Automated Knowledge Discovery or Integration: A Systematic Review of Data Mining in Knowledge Management

By

Ephenia Peu

Thesis presented in fulfilment of the requirements for the degree of Master of Philosophy (Information and Knowledge Management) in the Faculty of Arts and Social Sciences at Stellenbosch University

Supervisor: Dr C. H. Maasdorp

Department of Information Science

December 2021

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2021

Summary

Data mining and knowledge management activities have been crucial for making sense of the vast amounts of data, information, and knowledge created in organisations. Data mining comprises the collection, categorisation, and analysis of data to find useful patterns and establishing solutions based on those patterns. Integrating data mining into knowledge management has had little exploration and attention. The thesis aims at this gap and investigates the role of data mining in the knowledge management literature in both quantitative and qualitative studies between 2000 to 2017.

A systematic literature review identified and analysed published articles utilising data mining in knowledge management to reveal the trends in the field. The initial search was conducted on four interdisciplinary databases and an article selection process that involved inclusion and exclusion criteria and a quality assessment using a checklist yielded 54 articles for analysis.

Six themes were identified in a thematic analysis where the articles were coded using Atlas.ti software: 1) technical advances improve access to and transformation of knowledge, 2) the knowledge base as the basis for improved product and service development, 3) the use of big data analytics for customer relationship management, 4) the role of data and information assets for decision support, 5) combining automation and human expertise to improve efficiency, and 6) the effectiveness of data mining applications as guided by the specificity of the knowledge management task.

Finally, the themes resulting from the coding are mapped on the stages of the knowledge management process. The discovery and capture stages concern data mining techniques for knowledge discovery; the process stage uses the knowledge base and decision support to access knowledge for action; and the share and benefits stage is the domain of learning and capacity development.

Opsomming

Data-ontginning en kennisbestuursaktiwiteite is deurslaggewend om sin te maak uit die groot hoeveelhede data, inligting en kennis wat in organisasies geskep word. Data-ontginning behels die versameling, kategorisering en ontleding van data om bruikbare patrone te vind en oplossings op grond van daardie patrone te vestig. Tot op hede, is relatief min aandag gegee aan die integrasie van data-ontginning en kennisbestuur. Die tesis fokus op hierdie gaping en ondersoek die rol van data-ontginning in die kennisbestuursliteratuur in beide kwantitatiewe en kwalitatiewe studies tussen 2000 tot 2017.

'n Sistematiese literatuuroorsig het gepubliseerde artikels wat gebruik maak van data-ontginning in kennisbestuur geïdentifiseer en ontleed om die tendense in die veld te identifiseer. Die aanvanklike soektog is op vier interdissiplinêre databasisse gedoen en 'n seleksieproses met insluiting- en uitsluitingskriteria en 'n kwaliteitbeoordeling het 54 artikels vir ontleding opgelewer.

Ses temas is in 'n tematiese analise, waar die artikels met behulp van Atlas.ti-sagteware gekodeer is, geïdentifiseer: 1) tegniese vooruitgang wat verbeterde toegang tot en transformasie van kennis moontlik maak, 2) die kennisbasis as oorsprong vir verbeterde produk- en diensontwikkeling, 3) die gebruik van data-analise vir kliënteverhoudingsbestuur, 4) die rol van data en inligtingsbates vir besluitsteun, 5) die kombinasie van outomatisering en menslike kundigheid vir doeltreffendheid, en 6) die bepaling van data-ontginningstoepassings deur die spesifisiteit van die kennisbestuurstaak.

Laastens word die temas as resultaat van die kodering op die stadiums van die kennisbestuursproses toegepas. Die ontdekkings- en vasvangstadiums handel oor data-ontginningstegnieke vir kennisontdekking; die prosesstadium gebruik die kennisbasis en besluitsteun om aksie-kennis te verkry; en die deel- en voordelestadium is die domein van leer en kapasiteitsontwikkeling.

Table of Contents

List of Tables

List of Figures

Table of Abbreviations

ACM	Association for Computing Machinery
AI	Artificial intelligence
ANN	Artificial Neural Network
AQI	Adaptive Query Interface
AR	Association rules
BI	Business intelligence
CART	Classification and Regression Tree
CDSS	Clinical decision support system
CI	Competitive Intelligence
CRISP	Cross Industry Standard Process
CRM	Customer Relationship Management
DMKD	Data Mining and Knowledge Discovery
DBMS	Database Management Systems
DSS	Decision Support System
ERP	Enterprise Resource Planning
FARM	Fuzzy Association Rule Mining
HEI	Higher Education Institutions
ICT	Information Communication Technology
IoT	Internet of Things
IT	Information Technology
KAS	Knowledge Acquisition System
KB	Knowledge bases
KBE	Knowledge-based engineering
KDD	Knowledge Discovery in Databases
KMP	Knowledge management process
LVQ	Learning Vector Quantization
MBMS	Model Base Management System
NLP	Natural language processing
NN	Neural network
OLAP	Online Analytical Processing
OS	Organisational Structure

PKM	Personal Knowledge Management
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RFID	Radio Frequency Identification
SEMMA	Sample, Explore, Modify, Model and Assess
SKMS	Service Knowledge Management System
SR	Systematic Review
SVM	Support Vector Machine
TAPUPAS	Transparency, Accuracy, Purposivity, Utility, Propiety, Accessibility,Specificity
XML	Xtensible Markup Language

Chapter One

1.1 Introduction

Knowledge management (KM) is a process that has become synonymous with achieving organisational goals. It is a process that offers a blueprint of what and how the management of knowledge can be achieved. KM has seen development over time, which has been in line with the growing demands and needs of people and organisations. It is a discipline that cuts across different industries and disciplines, such as engineering, health, education, finance and marketing. Due to the nature of the diverse industries, the approach and application of KM are varied according to their understanding, the types of business being run and what its needs are. To embrace KM, organisations must first understand what data and knowledge capabilities are available and accessible within their reach. They must know how to use these assets for decision-making, hoping to gain some competitive advantage. KM can define what the knowledge means, but to understand what the data means, applying technologies, such as data mining (DM) and analytics are required for interpretation.

In understanding what KM and DM entail, it is important to highlight the differences between knowledge, information and data. The three concepts are linked but are understood differently in different environments and that "data isn't information and that information isn't knowledge" (Chen et al., 2008). In addition, each concept must include different elements to be classified as is. According to Liew (2007) and Liew (2013), symbols and signal readings make up data, which is to be recorded, captured and stored, with its purpose as recording situations or activities, to understand what is happening in the actual environment. Information comes from both current and historical sources with relevant meaning for decision-making, and knowledge is the ability to recognise, the capacity to action, and understanding what is contained within the mind to increase value. These definitions and understanding of the concepts are further supported by authors, such as Alavi and Leidner (2001a), who put forward that: "Knowledge is the result of cognitive processing triggered by the inflow of new stimuli, that information is converted to knowledge once it is processed in the minds of individuals and knowledge becomes information once it is articulated and presented in the form of text, graphics, words, or other symbolic forms. A significant implication of this view of knowledge is that for individuals to arrive at the same understanding of data or information, they must share a certain knowledge base." Data, information and knowledge are defined by Chen et al., (2008) as data is representations of models computerised and attributes of actual or simulated units; the results from a computational process are represented as data then interpreted as information, such as assigning meanings to the data through statistical analysis; and knowledge as the computer-simulated cognitive process results such as reasoning, perception, learning and association. Clarity on the three concepts allows the organisation to understand the activities and processes with its implementation.

Nemati, Steiger, Iyer and Herschel (2002) highlight that the creation of KM happens through the synergistic relationship between tacit and explicit knowledge. The creation is apparent through a four-step process of socialisation, articulation, understanding and integration or internalisation in the organisations. The motivation for this belief and practice is rooted in that explicit knowledge is a product of tacit knowledge. The production of explicit knowledge takes place through different platforms designed to capture the tacit knowledge and through different activities, leading to the developing of clear knowledge. That process can leverage the collection of explicit knowledge to create new knowledge. The activities become part of the knowledge creation cycle applied as part of a business strategy.

