, National University of Science & Technology, University of Engineering & Technology, Government College University, COMSATS, University of Bahawalpur, University of Lahore, University of management and technology, University of Karachi	1. List of corresponding and first authors from the identified relevant literature 2. Other experts from relevant journals such as the Journal of Big Data, IEEE Transactions of Big Data, Big Data Research, IEEE Access, Journal of Data and Information Quality, Big Data, Applied System Innovation, Data Science Engineering, Journal of Information and Knowledge Management Systems, Applied System Innovation

who agreed and fulfilled the expert selection criteria has been finalized, and the selected experts were informed.

### **Inviting the experts**

The consent form has been sent to the participants for their honest appraisal and commitment. It has been mentioned in the invitation letter that participation is completely voluntary and will be kept anonymous. At this stage, two experts quit participation due to their busy schedules. Therefore, the first questionnaire has been sent to the ten experts for model validation.

### **Experts panel**

According to the selection criteria, the experts' panel comprised researchers and/or academicians from the related domain. However, experts with industry and/or academic backgrounds have been preferred for fair and reliable results. Among the ten experts, two (E1 & E2) belong to the heterogeneous groups, i.e., having experience in both industry and academia. However, all experts belong to different geographical areas that generated more versatile results. The selected experts were highly skilled and experienced, with a minimum of 5 years and a maximum of thirty years of experience in the related field. They have a strong research profile of peer-reviewed journal articles in the relevant field of the present research. The names of the panel experts are kept confidential due to personal data protection and confidentiality. Therefore, codes have been given rather than names. Table 3 summarizes the profiles of the experts. The experts with industry backgrounds, E1 & E2, have thirteen years and eleven years of experience, respectively. At the time of the interview, E1 was serving as vice president in a data science consultancy and management company, whereas E2 in NASA and OnTrak was a senior data scientist. The experts with academic backgrounds include one professor, two associate professors, five assistant professors, and two senior lecturers. Experts from different countries like USA, Malaysia, Pakistan, Mauritius, UAE, UK, and Germany have shown interest, whereby 40% from Pakistan, 20% from Malaysia, 20% from the USA, 10% from Mauritius, and 10% from UAE have participated in this research.

**Table 3** Profiles of expert panel of Delphi study

Expert	Position	Country	Background	Experience in years	Publication
E1	Vice President	USA	Industry	13	11
	Associate Professor		Academia	10	
E2	Senior Data Scientist	USA	Industry	11	6
	Professor		Academia	23	
E3	Senior Lecturer	Mauritius	Academia	15	5
E4	Associate Professor	Malaysia	Academia	30	20
E5	Assistant Professor	Pakistan	Academia	16	3
E6	Assistant Professor	UAE	Academia	5	20
E7	Assistant Professor	Pakistan	Academia	6	5
E8	Assistant Professor	Pakistan	Academia	8	6
E9	Assistant Professor	Pakistan	Academia	13	4
E10	Senior Lecturer	Malaysia	Academia	12	4

### Delphi process

In this research, the Delphi process was completed in three steps: prepare & define, identify & invite, and collect & analyze, as depicted in Fig. 4. The Delphi protocol was developed at the first step of prepare & define, and the first questionnaire has been drafted. After completing the first step, the Subject Matter Experts were identified and invited according to the selection criteria and design considerations determined in the Delphi protocol using the KRNW process. The experts have been communicated to provide feedback on the proposed model following the interview questions. The collected data, i.e., experts' opinions, have been stored, and the consensus has been measured in the third step. The statements of disagreements and improvement have been analyzed, and a questionnaire on the revised model was prepared for the next round of Delphi. This process continued until the consensus achievement for all the statements for the preliminary model and the model had been validated. The consensus has been achieved after the second round of this Delphi. Therefore, the improved model after round two was considered a validated and accepted model.

The process to generate results from both rounds of Delphi has been discussed in the following subsections.

#### First round of Delphi

The feedback on an open-ended questionnaire from the expert panel has been analyzed using the directed content analysis method, as shown in Fig. 5. Deductive coding has been applied for predetermined codes and to identify new codes in directed content analysis. In the first step of preparation in directed content analysis, the key concepts have been identified from the first questionnaire for the preliminary model and turned into the initial codes. The interview transcripts, i.e., the questionnaire, have been used as a unit of analysis. The categorization matrix has been designed for the identified four main categories of questions and related subcategories [45, 54]. These categories include (1) identification of factors, (2) placement of factors in components, (3) identification of factors for usability rules, and (4) arrangement of components to formulate the model. A flat coding frame has been used in this round of Delphi as the specificity and importance of each node lie at the same level [55, 56]. The question categories have been used as node categories and questions under each category were used as codes. The supportive