According to Demarest (1997), the consensus is that KM requires aspects, such as people, culture, processes, operations and technical considerations. Gold, Malhotra & Segars (2001) posit that the KM process pays attention to knowledge flows and how they move within the creation, sharing and distributing phases. The field of KM is relatively developing (Han, Kamber & Pei, 2012, p. 2) and still confuses organisations in terms of how it is viewed with its overall usefulness in an organisation. Additionally, the research is lacking regarding how KM operates in unpredictable developments and its role to prepare for the future.

When considering KM, an organisation must maximise the benefits of the resources available to the business and use its capabilities for successful operations. Organisations aim to achieve a competitive advantage using their resources and matching them to their capabilities (Grant, 1991). The resources and capabilities can be viewed using two theoretical frameworks, the Resource-Based Theory and the Dynamic Capabilities Theory. The Resource-based Theory has the perspective that the combination of resources and capabilities creates a strong base for long-term strategy. According to the theory, every organisation has in its possession a set of unique resources and capabilities that could, based on their allocation and utilisation, positively contribute to the performance of the organisation (Hitt Ireland & Hoskisson, 2017, p. 16). The theory focuses on the tangible and intangible resources and capabilities controlled and used by organisations to build core competencies that make organisations rare, valuable and difficult to imitate (Pires & Trez, 2018) that can set the organisation apart. The Dynamic Capabilities Theory used previously made decisions from which processes depend on, and organisations go through the integration and reconfiguration of their resources and capability's portfolio to respond to current positions to build long-term competitive advantage (Teece, Pisano & Shuen, 1997). According to Teece (2007), there are three types of dynamic capabilities: sensing (shaping) opportunities and threats, seizing opportunities, and managing threats and reconfiguration. Organisations can use combined resources and capabilities to address challenges, find solutions and improve the business.

One capability that is considered important in discovering knowledge is the availability of technologies in an organisation. One technology is the application of DM. DM benefits result from operationalising its results via a business strategy to achieve a specific objective (Hirji, 2001). Organisations need an understanding of whether the practice of KM is purely logistical or if the process has or requires additional activities such as DM (Devedzic, 2001), which at times other research refers to it as Knowledge Discovery in Databases (KDD).

DM and KDD are often used interchangeably, but their activities are different, but DM is a step in the KDD steps, where KDD extracts information that is valuable for the available data, and DM performs modelling and analysis using techniques (Alasadi & Bhaya, 2017; Selamat, Prakoonwit & Khan, 2020). The main focus in this study is the use of DM, since knowledge discovery is about a sequence of a process made up of steps to discover new knowledge (Han et al., 2012). DM is a process of discovering knowledge and compelling using analytical techniques patterns from sizeable amounts of data. DM can gather patterns from different sized data, such as information repositories or the web, as long as the data is significant enough in size for the answer needed (Ranellucci, Poitras, Bouchet, Lajoie & Hall, 2016). According to Singh and Kumar (2017), there are several enablers of DM, such as the availability of Information

Technology, the World Wide Web (the Internet), an effective customer relationship management, database processing developments, data warehousing, machine learning, KM, developments in computer hardware and software, strategic business initiatives, and Knowledge Discovery in Database. These elements are necessary to activate DMs actions and make sure that it efficiently works for the organisation.

KM system design has evolved into an innovative design (Quan, Xiao, Ji & Zhang, 2021). The change is due to the continuously changing environment and the growing demands of dynamic services. They have become more aligned with business strategies and Information Technology (IT), which can maximise business knowledge. IT can facilitate each knowledge activity of capture and creation, sharing and dissemination, and acquisition and application (Crawford, Leonard & Jones, 2011). The IT capability is the link between many processes within organisations and the reason they function relatively well (Turulja & Bajgoric, 2018). These changes highlight the considerable integration that takes place between different processes due to the evolution of business needs, and also recognising that the same formula will not serve everyone.

The resources and capabilities elements highlight the importance of finding tools and methods that can form part of the KM process (KMP) on a much greater scale, such as playing a role in the organisational strategies, as part of the business system, where knowledge creation, storage and sharing become part of the life cycle. The decision-making and solution finding often come from a functioning business system (Lundvall, 1999).

There are existing gaps in the research about how and when discovering knowledge occurs and during what phase of a project. Strategically planning to use the KM process can support finding additional methods of discovering knowledge by exploring the data and knowledge that are continuously collected from the daily organisational activities. Especially if the organisation does not have an existing guiding framework on the use of discovered knowledge. Industries have come a long way in how they use the knowledge discovered within their operational processes, which is then transformed into knowledge and then integrated into daily business activities. Some industries have implemented system integration for knowledge discovery, such as DM and KM, business intelligence and data warehousing, customer relationship management, and knowledge engineering. By aligning and integrating the findings from processes, such as DM into the organisation's operations to achieve set objectives, it creates potential DM benefits from which new opportunities can be unearthed.

With the vast amounts of data generated from day to day organisational activities, pragmatic methods for effective handling are required. Such methods include applying different technologies that can answer their business intelligence questions and respond to their needs. Generated and collected data comes in different variations, and because of that diversity, organisations have different platforms to collect and keep, such as business databases. Organisational data storage contains data, like consumer purchasing patterns, sales, transactions, accounting, human resources, which can be mined to discover more knowledge.

Hema and Malik (2010) point out that DM techniques can be applied and implemented to existing hardware and software already in place. The process enhances the existing information resources for redevelopment so that integration with new products and systems takes place and brings new and different activities to the fore. Resource or processes integration involves connecting or incorporating different units, technologies, activities, or strategies to provide efficiency (Spencer-Oatey & Dauber, 2019). As much as integration can create opportunities for an organisation, it can also be a challenging decision because linking different existing systems, with the lack of structured data that is coherent and can be fused, may not necessarily work for the organisational units (Berente, Vandenbosch & Aubert, 2009). Thus it is important to strategically plan an integration project with the necessary information for alignment.

DM and KDD allow for the identification of patterns. The aim was to understand to what degree DM plays a role in the KM process and what the view was in this regard. DM is an interdisciplinary process, and it was relevant for the study since the literature was inter and multidisciplinary in nature. The reason for selecting DM was rooted in its use and how that use can indicate whether its activities are tied to the support for KM processes. Another aspect of the study was to investigate the role organisational processes play in accommodating such practices in their environment. There is often some grey area in the level of understanding of such systems and how implementation occurs in everyday business activities. It can be challenging for businesses to recognise and understand what KM can do and how far-reaching it can go for the organisation's overall performance using DM.

A systematic review was employed to obtain the patterns and trends of DM in KM from publications from 2000 - 2017 based on the search strategy developed. It was especially suitable for this study due to the expansive research publications that have been produced under the two concepts. Systematic Reviews are a means of condensing and putting forward overviews of current and historical knowledge extracted from an extensive body of literature (Aromataris & Pearson, 2014). The aim of systematic reviews goes beyond just summarising everything there is about a particular issue. The aim is to answer a specific question or test a specific hypothesis (Mulrow, 1994; Petticrew & Roberts, 2006, p. 10). Systematic review methodologies are common in the medical and health sciences, where decisions about patient health and public health policies can be developed using the method (Pluye & Hong, 2014). There has been a rise in using systematic reviews in the social sciences, where new findings can be identified concerning a social policy, unlike if a more traditional literature review was used (Kocher & Riegelman, 2018; Victor, 2008). With a systematic review, common themes and aspects that needed clarification were identified, and the research gaps on the topic. Additionally, the systematic review provided unbiased factors associated with DM and KM implementation (Mallett, Hagen-Zanker, Slater & Duvendack, 2012).