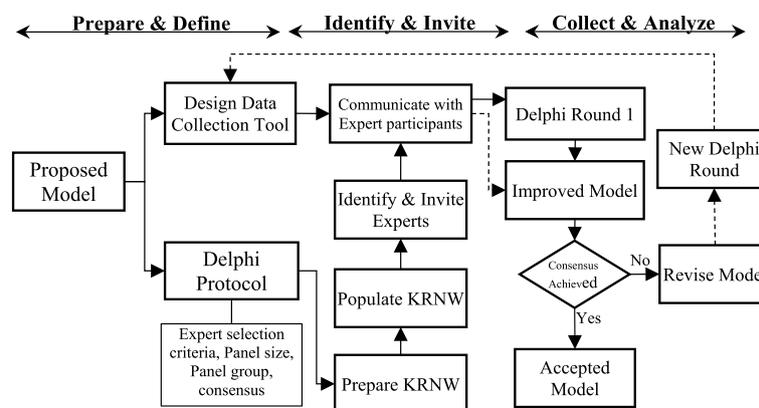
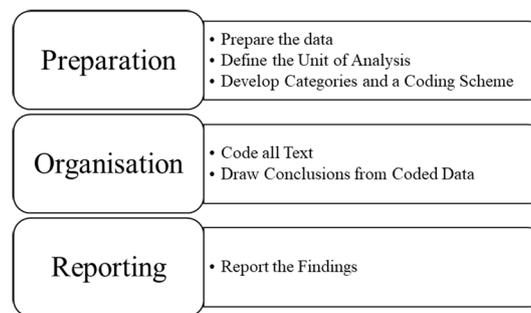


Fig. 4 Delphi process



**Fig. 5** Directed content analysis process. Source: [45, 54]

experts' responses (agreements) have been coded under the relevant code whereas the disagreement and UAC statements have not been coded. However, new codes from agreements and disagreements were identified and stored under related categories.

At the second step of analysis i.e., organization, the categories and codes from the questionnaire have been identified and stored in NVivo. The responses of ten experts were stored in NVivo for analysis, the case nodes were examined to find patterns, and data was coded based on a prior coding scheme. New codes have also been identified and stored under relevant categories. In the reporting phase, the results of the first Delphi round have been presented and data for the second round has been collected. Then APMO as the consensus method to measure the agreements and disagreements has been applied, as discussed in Section. "[First Round of Delphi](#)".

### ***Second round of Delphi***

In the second round of this Delphi, the second-round questionnaire has been prepared according to the improvements in the model in the first round of Delphi. The second-round questionnaire included questions related to all four categories where consensus was not achieved in the first round. The experts of the first round were requested again to participate in the second round of Delphi, whereas nine experts out of ten participated in this round. The results of the first round and revised model have been shared with the expert panel along with the second-round questionnaire. The participants have been requested to provide in-depth responses to the open-ended questions according to their perception, knowledge, and experience.

All responses from the expert panel on the second-round questionnaire have been then reviewed for completeness and accuracy and stored in a secured storage location. The responses of experts have been analyzed to identify the disagreements and improvements suggested. Further, the analysis of the second round has been conducted based on the prior coding in NVivo, and the APMO cut-off rate has been applied as a consensus measurement method. The results of the second round have been discussed in Section. "[Second Round of Delphi](#)".

## **Results and discussion**

This section presents the results of both rounds of Delphi to validate the proposed usability enhancement model.

### First round of Delphi

The feedback of the experts has been analyzed for all four categories of questions in the first round questionnaire, and it has been determined whether the consensus has been achieved. The responses under the agreement category have been coded in NVivo for the predetermined list of codes whereas the disagreements have been listed to calculate the majority of disagreements for APMO. The questions with 80% or above agreement have been counted as majority agreements, while the questions with more than 20% disagreement have been added to the majority disagreements. Then, the majority agreements, disagreements, and total opinions were summed up to measure the APMO cut-off rate.

Based on the findings, 199 majority agreements while 13 majority disagreements have been received. Regardless of the majority agreements and disagreements, 299 total experts' opinions have been received. APMO cut-off rate for the first round of Delphi has been calculated i.e. 70.9%, as presented in Table 4. The APMO rate shows that if the percentage of agreements for each question is greater than 70.9%, then it is considered that consensus has been reached.

Based on the APMO cut-off rate, a consensus has been measured for each question, as presented in Table 5.

According to the results of the first round, the components of the model that could not reach a consensus have been improved and validated again after revision in the second round of Delphi. The experts' opinions on the first round questionnaire are discussed in the subsequent sections based on the results given in Table 5.

### Identification of factors

This category of questions investigated the correctness of identified usability issues, key processes, and determinants. Out of ninety responses from experts on the identification of factors (usability dimensions and determinants), eighty-eight responses have been coded whereas two responses were UAC. The agreements have been added to the predetermined codes in NVivo whereas the disagreements have not been coded.

Data availability, data relevance, and data completeness were three usability issues identified as unstructured text challenges in big data analytics [6]. Experts' opinion was collected on the correct identification of these usability issues. All experts agreed with the data availability issues of unstructured text data in the analytics process. Seven out of ten experts agreed on the data relevance issues, whereas three experts disagreed. Similarly, the data completeness issue of unstructured data has been approved by eight experts as a usability issue.