1.2 Problem statement

Research articles have been published in KM and DM, respectively, which have emanated from the understanding and experiences of different organisational data and knowledge practices. However, the volatile business landscape has offered an opportunity for greater understanding and implementation of KM programmes due to the development in DM. Publications have not reflected the relationship between the two processes, especially highlighting whether DM has impacted KM and the impact thereof.

1.3 Purpose of the study

The purpose of the study is to understand the process of discovering knowledge by identifying the relationships and contributions of DM to the field of KM and revealing the different types of themes found over the timeframe covered by the study. This occurs through analysing the contribution to the KM process and the analysis of research articles that focus on the relationship between DM and KM. The study aims to review publications that have investigated, have applied the processes, identified the relationship between DM and KM, the role of DM in KM, and establish the significance of the contribution.

The research objectives are:

- To apply a thematic analysis on multidisciplinary literature published between 2000 and 2017.
- To highlight all the research patterns and trends from the reviewed articles regarding the relationship DM has with KM and the relationship between DM and KM.
- To identify the factors that exist in making DM and KM work together.

The study should be able to answer the following questions:

- What themes are prevalent in the study's unit of analysis?
- From the patterns discovered and trends highlighted, what do they illustrate about the nature of the relationship link between DM and KM?
- Is the process of discovering knowledge an automated activity, or is it an action that requires integration into the process of discovering new knowledge?

1.4 Significance of the study

There is a wide range of research on DM and KM as significant processes. From the initial review of the literature, it was observed that there are studies done to understand the relationship between the two processes and how DM was related to KM over the years. Using KM takes place in many organisations' practices and activities. However, the efficiencies of the activities and aspects of DM as part of the company's KM process and framework have not been thoroughly investigated. By identifying the key research gap, this paper aims to carry out an in-depth review of the relationship between DM and KM. The study investigated and identified the patterns in applying DM in KM, by looking at two keywords, integration and automation. The keywords are defined: (a) Integration is generally defined as combining the different parts or aspects to make one whole. The study looked at the diverse elements of DM and how they integrate with the various aspects of KM to support the study's stance on the relationship. The other definition: (b) Automated is defined and understood to be converted to operate automatically as defined by the Online Oxford English Dictionary as “the action or process of introducing automatic equipment or devices into a manufacturing or other means or facility.”

The study provides a critical look at DM in KM and the relationship between DM and KM. The timeframe of 2000 - 2017 was relevant because it was over a period when KM was on the spectrum of popularity. The study highlights the different approaches, challenges, successes and outcomes that build upon the trends currently taking place in the evolving space of DM and KM. The study's outcome can further support future decision-makers in anticipating and developing criteria required to implement projects that involve DM and KM.

1.5 Rationale for study approach

The literature that informed the systematic review includes the time frame of the year 2000 to 2017 and the diverse industries due to the nature of the topic. During this period, the publication rates for KM and DM were very popular and relevant because many organisations were adopting the processes to enhance their positioning in the competitive environment, and researchers published more articles on the topics. There was an increase in academic publications about KM between 2000 and 2008 (Koç Kurt & Akbıyık, 2019). The study's time frame considered the developments that were inspired by publications on the processes and to understand implementation activities from the theory. This position is supported by Ribiere and Calabrese (2016), who posit that there is still an increase of publications related to KM, even though there has been a steady decline in interest from 2005 to 2015 by using Google trends, where the usage trend averaged 42% from 1996 to 2010. According to Rigby and Bilodeau (2007), in Bain’s global 2007 management tools and trends survey, ten tools were highlighted as prominent in terms of usage. These ten tools included: 1. Strategic Planning; 2. Customer Relationship Management; 3. Customer Segmentation; 4. Benchmarking; 5. (tie) Mission and Vision Statements, Core Competencies; 7. Outsourcing; and 8. (tie) Business Process Reengineering, Scenario and Contingency Planning, Knowledge Management.

DM has progressed from 1990 to current times. It is widespread across industries, which is used where it can mine data, such as consumer behaviour, credit card transactions, stock market movements, national security and clinical trials (Kumar & Bhardwaj, 2011; Sharma, 2014). In a survey conducted by Rigby and Bilodeau (2017), advanced analytics, which is associated with DM, was one tool that appeared in the trending usage, which had a steady climb since 2000. The surveys did not highlight if organisations use DM in KM to discover knowledge and if benefits have been realised, even though there has been wide adoption of DM and KM. This observation is supported by Parlby (2000), who reports that organisations have accepted that KM is an accepted part of the business agenda and that organisations have been better off since implementation, but the organisation could not articulate the benefits of KM. Therefore, since the uptake informs the study's timeframe of KM since 2000, key aspects that could be identified would be the different factors that have influenced the organisation's adoption and use of DM in KM, the overall and composite outcome of the units of analysis, the perspectives that have developed, and what these perspectives say about the overall application in discovering knowledge. Combined with the diverse industries and organisations, the study represents different outcomes in terms of the general view and practice in the area of KM and DM. The application of a systematic review method was informed by its ability to provide an overview of the topic even when there are large quantities of information from evaluating publications. According to Petticrew and Roberts (2006, p. 45), systematic reviews cannot address questions about effectiveness. Still, they can answer questions about associations between characteristics of the population, risk factors, and explore associations between risk factors or predictors and outcomes. A systematic review is appropriate for the study because the method can provide answers to the study's objectives.

1.6 Related Studies

The systematic review investigated two previous related research publications: (1) Silwattananusarn and Tuamsuk (2012); and (2) Dastyar, Kazemnejad, Sereshgi and Jabalameli (2017).

The study by Silwattananusarn and Tuamsuk (2012) investigated the development of KM in organisations through DM techniques by reviewing ten research articles from one database. They argued that DM plays a pivotal role in KM. The focus of the research was based on publication rates and applying DM on developing organisational knowledge by looking at collected data from questionnaire-based surveys. In the Silwattananusarn and Tuamsuk (2012) study, the review looked at studies between 2007 and 2012. It pursued to highlight the application of DM techniques for KM to state the importance of knowledge discovery. Where the findings were reported and categorised according to the DM tasks. Their synthesis focused on (a) the knowledge resource- which is the type of industry the knowledge is from; (b) knowledge types- the collected and stored knowledge that forms part of the operations; (c) DM tasks- the techniques you find; and (d) DM techniques applied to KM in the study- the algorithms used. The study's limitations include a heavy focus on DM more than the equal representation of DM and KM, and using only ten articles for the review did not provide a wide view on the topic.

The study by Dastyar et al. (2017) looked at the investigation of applying DM techniques to develop KM in organisations by looking at the same articles as those reviewed by Silwattananusarn and Tuamsuk (2012). The authors attempt to provide organisations with the steps that lead to automation of their activities by implementing organisational information and transforming them to require knowledge. The motivation was establishing DM as a tool that supports organisations to set goals for KM. The article falls short in providing the steps, and its limitations are similar to those of Silwattananusarn and Tuamsuk (2012) because the same articles are reviewed.

This current systematic review differs from the previous reviews in the rigour of the approach employed. A wider investigation into the topic that included consulting four databases to identify a large pool of publications because of the timeframe informing the study and focused on both DM and KM equally, was conducted and completed. The thematic synthesis discovered different themes, which could establish commonalities in the cases where DM is applied to KM, whether empirically or conceptually. This study provides a more detailed overview of the topic, especially for adoption.

1.7 Organisation of the thesis

Chapter 1 introduces the topic and the intention of the study. It then presents the problem statement, the purpose statement, the goal of the overall study, the rationale of the study, and related studies.

Chapter 2 presents the literature review, which presents the background literature on DM and KM, respectively, as processes that play significant roles to give context to the study. It reflects on the application and challenges of the processes.