Furthermore, three key processes have been identified along with usability issues to form usability dimensions for unstructured text in big data analytics such as extraction, transformation, and representation [6]. Eight of ten experts agreed with data extraction

**Table 4** APMO measurement for Delphi round one

Total majority agreements	199
Total majority disagreements	13
Total opinions	299
APMO cut-off rate	$= [(199 + 13) / 299] * 100 = 70.9\%$

**Table 5** Results of the first round

Questions	Agreed		Disagreed		UAC		O	C
	#	%	#	%	#	%		
Identification of factors								
Availability as a usability issue	10	100	0	0	0	0	10	Yes
Relevance as a usability issue	7	70	3	30	0	0	10	No
Completeness as a usability issue	8	80	2	20	0	0	10	Yes
Extraction as a key process	8	80	2	20	0	0	10	Yes
Transformation as a key process	10	100	0	0	0	0	10	Yes
Representation as a key process	8	80	0	0	2	20	8	Yes
Context as a usability determinant	9	90	1	10	0	0	10	Yes
Structure as a usability determinant	10	100	0	0	0	0	10	Yes
Semantics as usability determinant	9	90	1	10	0	0	10	Yes
Placement of factors								
Availability in model	8	80	2	20	0	0	10	Yes
Relevance in model	9	90	1	10	0	0	10	Yes
Completeness in model	9	90	1	10	0	0	10	Yes
Extraction in model	9	90	1	10	0	0	10	Yes
Transformation in model	10	100	0	0	0	0	10	Yes
Representation in model	9	90	1	10	0	0	10	Yes
Context in model	8	80	2	20	0	0	10	Yes
Structure in model	9	90	1	10	0	0	10	Yes
Semantics in model	8	80	1	10	1	10	9	Yes
Formulation of usability rules								
Determinants for dimensions	64	71	0	0	18	20	72	Yes
Arrangement of model components								
Arrangement of usability issues	6	60	4	40	0	0	10	No
Arrangement of key processes	7	70	3	30	0	0	10	No
Formulation of usability dimensions	8	80	2	20	0	0	10	Yes
Arrangement of usability determinants	8	80	2	20	0	0	10	Yes
Formulation of usability determinant	7	70	3	30	0	0	10	No

O Total Opinions, C Consensus Achieved?

as a key process in the analytical pipeline. It has been accepted that extracting the useful and discarding the useless data is vital to improve the usability of unstructured data. In the case of the identification of data transformation as a key process, all experts agreed that the effective transformation of unstructured text into a usable form is important to improve the process of analytics. It has been analyzed from the opinion of eight experts that effective data representation is equally important as data extraction and transformation whereas two UAC responses were received as two experts did not respond to these questions. As presented in Table 5, it has been inferred that the identification of all three key processes has achieved the consensus. However, a consensus was not achieved for data relevance as a usability issue, whereas the remaining two usability issues, availability, and completeness, have achieved the consensus. Therefore, all the questions regarding the identification of factors for the proposed model have been approved by the experts except 'relevance'. The responses from the experts on this node have been analyzed. It has been stated by E2, "no, the irrelevant data can be filtered during feature extraction". Also, E5 mentioned, "not too much because if irrelevant data is present, we

can further clean it and it might be possible it is useful as well". E6 stated in the expert opinion that "I think relevance could have two dimensions, i.e., from a user perspective and available data perspective".

Experts' opinions have shown that irrelevant data needs to be filtered, and it requires an effort to make it usable. Considering these aspects, the relevance could have two perspectives i.e. data perspective and user perspective. Therefore, the component "usability issues" in the proposed model has been improved with two subcategories of relevance and illustrates the change in the respective subcomponent. The experts' opinions on the identification of usability determinants have been collected and analyzed, as presented in Table 5. In the opinion of all experts except E4, context and its variants significantly influence the enhancement of unstructured text usability. E4 mentioned, "different types of context might end up in failure of data integrity in final structured data". The usability of data highly depends on the user's requirements. However, integrity is a significant constraint to be considered while improving usability when integrating data. In this regard, usability rules should be designed accordingly after considering all the constraints to overcome such issues. The responses also indicated that all experts agree with the structure as a usability determinant. For semantics as a usability determinant, nine experts agreed that semantics significantly addresses usability issues. However, no explanation has been provided for one disagreement. The analysis of experts' opinions has shown that consensus has been achieved for the identification of all usability determinants. Therefore, no improvements/revisions were required in this component.

### ***Placement of factors***

This category of questions dealt with the appropriate placement of identified factors into the components. The responses of ten experts on nine questions of this category have been coded and summarized in Table 5. The effective characterization of data availability has been approved by the eight experts, while two experts disagreed. One reason for disagreement was presented as "must add some subprocess for handling data scalability". The response has been coded as disagreement and a subprocess of data scalability has been added. Nine positive and one negative response have been received for the data relevance usability issue from E3 such as, "no, there is no significant correlation/association". It is mentioned that relevance had not been significantly represented in terms of usability in the proposed model. Similarly, nine agreements and one disagreement have been received for data completeness as a usability issue in the proposed model. It was mentioned in the answer that data completeness has not been significantly described in the model.