Chapter 3 presents the research methodology, with the research objectives stated. The study conducts a systematic review methodology by following the systematic review protocol throughout the different phases of the methodology, using a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) model. Details are provided for the rationale behind the search strategy and defining the inclusion and exclusion criteria used. The study highlights the process of reaching and identifying the unit of analysis and the methods used to analyse and synthesise the data. The chapter concludes with the application of thematic analysis.

Chapter 4 provides the results from the thematic analysis applied to the unit of analysis and the synthesis thereof. The section provides conclusions, and the chapter concludes with a summary of the results.

Chapter 5 discusses the analysis of the results and the stated limitations. The section recommends future studies, and the chapter conclusion is presented.

Chapter Two

2.1 Background literature

This study draws from existing literature across diverse disciplines and industries, where DM and KM processes are investigated or implemented. Key elements have been identified based on the perspectives highlighted in the literature. The main element acknowledges that KM is a critical part of many organisations. Knowledge management should be tailored according to the needs of the business, which can be established through the utilisation of various tools and techniques. These techniques may include the varied tasks within DM, depending on the solutions offered by the relevant products and services. This chapter covers literature related to KM and DM. It highlights the definitions of the concepts, the application and the challenges of use.

2.2 Knowledge Management

KM is considered a catalyst for achieving competitiveness and superior business performances within a firm's management (Turulja & Bajgoric, 2018). KM has evolved over the years, and its application is still viewed as a young discipline with room for progress (Koç et al., 2019). There is no singular definition for the process, as it means different things to many firms, speciality industries and scholars. According to Hussinki Ritala, Vanhala and Kianto (2017), KM is made up of processes and activities that assist the organisation in generating, acquiring, discovering, and organising knowledge, then applying and distributing it among the individuals in the organisation, and transferring the collected information and experiences and using them in management activities, such as strategic planning, decision-making, and job procedures. Grey (1996) and Bahra (2001) assert that KM is "a collaborative and integrated approach to the creation, capture, organisation, access, and use of an enterprise's intellectual assets" - and that it is a process that systematically locates, chooses, arranges, extracts and presents information to employees so it improves their understanding in a specific area of interest, and through the same actions, firms can understand their own experience through gaining insight.

According to Alavi and Leidner (2001a), there are several perspectives that knowledge may be viewed from (1) a state of mind - which enables individuals to broaden their individual knowledge and apply it in support of the organisation, (2) an object - as something that can be accumulated and alter, (3) a process - using available expertise, (4) a condition of having access to information - in an organisation, knowledge must be organised to facilitate access to and retrieval of content, or (5) a capability - knowledge can influence future decision-making. In KM, there are knowledge types categorised as either being tacit or explicit. Tacit knowledge resides in the minds of people, while explicit knowledge represents knowledge in a format that can be easily expressed, such as in words and documents (Rolf, 2004; Smith, 2001). According to Bharara, Sabitha and Bansal (2017), the extraction of implicit and explicit knowledge by experts defines KM. The extraction takes place within the organisation and from the processes that play a role in the generation, storing and reusing the knowledge. The governance and optimisation of an organisation's intellectual assets of both tacit and explicit knowledge are at the heart of the definition of KM (Selamat et al., 2020). There is a challenge in equally taking advantage of the knowledge types because there is more emphasis on explicit knowledge when there could be more focus on the usage of tacit knowledge, since it has a greater advantage in encouraging organisational innovation (Davenport, De Long & Beers, 1998; Selamat et al., 2020).

The day to day performance of businesses depends on how accessible, relevant knowledge is to their employees and whether they can interpret it in a manner in which they can utilise it (Smith, 2001). Models and techniques have been developed to improve the sharing, use and dissemination of such knowledge. These models and techniques depend on the methods used to structure the information, and the two aspects together result in the existence of knowledge.

The action of KM is still not widely accepted among decision-makers, and thus they may find it challenging to apply (Parlby, 2000). These challenges are amplified since KM requires that businesses have infrastructure and policies to allow for cross-functional knowledge exchange. Different authors see KM practices from varied perspectives. Authors, such as Xie, Zou & Qi, (2018), have the view that dimensions that are more comprehensive comprise of acquisition, assimilation, transformation and exploitation. While others, such as Al-Emran et al. (2018), view knowledge creation, transfer and application as the KM processes that plays a key role.

According to Heisig (2009) and Bolisani et al. (2015), systematic knowledge handling should be a deliberate and planned activity in an organisation and not just "handling knowledge" because the latter presents nothing new. The intentionality of KM being a strategy comes into play when organisations and top management need to plan long-term and decide which resources are required for that type of planning activity. Not all studies support the strategic planning of KM but argue that not all organisations are structured the same way, and others uncover their need for KM at a later stage, while others manage their knowledge differently (Bolisani et al. (2016); Desouza & Awazu, (2006); Wee & Chua, (2013)). One of the fundamental aspects of KM is to ensure that available knowledge is applied to benefit an organisation, which should form part of the strategic planning. Integration of KM processes with the corporation’s system can fast track easy access to knowledge by the different teams (Bach and Alessa, 2014). One of the key goals of KM is making use of the available knowledge to benefit an organisation with a focus on fundamental aspects, such as knowledge integration to find solutions to organisational problems (Ode & Ayavoo, 2020).

Abualoush et al. (2018) assert that there has been rapid growth and importance placed on the role knowledge plays as a core unit of wealth based on the expertise, skills, creativity, and the ability to generate new knowledge by individuals, due to the increased importance of technology and information systems. Tangible resources available in the organisation enable the organisation's performance, but the performance relies not only on the tangible resources but also on the intangible resources, such as the effectiveness of management efforts of knowledge, keeping track of developing technology and the adoption of advanced data collection and analysis systems (Obeidat, Tarhini, Masa'deh and Aqqad, 2017). The importance of knowledge is emphasised by Abusweilem and Abualoush (2019). In addition to human resources and capital, they assert that great value is seen and placed in knowledge as assets in modern organisations. It has consequently, become a factor of production that has proven to be important.

Organisations should have a defined scope for KM practices to operationalise it. According to Oktari, Munadi, Idroes and Sofyan (2020), definitions of KM practices have been grouped based on being i) People-oriented KM, which is associated with innovation, ii) Process-oriented KM supports the advancement of the organisation's performance, iii) Technological-oriented KM is a primary driver of knowledge acquisition, knowledge creation, and knowledge sharing, and iv) Goal-oriented KM has the intention of enhancing the performance of both the organisation and individuals. A system, such as KM consists of a complicated mix of organisational knowledge and culture, people, structures, and technology infrastructures. KM became technology driven because of the availability of information technologies that can be applied to knowledge to make it more effective, such as expert systems, business intelligence, knowledge mapping collaboration, distributed learning, opportunity generation, knowledge discovery and security (Gold et al., 2001). Melville, Kraemer and Gurbaxani (2004) emphasise that organisations must adapt to the undeniable development of new technology being experienced in all environments because the connections created between and within organisations are changing the way they gain and buy factors of production, turning them into products and services and then distributing them to its customers.

There is a heavy emphasis on employees understanding the nature of knowledge and its uniqueness, but to understand these aspects requires that people identify potentially relevant knowledge and possibly outside their regular scope. Faniel and Majchrzakb (2007) argue that this can create difficulties because these individuals need to develop new expertise in evaluating the reliability of the knowledge, whether it is suited to solve problems, or whether it fits with solving the current problems.

According to Karami, Alvani, Zare, and Kheirandish (2015), knowledge is the only reliable source to achieve that goal in creating a competitive advantage and sustainability within the organisation. They argue further that it is a "strategic, irreplaceable and worthy value creator for stakeholders that boosts the production of innovative goods and services". The hyperactive environment shows that organisation's stand out because of their innovation and offering, whether products or services. The ability to effectively and efficiently use the organisation's knowledge both in processes and for distributing to employees can contribute to the organisation's differentiation. KM can counteract competitive dilemmas experienced by organisations by elevating the ability of the organisation’s innovation to lead a performance rooted in sustainability (Selamat et al., 2020).

2.3 Application of Knowledge Management

KM applications can take place in various sectors and different formats. KM can help the organisations drive their strategy, resolve issues and challenges quickly, encourage and align best practices and highlight and promote the knowledge embeddedness in the operations and services provided. Knowledge management implementation is not easy, and recent studies show that some KM projects still fail and require reflection, much needed research, and providing for the required infrastructure and critical success factors for knowledge management implementation (Coakes, Amar, and Luisa Granados, 2010; Karami, et al., 2015). Stakeholder expectations naturally become high due to what is thought the system will provide. According to Mamcenko and Beleviciute (2007), both DM and KM develop and advance at the convergence of information technologies (databases, knowledge-based systems, and machine learning), data analysis and statistics, and the management and business sciences. The right technologies and platforms can offer the potential for successful KM, and DM approaches.

Organisations can take advantage of tools such as the intranet, the availability of metadata used to find and organise the existing information and knowledge, using customer feedback or using the collection of the different experiences from organisational staff who are not from client intensive organisations, to have a record of what they have done on their projects, whether it be the decisions, activities and challenges (Alavi & Leidner, 2001b). All this knowledge can help firms make more informed decisions and increase their tasks and everyday activities (Xie et al., 2018).

For a successful KM application, the business needs to understand the technologies available for use. The practice of KM can fall under organisational learning, even if differentiation management is in practice. As long as there is constant learning, knowledge would be part of the day to day operations and the overall organisational KM process. According to Liao (2003), applying KM technologies develop toward an expert orientation, and this issue is a problem-oriented domain. KM can be applied through an expert system and, in some cases, with the use of a knowledge base. An expert system is a programme designed based on how humans make their decisions, judgement, and behaviour (Bharara et al., 2017). It is a computer programme that integrates artificial intelligence (AI) technologies that make it possible and offer solutions based on actual real-life scenarios. The system is designed based on the expertise the particular organisation specialises in. There are other applications that the expert system can implement, such as visualisation, human resource management, project management, information retrieval and knowledge engineering (Liao, 2003). KM systems have improved the productivity and efficacy of organisations, mainly through the contribution to developing Decision Support

Systems (DSS). Integration of KM processes with the corporation's system can fast track the access of knowledge easily by the different teams (Bach and Alessa, 2014). Neaga and Liu (2014) highlight that the advancement of new technologies takes place all the time, and KM frameworks have been used to integrate information mining, data warehouses, Customer Relationship Management (CRM) applications and Enterprise Resource Planning (ERP).

According to Abualoush, Masa'deh, Bataineh, and Alrowwad (2018), KM infrastructure factors include Organisational Culture (OC), IT Infrastructure (IT), and Organisational Structure (OS). These form the basis for reinforcing knowledge management processes within organisations. IT tools form part of the technological infrastructures, which includes hardware, software and protocols, and can further possibly represent electronic versions of knowledge in the organisation and can enable the simplification of knowledge exchange. IT plays a role in the effectiveness of KM because it can facilitate structured and cross-organisational collaborations and enable computing power, such as technologies like DM and AI (Mishra, Kishore & Shivani, 2018).

Instances of simplifying knowledge exchange require a resource management system, such as a knowledge warehouse. Knowledge warehouses are supported by modern and relevant information management and network technology. The level of KM utilisation depends on the field, and the strategy used. The technologies utilised may use the KM framework, which includes existing processes, the people within the firm, the technology used to operate and the available content regarding their day to day activities (Koç et al., 2019). These ideas fall within the KM process, where there is a cycle of development and acquisition of knowledge. The process will then be refined according to the business's needs. Knowledge can then be transferred, shared and stored in the 'memory' of the organisation, and after that collated for decision-making.

The recurring theme of implementation in the KM framework includes knowledge creation, knowledge assets, expert systems, systems thinking, methods and techniques. For an organisation to continuously find ways to create and discover knowledge, there needs to be a knowledge management process (KMP), which ensures a constant flow of knowledge. Nguyen (2018) supports the assertion by highlighting that the KMP phases comprise knowledge identification, knowledge creation, knowledge storage, knowledge transfer, and knowledge utilisation. These phases can ensure advantages are leveraged and in turn, create a competitive advantage. Figure 2.1 depicts the different KMP stages.

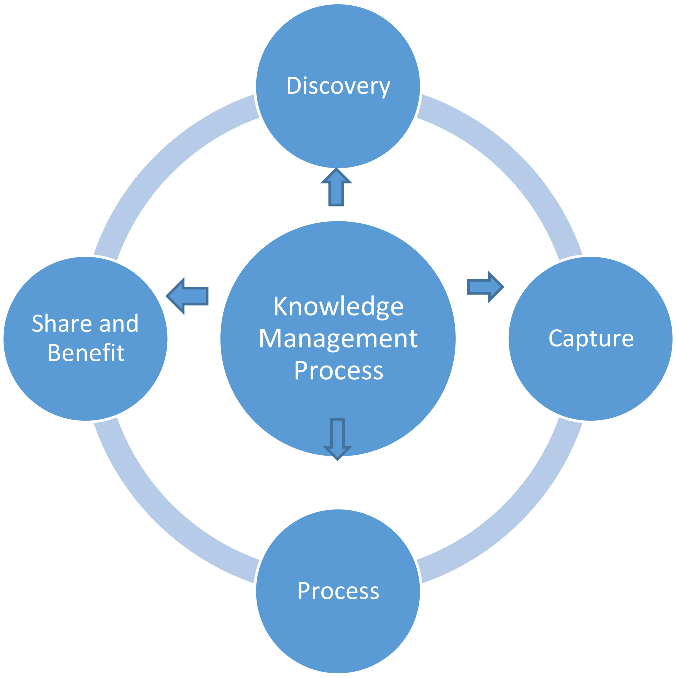

Figure 2. 1: Knowledge Management Process

The process shows that the actions that form part of KM are linked. Each idea represents the activity that takes place in the process, such as:

- Discovery - involves identifying the knowledge in the firm.
- Capture - looks at categorising and mapping old and new knowledge, storing it in a warehouse, and applying the metadata to it to be easily managed and accessed.
- Process - applies the analysis and assessment of the data, information, and knowledge captured.
- Share and Benefit - involves giving the stakeholders knowledge access and establishing the potential benefits and who will benefit from it.

These activities should take place on a platform that supports the organisation's activities and decisions. Zhang and Liang (2006) suggest that a knowledge warehouse can provide an intelligent analysis platform to decision-makers by enhancing the retrieval and sharing of knowledge across the organisation while facilitating, capturing and generating knowledge. The need for a knowledge warehouse is born of a challenge of how organisations can make use of knowledge and previous experiences to support decision-making on a managerial level, since the intellectual assets are much more influential and exist in the minds of management and staff, and in the external environment in the organisation. A knowledge base or warehouse can capture, code, organise, store and analyse knowledge in different types, such as explicit and tacit knowledge; and in different forms, such as plainfile, binary object or model (Dang & Yuan, 2011).

Karami et al. (2015) posit that KM system implementation should reflect what is considered critical success factors, which are identified as:

- Organisational culture dimension – includes teamwork, knowledge sharing and trust culture.
- Human resource management dimension – employee remuneration.
- Incentives and development, knowledge-oriented evaluation, recruitment and promotion based on knowledge competencies.
- Goals and strategies dimension – includes organisation's strategy alignment, strategic KM, upper management support, appropriate allocation of the budget, and empowered staff.
- Information Technology dimension – includes IT infrastructure, consistency between system and users, frequency of technological tools, IT and KM compatibility and information resource distribution networks.
- Organisational factors dimension – includes a typical and versatile knowledge structure, dedicated KM team, knowledge of products and services and enablement to collect employee and customer feedback.