In response to the placement of data extraction and data representation as key processes, nine positive and one negative feedback have been received. All experts agreed that usability issues during the transformation process would show an affirmative impact on usability enhancement in the model. Eight out of ten respondents agreed with the organization of usability determinants in the model, as presented in Table 5. It has been agreed by nine out of ten experts that the proposed model identified the role of structure for unstructured data significantly in terms of data usability. Similarly, eight positive responses have been received from the ten experts about the role of semantics as a usability determinant in the proposed model. One disagreement and one UAC have

also been received. A consensus has been achieved for the placement of all the usability dimensions and determinants in the proposed model. Therefore, no improvements were required for this category.

#### **Formulation of usability rules**

Different usability determinants with variants are required for each dimension to formulate usability rules. This category of questions dealt with validating identified usability dimensions for each determinant. A total of eight responses have been received for the questions in this category, while two experts mentioned UAC for this part, as shown in Table 6. From the remaining eight responses, experts agreed with the determinants for dimensions 1, 4, 5, and 9. However, disagreement with the suggestion has been provided for 2, 3, and 8, whereas the remaining responses have been added to the disagreement for APMO. The suggested determinants for dimensions 2 and 8 are context and structure with semantics instead of proposed determinants such as context and structure. Similarly, the proposed determinants for dimension 3 are context and semantics, while adding structure to this dimension was suggested. The suggestions have been added to the usability rules. Based on the APMO cut-off rate, it has been analyzed that consensus has been achieved for the formulation of usability rules, as presented in Table 6. So, no further improvements were required. However, the suggestions of experts are added to the findings for usability rule formulation.

#### **Arrangements of model components**

Based on the APMO cut-off rate, it has been analyzed that consensus has not been achieved for the questions on the arrangement of components, as presented in Table 5. It has been analyzed that four experts did not agree with the arrangement of modules in the proposed model. E2 mentioned that scalability is an important usability issue that should be added to the existing list of usability issues. Similarly, it has been suggested by E5 that “there should be a correctness module added in the data usability issues which are related to the veracity of data”. Therefore, these two usability issues have been added to the subcomponent ‘usability issues.’ An iconic pictorial representation of each usability issue, as recommended by E10, has also been added. The icons for each usability issue have been designed using icons, symbols, and shapes from Microsoft Word. Based on

**Table 6** Number of agreements, disagreements, and UAC for ‘formulation of usability rules’

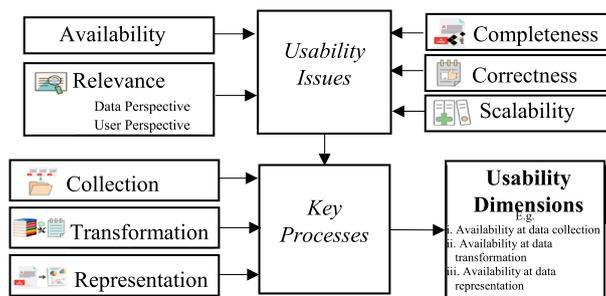
Usability dimension		Usability determinants		Expert opinion		
				Agree	Disagree	UAC
1	Availability	Extraction	Context	8	0	2
2		Transformation	Structure, Context	6	2	2
3		Representation	Context, Semantic	6	2	2
4	Relevance	Extraction	Context	8	0	2
5		Transformation	Structure, Semantic	8	0	2
6		Representation	Context, Semantic	7	1	2
7	Completeness	Extraction	Context	6	2	2
8		Transformation	Structure, Context	7	1	2
9		Representation	Context, Semantic	8	0	2

the experts’ opinions about the subcomponent’s ‘usability issues,’ the subcomponent has been redesigned.

The consensus on the ‘arrangement of key processes,’ was not achieved. Two experts E3 and E8 suggested that data extraction is not an appropriate word to reflect its function, commenting as “extraction may involve extracting useful information from the data for better analysis. Collecting data from different resources is not extraction rather than it’s the collection of data”. Therefore, the data extraction key process has been replaced with the data collection. As mentioned earlier, E10 recommended including pictorial icon representation for each key process. Therefore, considering all three suggestions, the key processes subcomponent of the proposed model has been revised. The consensus for questions regarding the formulation of usability dimensions has been achieved. The formulation of usability dimensions comprised of two subcomponents i.e. usability issues and key processes. However, consensus was not achieved for these two subcomponents. The disagreements on the formulation of these two subcomponents as usability dimensions have been considered for improvement. It has been analyzed that there was a need to improve the input flow in the usability dimensions component. A usability issue and a key process make one usability dimension. There were two parallel inputs from ‘usability issues’ and ‘key process’ in the proposed model. However, it should be in a flow to select one usability issue and then a key process. Considering this expert opinion, the usability dimensions component of the proposed model has been improved, as shown in Fig. 6. Moreover, highlighting some usability dimensions as examples would enhance its understandability, as suggested in expert opinion, has also been considered in model revision.