Heisig (2009) highlights that a successful KM will have the following factors: 1.) The human aspect of the organisation (people, culture, leadership), 2.) Organisational structures and processes, 3.) The infrastructure and applications are determined by technology, and 4.) The strategies, goals and measurements are determined by management, supporting what can be observed as a standard for KM success.

The literature highlights that an organisation must have a set of criteria to follow to implement KM and measure, whether the project succeeded. The project's outcome can then offer an overview of what needs changing or adjustment to improve KM.

2.4 Challenges of Knowledge Management Application

How people interact with knowledge differs from person to person, organisations to organisations, and each will have their preference in approach. Sangiorgi and Siboni (2017) posit that for an organisation to perform, it needs to access and apply various resources to reach its objectives and goals. Knowledge is such a resource, and its creation occurs in varied ways and diverse situations; it can be restructured or redeveloped. The development depends on what the organisation needs the knowledge for or what decisions need to be made. This can be challenging to establish criteria for development.

A well-designed, developed KM approach translates into better responsiveness in the different phases needed to make the whole process work. Faniel and Majchrzakb (2007) highlight that KM technology fails in showing how to access the organisation's expertise, which is more important to the organisation, even though it can help find the sources of explicit knowledge, and thus the technology has little to offer. According to Ribiere and Calabrese (2016), organisations struggle with seven significant factors in KM application: the culture of the organisation, measurement/benefits, strategy, OS, governance and leadership, issues related to IT, and To anticipate and curb these issues, a clear policy about implementing a KM process is required. With the process in place, it creates opportunities for other systems to coexist. There are challenges in developing a KM framework in an organisation, which can be in the form of (Kalkan, 2008):

- The leadership's view and understanding of KM.
- Motivating and encouraging people (staff) to buy into the process, as the team may perceive that it is additional work. The team members may believe they need to learn how to use the tools required for the implementation.
- The changing technology used for and in the process.
- Identifying who the knowledge contributors are.
- Which knowledge is important?
- Interpreting the data.
- Maintaining the relevance of the information, its accuracy and keeping it up to date.
- Making the knowledge available, by which parties and to whom.

- Whether the business knows its own goals concerning the use of KM.
- It delivers tangible business benefits that support organisational objectives and priorities.

The answer to ensuring these issues do not overshadow planned implementation is that the appropriate technologies must be in place to facilitate better communication and capture activity knowledge for the knowledge base. In addition, Parlby (2000) highlights that organisations struggle to see the benefits of KM because there is a lack of uptake by users due to communication being unclear and lacking; it was not integrated into the everyday activities; users are discouraged because of either not having sufficient time to learn or because of the system is complicated; training is not offered, and users do not see personal benefits; persistent technical problems; and top management does not fully support it.

KM is more of an issue related to people and performance because the better people know and understand, the better an organisation will perform. Organisational Performance is the measure of output based on the set goals of the organisation. The concept includes three specific areas of firm outcomes: (a) financial performance (profits, return on assets, return on investment, etc.); (b) shareholder return (added economic value, total shareholder return, etc.); and (c) product market performance (sales, market share, etc.) (Richard, Devinney, Yip & Johnson, 2009). Measuring the performance of a KM project should be a relevant decision when conducting strategic planning. The presence of measurable objectives is a key part in creating an environment where there is employee engagement and commitment towards the organisation, which can lead to the organisation's performance (Abubakar, Elrehail, Alatailat and Elçi, 2019). Objectives such as profitability, staff retention and financial benefits can be used as indicators of performance assessment.

The lack of strategic planning and a failure to integrate KM into the organisation’s day to day operations will not help an organisation see or understand the benefits of the process. The organisation cannot realise KMs possibilities on the impact on profit, share price, employee retention and development.

2.5 Data Mining

DM is defined in applicable and practical terms as: "the analysis and non-trivial extraction of data from databases to discover new and valuable information, in the form of patterns and rules, from relationships between data elements" (Hirji, 2001). DM research can be categorised as i) Data Mining Functions - can be used to define the pattern types or knowledge to be discovered during the data mining process, ii) Data Mining Techniques - the number of data mining techniques or approaches that determine the performance of the data mining tasks, iii) Data Mining Algorithms - also known as methods, carry out data mining functions based on data mining techniques, iv) Data Mining Domains - data mining can happen in a set of a domain, such as business data mining, educational data mining, and web mining, and v) Data Mining Applications - it is where one or more data mining function can be used based on a set of application areas such as fraud detection (Gupta & Chandra, 2020).

With significant growth in interest for the DM applications, a standard and specifications were developed to define the necessary steps of DM for knowledge discovery. DM application for business operations early adopters came up with the Cross Industry Standard Practice for data mining (CRISP-DM) model. CRISP “provides a nonproprietary and freely available standard process for fitting DM into the general problem-solving strategy of a business or research unit” (Larose & Larose, 2014, p. 33). The model is iterative and adaptive and contains six phases, which form a DM project’s life cycle. According to Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer and Wirth (2000), the six phases in CRISP include Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation and Deployment. Bosnjak, Grljevic and Bosnjak (2009) suggest that CRISP-DM as a framework for discovering knowledge applies in many scenarios. This framework provides guidance in the application and use of DM and highlights it as a contributor to KM. The model has become the basis for using and applying the different aspects of discovering new and old knowledge, making it easier for organisations to lay out their business strategies. This encourages a better understanding of data for easy preparation, modelling and evaluation, so deployment can take place to support the firm reach its goals.

One of the key aspects of DM is that its development goes beyond human intelligence and human activity to exhibit fewer errors. Unlike the culture of KM, where humans manage and share with no intervention of technology. DM has become an action understood as a tool to extract and mine the patterns identified in enormous amounts of data. The field of DM and its functions, such as classification, prediction, association, and clustering, have developed to contribute to daily business operations. DM exists as a method to extract the patterns within the KM (Wahab & Rahman, 2018). According to Han and Kamber (2001, p. 1), IT has transformed so much over the years that DM is perceived due to the natural evolution of that change. The advancements in technology make DMs process possible automated, which makes it common to use used in different areas, including the marketing area (Bach & Alessa, 2014). IT capabilities are a prerequisite for the existence of systems that work for the organisation. The capability of IT is categorised within three-dimension: IT knowledge-organisation's understanding of the abilities of existing and new technologies and the employees' capacity to use them; IT operations- implemented activities supported by IT and IT infrastructure - the resources and tools that contribute to the acquisition, processing, storage, dissemination and use of information (Crawford et al., 2011).

DM technologies are grouped into classification, statistics, regression, association and clustering (Su & Wu, 2021). The types of techniques used and the features selected determine DMs efficiency, so it is important to identify which combination of best-performing algorithms works well with significant features (Amin, Chiam & Varathan, 2019). Of the many technologies identified through DM, OLAP has featured as a commonly used technology. There are great capabilities within OLAP (Online Analytical Processing) used for analysing data in several ways, supporting the provision of quick answers to complex queries based on the needs of management (Abusweilem & Abualoush, 2019). The data exploration concept can be utilised in DM to combine significant quantities of unrelated data, establish practical correlations, and retrieve meaningful information from the data (Su & Wu, 2021).

Decision-making is supported by several systems and processes, and one of those is a data warehouse. According to Zhang and Liang (2006), executive decision-making is supported by activities of a data warehouse because it provides information needed by extracting, cleaning and storing large amounts of operations data and OLAP querying on the data can take place to support the decision-making (Abusweilem &

Abualoush, 2019). They argue further and highlight that a data warehouse can only give a fraction of the needed information. The enterprise's knowledge assets provide the vast majority of the needed information. Therefore, the traditional data warehouse comes up short in providing adequate knowledge for decision-making purposes (Apte, Liu, Pednault & Smyth, 2002). In an ideal situation, the combination of both the data warehouse and knowledge warehouse would have potential because they would include their different data and knowledge types in one system to support decision-making (Bharara et al., 2017).

Other sectors are progressively moving towards leveraging their data when traditionally, they never pursued such action. With industries, such as universities, using Online Analytical Processing (OLAP) is prevalent, where the process analyses the data within businesses - data, such as percentages, profit margins, and trend analyses. In the context of universities, OLAP makes applications in different aspects possible. To understand what the pospective and current students are searching for on platforms such as university websites, OLAP can help reveal information about which hostels they would like to stay in or which courses, they are interested in enrolling for. The application of OLAP can turn a simple entry such as a date into useful information for the university to use in planning for the future intake or strategies.

DM is synonymously interchanged with Knowledge Discovery in Databases or the abbreviation KDD, but DM is considered a step in KDD by some (Geist, 2002; Luo, 2008). Others argue that KDD with Cross Industry Standard Process for data mining (CRISP-DM) and Sample, Explore, Modify, Model and Assess (SEMMA) model, are process models for DM (Selamat et al., 2020). KDD explains a step by step process of sifting through vast amounts of data for existing and future knowledge discoveries. This data should be analysed appropriately, processed and interpreted to extract the knowledge applied in organisations. The firm can decide how the new knowledge may form part of the KM activities and decision-making and incorporate it into their business processes. Kvassov and Madeira (2004) state that the Knowledge Discovery process is commonly understood as a sequence of interactive steps, such as data selection, data cleaning, data transformation, data integration, modelling, pattern evaluation, and knowledge representation. According to Martínez-Plumed, Contreras-Ochando, Ferri, Orallo, Kull, Lachiche, Quintana and Flach (2021), DM is more goal-oriented and focus on the process is at its core. It is a process with its genesis from a business goal that is relatively clear, with readily available data that has been collected for further computational processing.

The general perception of DM has traditionally been one of a very technical process where it needs active organisational involvement. Agarwal (2014) puts forward that in business decision-making, valuable inputs are drawn from systems, such as Business Intelligence and Analytics, which allow for the combination and transformation of data into information. Over time, there has been a growing change in the connection and relationships of DM to different aspects of a business model and process, such as KM. The move has mainly been due to the growing difference in the environment, the competitors and competition, the needs and demands of customers, technology and thus, the solutions needed in meeting the move accordingly.

The change in the overall environment constantly pushes the organisation’s processes to adapt. Due to unintended consequences, integration became inevitable, creating more possibilities for solution finding in response to the ever-changing environment. There has been some acknowledgement of the connection between DM and KM. Han and Kamber (2001, p. 5) highlight and argue for how DM should have been named “knowledge mining from data” because of the activities in the process, but that would not be a real emphasis on the role data plays in KM.

2.6 Evolution of data mining

DM has evolved over the years, where the process is known and associated with technologies, such as analytics, predictive analytics, machine learning, Big Data and KDD (Gupta & Chandra, 2020). There are technologies from IT and AI that are identified as key enablers for KM as listed by Devedzic (2001), such as intranets, knowledge-based systems, Ontologies, groupware, document retrieval, pointers to people, decision support, Xtensible Markup Language (XML), browsers, DM, intelligent agents, and databases are considered to be the major IT/AI components in the KM field. Several steps make up the process of DM, where useful information and subsequent knowledge are identified in large databases. These steps include using AI, statistical, mathematical and machine learning techniques to obtain the needed understanding and knowledge from data.

There are main stages or phases of DM that form part of the process shown in Figure 2.2.

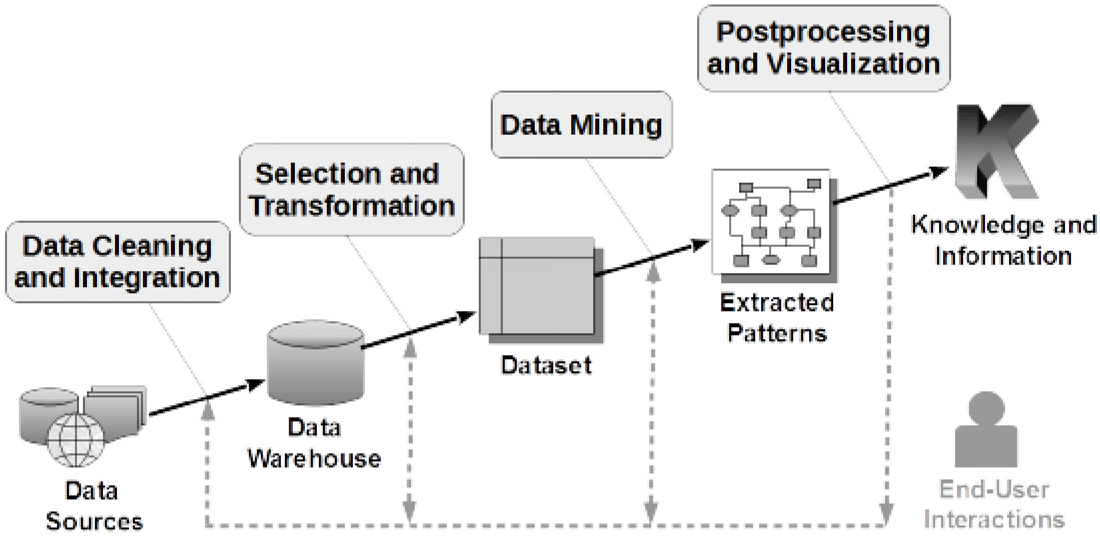

Figure 2. 2: Main phases of a data mining process

Source: Inthasone, Pasquier, Tettamanzi & da Costa Pereira, 2014

The DM phases show the process from the moment data is identified in the different sources until it goes through cleaning and integration to be added to the data warehouse for further processing. Then it goes through a selection and transformation stage to turn the data into datasets for mining. DM is employed, where modelling and algorithms can be applied to the datasets so patterns can be extracted. The results can go through post-processing and then visualised so knowledge, and information can be established.

A better understanding of the difference between DM and statistics is needed, as they can reveal different patterns and details from the data. DM differs from statistics because it focuses on various data fields, not just varied numbers, as in statistics (Huang, Chen & Lee, 2007). The idea behind the processes is to understand the past, understand previous decisions that were made, and then look at what these details can reveal about the firm's future. Based on Wu, Lee, Tseng and Chiang (2010) understanding, yesterday's success rules can be outlined by statistical methods, while DM techniques can explore tomorrow's success clues. Using the different methods and algorithms allows for different questions to be answered and solutions provided.

The general understanding regarding DM techniques, according to King (1999), is that they are primarily for:

- Predicting trends and behaviours - where the techniques are used to sweep through historical data to provide insight into future activity, such as a company wanting to analyse a direct marketing campaign and assessing who responded to which offer when, and used it to predict the results of future campaigns.

- Discovering previously unknown patterns - where new patterns may be discovered or identified and, for example, a company realising that phishing scams happen at a certain period of the year and specific times through new patterns.

The Techniques:

For the vast amounts of data to be mined and codified for pattern identification and extraction, applying numerous techniques and models occurs to develop conclusions and for the creation of knowledge. The applications can use different techniques and models. The data interpretation will change depending on the goal at hand. The data used for this process need not be new; it can be collected. Neaga and Liu (2014) highlight the ten DM algorithms that are considered highly influential, ranked in no particular order. These techniques are as follows: k-means, C4.5, SVM (Support Vector Machine), EM (expectation-maximisation), Apriori, PageRank, AdaBoost, Naïve Bayes, kNN (k-nearest neighbours), and CART (Classification and Regression Tree). These algorithms are commonly used for clustering, classification, regression, network analysis and association rules.