The arrangement and formulation of usability determinants received disagreements from four experts. The inputs and outputs of the subcomponents were required to improve. Two experts have strongly suggested that the selection of variants of determinants should easily be understood as the determinants might vary among users and tasks. It has also been analyzed that the ability of existing tools to manipulate data is required to implement determinants. Therefore, it has been added to the usability determinants component.

Furthermore, mentioning the variants of usability determinants in the corresponding subcomponent strengthens its understandability, as recommended by E10. Also, E10 mentioned that there should be a pictorial representation for each usability determinant. Considering the experts’ opinions, the usability determinants component has been further improved, as shown in Fig. 7.



**Fig. 6** Revised component ‘usability dimensions’

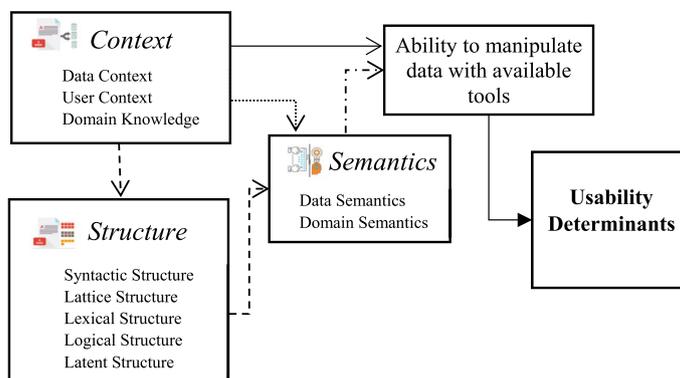


Fig. 7 Revised component 'usability determinants'

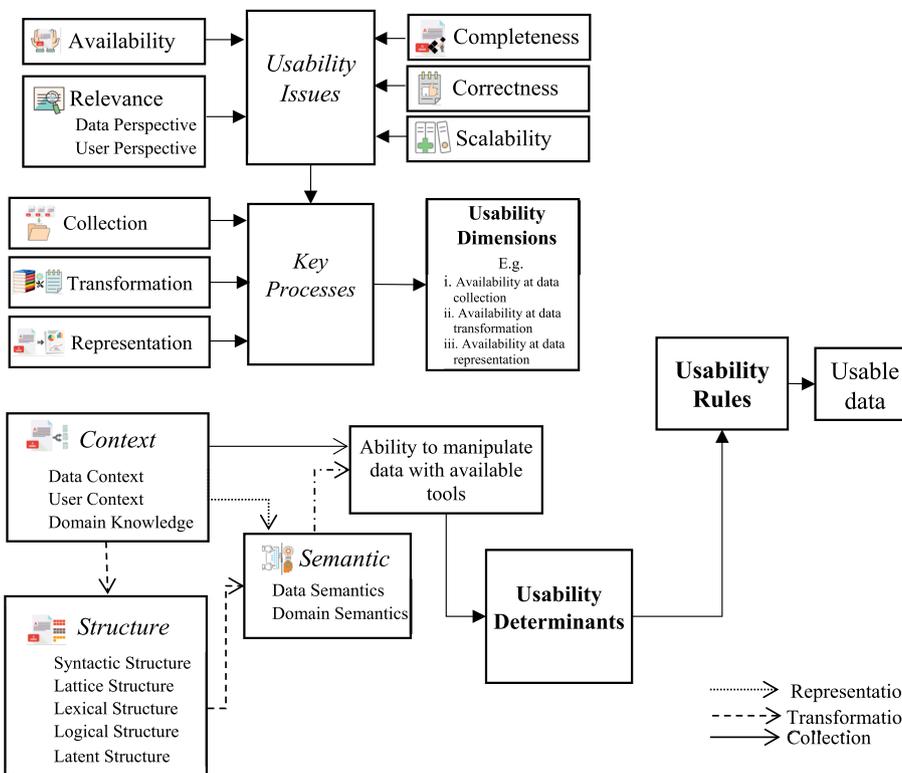


Fig. 8 Revised model after the first round of Delphi

**Revised model**

Based on the analysis results of experts' opinions, the proposed model was revised, as shown in Fig. 8. The second round of Delphi was initiated to validate the revised model.

**Second round of Delphi**

The feedback from nine experts has been collected and analyzed to verify the consensus achievement. Applying APMO as a consensus measurement method for the second

**Table 7** APMO measurement for Delphi second round

Total majority agreements	117
Total majority disagreements	0
Total opinions	117
APMO cut-off rate	100%

**Table 8** Results of the second round

Questions	No	Agreed		Disagreed		UAC		O	C
		#	%	#	%	#	%		
Scalability as a usability issue	M1	9	100	0	0	0	0	9	Yes
Correctness as a usability issue	M2	9	100	0	0	0	0	9	Yes
Relevance sub-dimensions	M3	9	100	0	0	0	0	9	Yes
Pictorial icons of usability issues	M4	9	100	0	0	0	0	9	Yes
Extraction is replaced by a collection	M5	9	100	0	0	0	0	9	Yes
Pictorial icons for key processes	M6	9	100	0	0	0	0	9	Yes
Change in flow in usability dimensions	M7	9	100	0	0	0	0	9	Yes
Mentioning example usability dimensions	M8	9	100	0	0	0	0	9	Yes
Change in flow in usability determinants	M9	9	100	0	0	0	0	9	Yes
Consideration of different perspectives	M10	9	100	0	0	0	0	9	Yes
Pictorial icons for usability determinants	M11	9	100	0	0	0	0	9	Yes
Mention variants in the respective module	M12	9	100	0	0	0	0	9	Yes
Addition of existing tools capability	M13	9	100	0	0	0	0	9	Yes

O Total Opinions, C Consensus achieved?

round as well, the responses of experts have been categorized into agreement, disagreement, and UAC. The number of majority agreements and disagreements has been calculated for each question and divided by total opinions to measure the consensus.

Table 7 presents the APMO cut-off rate for the second round. Based on the APMO cut-off rate, it has been determined whether the revised model has achieved the consensus.

The APMO rate for the second round was 100% to measure the consensus. The APMO rate has shown that if the percentage of 'agreement' for each question is 100%, it would be considered as the consensus has been reached. The APMO rate for consensus measurement has been applied to the questions, and the results have been determined, as presented in Table 8. According to the results, the consensus was achieved for each of the statements of modifications in the model. Therefore, based on the first round results, the revised model has been considered a usability enhancement model.

According to the results, the consensus was achieved for each of the statements of modifications in the model. Therefore, based on the results of the first round, the revised model has been considered a usability enhancement model, as shown in Fig. 8.

**Validated model**

This section presents the results after model validation in relevance to the usability enhancement of the unstructured text data. The usability enhancement model was revised using Delphi. Two more usability issues of unstructured text have been identified. Therefore, five usability issues of unstructured text have been identified: data

availability, data relevance, data completeness, scalability, and data correctness. These usability issues occur in the first three phases of the big data analytics process such as data collection, transformation, and representation. These findings formed usability dimensions with two subcomponents: usability issues and key processes. Further, this research validated the identification of three usability determinants with various variants for unstructured text data. The third component of the usability enhancement model, usability rules formulation, has been derived based on mapping usability dimensions and determinants. The revised model has been considered validated because the consensus was achieved after the second round of Delphi, as shown in Fig. 8.

The usability enhancement model for unstructured text in big data provided a strategy for data analysts and scientists to use unstructured data according to their requirements. The three fundamental components of the model provide the basis for the usability enhancement for big data analytics according to the usage context. However, users can formulate their own usability rules according to their requirements. In this regard, the fundamental need is to identify the usability dimension and appropriate usability determinant(s) and map both to make unstructured data usable.

Unstructured text can also be found in manufacturing companies, factories, and even in class notetaking. The industries where a huge volume of unstructured data is being generated and stored, find it difficult to locate the required data such as newspapers, blogs, healthcare, medical reports, social media data, and many more. With the huge volume of unstructured data, companies find it difficult to make this data usable. In this regard, this research explains how the use of unstructured data is enhanced. The proposed model applies to different domains such as automatic news summarization, customer care systems/ customer complaints response, and extracting relevant data from electronic health records and medical reports.

### **Conclusion and future work**

This paper presents the validation process for the usability enhancement model, developed using SLR. The Delphi technique is adopted to validate the proposed model. The findings of Delphi have shown that the results of Delphi support the results of qualitative synthesis of SLR. The proposed model has been revised according to the feedback of subject matter experts. The consensus has been achieved after the second round of Delphi and the revised model after the second round has been considered validated. The results of Delphi added new themes to the existing themes, identified in the SLR. It has been observed that the feedback of the subject matter experts was aligned with the findings of SLR. Therefore, the usability enhancement model is suitable and applicable to industries where unstructured text is generated in huge volumes. The usability enhancement model provides a strategy for big data analytics to unlock significant values from unstructured text according to the users' requirements.

This existing work will be extended further with case studies to generate an in-depth and extensive understanding of the implementation of the proposed solution. These case studies can be implemented in various domain-specific and general disciplines where data is generated and stored in free text form such as healthcare, social media, e-governments, banking, and alike. The proposed usability enhancement model is limited to unstructured free text, whereas unstructured big data is available in many formats.

Therefore, it is planned to conduct detailed investigations on the usability enhancement of other unstructured data types such as images and videos in the future. The usability enhancement of images and video data is a potential research area as a huge volume of this data is generated daily, especially on social media.

#### Abbreviations

SLR	Systematic literature review
APMO	Average percentage of majority opinions
UAC	Unable to comment
KRNW	Knowledge resource nomination worksheets

#### Author contributions

All authors contributed to this article for its technical content and organization. The research is conducted by KA under the supervision of RA. The write-up of this paper is done by KA in coordination with RA. Improvements in the technical content, language, and organization of the content have been made by RA and KSW.

#### Funding

Not applicable.

#### Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

#### Declarations

##### Ethics approval and consent to participate

Ethical approval has been taken by the UTAR Scientific and Ethical Review Committee before conducting the interviews. The participation was on a completely volunteer basis and all the participants signed the informed consent to participate in this research.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

Received: 29 May 2023 Accepted: 3 October 2023

Published online: 08 November 2023

#### References

1. Adnan K, Akbar R, Khor SW, Ali ABA. Role and challenges of unstructured big data in healthcare. In: Sharma Neha, Chakrabarti Amlan, Balas Valentina Emilia, editors. Role and challenges of unstructured big data in healthcare. Singapore: Springer; 2020.
2. Adnan K, Akbar R. Limitations of information extraction methods and techniques for heterogeneous unstructured big data. *Int J Eng Bus Manag*. 2019. <https://doi.org/10.1177/1847979019890771>.
3. Adnan K, Akbar R. An analytical study of information extraction from unstructured and multidimensional big data. *J Big Data*. 2019. <https://doi.org/10.1186/s40537-019-0254-8>.
4. Adnan K, Akbar R, Wang KS. Information extraction from multifaceted unstructured big data. *Int J Recent Technol Eng*. 2019. <https://doi.org/10.3594/ijrte.B1074.0882S819>.
5. Kim Y, Huh EN. Study on user-customized data service model for improving data service reliability. In: Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication, IMCOM 2017. New York, NY, USA: Association for Computing Machinery, Inc; 2017; <https://doi.org/10.1145/302227.3022254>
6. Adnan K, Akbar R, Wang KS. Development of usability enhancement model for unstructured big data using SLR. *IEEE Access*. 2021;9:87391–409.
7. Adnan K, Akbar R, Wang KS. Towards Improved Data Analytics Through Usability Enhancement of Unstructured Big Data. In: 2021 International Conference on Computer & Information Sciences (ICCOINS). IEEE; 2021. p. 1–6.
8. Nordin N, Deros BM, Wahab DA, Rahman MNA. Validation of lean manufacturing implementation framework using delphi technique. *J Teknol*. 2012;59(2):1–6.
9. Kennedy et al. Verification and Validation of Scientific and Economic Models. In: Agent 2005 Conference Proceedings. 2005. p. 1–15.
10. Inglis A. Approaches to the validation of quality frameworks for e-learning. *Qual Assur Educ*. 2008;16(4):347–62.
11. Cresswell JW. Research design: qualitative, quantitative, and mixed methods approaches. 4th ed. California: Sage Publications Inc; 2014.
12. Sekayi D, Kennedy A. Qualitative Delphi method: a four round process with a worked example. *Qual Rep*. 2017;22(10):2755–63.

13. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm*. 2016;38(3):655–62.
14. Brady SR. Utilizing and adapting the Delphi method for use in qualitative research. *Int J Qual Methods*. 2015;14(5):160940691562138. <https://doi.org/10.1177/1609406915621381>.
15. Aljumaili M, Karim R, Tretten P. Metadata-based data quality assessment. *VINE J Inf Knowl Manag Syst*. 2016;46(2):232–50. <https://doi.org/10.1108/VJKMS-11-2015-0059>.
16. Shanks G, Corbitt B. Understanding data quality: social and cultural aspects. In: 10th Australian Conference on Information Systems. 1999. p. 785–96. <https://pdfs.semanticscholar.org/2959/4536de86f084f9053743d48e5a2b0312b294.pdf>
17. Lee Jang, Gim Kim. A study on data profiling: focusing on attribute value quality index. *Appl Sci*. 2019;9(23):5054.
18. Waggoner J, Carline JD, Durning SJ. Is there a consensus on consensus methodology? Descriptions and recommendations for future consensus research. *Acad Med*. 2016;91(5):663–8.
19. Landeta J. Current validity of the Delphi method in social sciences. *Technol Forecast Soc Change*. 2006;73(5):467–82.
20. Shariff N. Utilizing the Delphi survey approach: a review. *J Nurs Care*. 2015;4(3):246.
21. Flanagan T, Ashmore R, Banks D, MacInnes D. The Delphi method: methodological issues arising from a study examining factors influencing the publication or non-publication of mental health nursing research. *Ment Heal Rev J*. 2016. <https://doi.org/10.1108/MHRJ-07-2015-0020>.
22. Alaloul WS, Liew MS, Zawawi NAW. Delphi technique procedures: a new perspective in construction management research. *Appl Mech Mater*. 2015. <https://doi.org/10.4028/www.scientific.net/AMM.802.661>.
23. Laick S. Using Delphi methodology in information system research. *Int J Manag Cases*. 2014;14(4):261–8.
24. Kim N, Joo SJ. The exploratory study on the method of information system introduction in SMEs using Delphi technique. *J Korea Ind Inf Syst Res*. 2013;18(1):47–58.
25. Ogden SR, Culp WC, Villamaria FJ, Ball TR. Developing a checklist: consensus via a modified Delphi technique. *J Cardiothorac Vasc Anesth*. 2016;30(4):855–8.
26. Flostrand A. Finding the future: crowdsourcing versus the Delphi technique. *Bus Horiz*. 2017;60(2):229–36.
27. Skulmoski GJ, Hartman FT, Krahn J. The Delphi method for graduate research. *J Inf Technol Educ Res*. 2007;6(1):1–21.
28. Fink-Hafner D, Dagen T, Doušak M, Novak M, Hafner-Fink M. Delphi method: strengths and weaknesses. *Metod Zv*. 2019;16:1–9.
29. Chalmers J, Armour M. The Delphi technique. In: Liamputtong Pranee, editor. *Handbook of research methods in health social sciences*. Singapore: Springer; 2019.
30. Mariam N, Woo Nam C. The development of an ADDIE based instructional model for ELT in early childhood education. *Educ Technol Int*. 2019;20(1):25–55.
31. Amidharo SS. Critical success factors for the implementation of a knowledge management system in a knowledge-based engineering firm. Brisbane: Queensland University of Technology; 2014.
32. Alias R. Development and validation of a Model Of Technology-Supported learning for special educational needs learners in Malaysian institutions of higher learning. Selangor: Universiti Teknologi MARA; 2016.
33. Trevelyan EG, Robinson N. Delphi methodology in health research: how to do it? *Eur J Integr Med*. 2015;7(4):423–8.
34. Avela JR. Delphi panels: research design, procedures, advantages, and challenges. *Int J Dr Stud*. 2016;11(1):305–21.
35. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Inf Manag*. 2004;42(1):15–29.
36. McPherson S, Reese C, Wendler MC. Methodology update: Delphi studies. *Nurs Res*. 2018;67(5):404–10.
37. Hosseinvand R, Dosti M, Tabesh S, Razavi SMH. Identifying entrepreneurial challenges with an athletic and recreational approach (case study: Lorestan Province). *Appl Res Sport Manag*. 2020;9(2):75–98.
38. Cooke E, Henderson-Wilson C, Warner E. The feasibility of a pet support program in an Australian university setting. *Heal Promot J Aust*. 2020. <https://doi.org/10.1002/hpja.411>.
39. Keil M, Tiwana A, Bush A. Reconciling user and project manager perceptions of IT project risk: a Delphi study. *Inf Syst J*. 2002;12(2):103–19.
40. Habibi A, Sarafrazi A, Izadyar S. Delphi technique theoretical framework in qualitative research. *Int J Eng Sci*. 2014;3(4):8–13.
41. Muhammad S. METHODS OF DATA COLLECTION. In: *Basic Guidelines for Research: An Introductory Approach for All Disciplines*. BZ Publication, Ed.) Chittagong, Bangladesh: Universidad Curtin. Obtenido de. 2016. p. 201–76.
42. Barrett D, Twycross A. Data collection in qualitative research. *EvidBased Nurs*. 2018. <https://doi.org/10.1136/eb-2018-102939>.
43. Erlingsson C, Brysiewicz P. A hands-on guide to doing content analysis. *African J Emerg Med*. 2017;7(3):93–9.
44. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15(9):1277–88.
45. Assarroudi A, Heshmati Nabavi F, Armat MR, Ebadi A, Vaismoradi M. Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. *J Res Nurs*. 2018;23(1):42–55.
46. Heiko A. Consensus measurement in Delphi studies: review and implications for future quality assurance. *Technol Forecast Soc Change*. 2012;79(8):1525–36.
47. Stewart BT, Gyedu A, Quansah R, Addo WL, Afoko A, Agbenorku P, et al. District-level hospital trauma care audit filters: Delphi technique for defining context-appropriate indicators for quality improvement initiative evaluation in developing countries. *Injury*. 2016;47(1):211–9.
48. Othman MR, Bruce GJ, Hamid SA. The strength of Malaysian maritime cluster: the development of maritime policy. *Ocean Coast Manag*. 2011;54(8):557–68.
49. Zhang X, Roe M. The Delphi research process. In: Zhang Xufan, Roe Michael, editors. *Maritime container port security*. Cham: Springer; 2019.
50. Zhang X. The United States container security initiative and european union container seaport competition. UK: University of Plymouth; 2018.
51. Arnold, Deborah; Sangrà A. E-leadership literacies for technology-enhanced learning in higher education. In: *Proceedings of the European Distance and E-Learning Network 2018 Annual Conference*. 2018. p. 1–9.

52. Allin B, Ross A, Marven S, Hall NJ, Knight M. Development of a core outcome set for use in determining the overall success of gastroschisis treatment. *Trials*. 2016;17(1):1–7.
53. Botma Y. Consensus on interprofessional facilitator capabilities. *J Interprof Care*. 2019;33(3):277–9.
54. Zhang Y, Wildemuth BM. Qualitative Analysis of Content. *Applications of Social Research Methods to Questions in Information and Library Science*. Retrieved; 2009. p. 421. [http://ils.unc.edu/~yanz/Content\\_analysis.pdf](http://ils.unc.edu/~yanz/Content_analysis.pdf)
55. Silver C, Lewins A. *Using software in qualitative research: a step-by-step guide*. 2nd ed. Thousand Oaks: Sage Publications; 2014. p. 211.
56. Wang S, Wang H. Big data for small and medium-sized enterprises (SME): a knowledge management model. *J Knowl Manag*. 2020;24(4):881–97.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---