Alasadi and Bhaya (2017) describe the techniques:

- ***Classification*** **-** is a DM technique that is more complex. The technique forces you to collect assorted characteristics together into distinct categories. Conclusions can then be drawn from, or serve some function, where you can explain or make predictions over a class value. The methods found in the technique are classification by decision tree induction, Neural Networks, Classification Based on Associations and/or Support Vector Machines (SVM), and Bayesian Classification.

- ***Pattern Recognition*** - The key driver of DM is finding patterns. These usually recognise some oddity in the data at regular intervals or an ebb and flow of a particular variable.
- ***Prediction/ Induction Rule*** - Prediction is used to project the varied data types you will see in the future, making it a valuable DM technique. An example of this is when historical trends are identified and understood enough to plan a somewhat accurate prediction of what will happen in the future.
- ***Regression*** - Regression is used to identify the likelihood of a specific variable, given the presence of other variables. It is also used primarily as a form of planning and modelling. It looks at previous data to explain a numerical value based on that data. The methods found in the technique are Multivariate Linear Regression, Multivariate Nonlinear Regression, Linear Regression and Nonlinear Regression.
- ***Clustering*** - Clustering is seen as similar to classification but involves organising or clustering portions of data together based on their similar characteristics. Some of the clustering methods include Density-based methods, Hierarchical Agglomerative (divisive) methods, Partitioning Methods, Model based methods, and Grid-based methods.
- ***Outlier Detection*** - In many cases, merely recognising, the underlying pattern cannot give you understandable data sets. Identifying anomalies or outliers in the data is a necessity.
- ***Association*** - This technique is used in tracking patterns. However, it is more specific to dependently linked variables. In this case, identify particular events or attributes highly correlated with another activity or characteristic. Some methods include Multilevel Association Rule, Quantitative association rule and Multidimensional association rule.

Each technique can work, where a system can accommodate such techniques, such as machine learning systems. Machine learning is a subset of AI, where a computer can learn from previous experiences and activities without being explicitly programmed to do so. Other non-high-end programmes used are prediction query builders, with a type of wizard that allows the structuring of models and the capability to view that model. The decision-makers need to remember that techniques can provide accurate information when the right logic is applied. The intent is very clear about what information is required from the data (Rygielski, Wang & Yen, 2002).

In some cases, data workers can apply more than one technique to find knowledge. It is not always possible to find what you are looking for from only one model or task (Liao, Chu & Hsiao, 2012). There are also other statistical models or algorithms used in DM, which are part of the techniques (Microsoft, 2018). These need to be developed based on the different criteria of the information needed. Algorithms in DM, especially from the machine learning lens, is a group of calculations and heuristics that create a model from data. A representative model is created by analysing provided data and looking for specific patterns or trends, which must be carried out using the algorithm. The mining model is created through optimal parameters when the algorithm uses the results of this analysis over many iterations. Extracting actionable patterns and detailed statistics comes about when these parameters are applied across the entire data set (Microsoft, 2018).

There are a few algorithms that can be used depending on what the need is. There is no best or "be all" algorithm because its ultimate goal is achieving a prediction. The following are some algorithms observed in the studied literature:

- Association Algorithm
- Multilayer Perceptron
- Clustering Algorithm
- Rough set theory
- Decision Trees Algorithm
- Time Series Algorithm
- SEM Approach
- Linear Regression Algorithm
- Regression Analysis
- Neural Network Algorithm
- Logistic Regression Algorithm
- Sequence Clustering Algorithm
- Naive Bayes Algorithm

These models aim to show us that they assist in predicting and forecasting future outcomes. The utilisation of these algorithms can occur in the following scenarios:

- Identifying prospective buyers.
- Forecasting future sales.
- Understanding customer buying behaviour and patterns.
- Leveraging customer knowledge about the brand and products.
- Determining product placement.
- Calculating risk within a group of clients/customers.
- Grouping at-risk patients.
- Investigating fraud.
- Tracking site activities by customers.
- Predicting successful students.

Application of algorithms occurs in many spheres that regularly generate data. Luan (2002) believes that all different models or algorithms used in the business sector directly apply to research within higher education, especially in institutional research.

In many environments driven by the discovery of knowledge, you find connecting ideas that make discovery and development possible. Knowledge workers and data miners can trace the logical steps in creating knowledge and involve levels of workflow towards the ontology. A growing number of developmental experience has taken place in terms of the application of DM in organisations. DM has more interests and uses in different industries, whether a research agency, higher education institution, health industry or business. DM is described by Luan (2002) as a type of predictive and exploratory data analysis. An analysis that works well, is one where systematic associations between variables are laid out when there is no (or partial) prior anticipation as to the nature of those relations.

Firms should follow the phases in the DM process as a guideline, which ensures that they be mindful when pursuing to implement the action. The decision means there should be a concerted effort in preserving data, having a clear idea of what the data will be used for, having a well-designed modelling technique and then providing answers to the identified business problem. Clarity and conviction of the decision can support businesses to conceive better decisions based on high probabilities by using methods and systems that are fact and trend-based (Kriegel, Borgwardt, Kröger, Pryakhin, Schubert & Zimek, 2007). These can be highly effective when dealing with a powerful client-centred platform.

Even though the method continuously grows, evolves, and improves, there are still misconceptions about what DM can do for organisations. However it is not always the answer to all the challenges experienced, both strategically and operationally. There are fallacies that Larose and Larose (2014, p. 6-7) highlight in their book, and this is how they have broken them down with what the reality for each fallacy is:

Table 2. 1: Data Mining Fallacy and Reality Comparison

Fallacy	Reality
To discover solutions to our obstacles, we can apply DM tools to data repositories to help with solving them.	Organisations need to invest in DM tools because they will not fall on their laps. It is not a process that is automatic, which will mechanically solve your problems "while you wait."
Autonomy is how the DM process operates, where little or no human oversight is required.	Expert human element and supervision are still necessary and required. Using DM software with no strategic planning will only prolong finding a relevant solution. Instead, it will give the wrong answer to the wrong challenge, applied to the wrong data. A scenario where incorrect analysis occurs must be avoided since policy recommendations are based on such analysis, proving to be a costly failure. Even after deploying a prototype, introducing new data often requires the same prototype to be updated. Human analysts must assess the quality monitoring and other evaluative measures continuously.
The setup of DM is such that it pays for itself.	The costing depends on a few variables, and the return rates vary. The move depends on costs, such as the start-up, data warehousing, analysis personnel, and preparation.
The design packages of the DM software are meant to be easy to use and intuitive.	Use can vary, especially the ease of it. The software can over-promise capability. There is no automatic resolution to your problems by just purchasing a DM software, installing it, and sitting back to watch it do all the work. For instance, specific data formats are a prerequisite for algorithms, where substantial preprocessing is required. Data analysts must apply their expertise by combining their knowledge on the subject matter, analytical skills and awareness with the overall business or research model.
The business or research problem causes will be identified by DM.	Uncovering patterns of behaviour can be distinguished through the use of the knowledge discovery process. It is up to the knowledgeable workers to identify the causes.
The expectation of automated cleanup of our messy database through DM is a given.	The organisation does not happen automatically. As the beginning phase in the DM process, data preparation often deals with data that has not gone through assessment or been used over an extended period. Organisations initiating a new DM operation will frequently confront the data problem that has not been attended to for years; it has become stagnant and in need of extensive updating.
Positive results are guaranteed when using DM.	Like any other process, there are no guarantees of positive results when seeking actionable knowledge through DM. A cure-all is not what DM is but rather for solving business problems. When used properly, it is due to people understanding the data requirements, the applied models, and the project's overall objectives. DM can provide actionable results and profit-making outcomes.

2.7 Application of data mining

There are two primary objectives in carrying out DM, (1) to give rise to predictions and (2) to uncover new and distinct insights (Selamat et al., 2020). Tracking of the behaviour of individuals and groups, processing of medical information, and several other applications have occurred because DM has been applied successfully for several years in the scientific, medical and business communities. According to Bosnjak et al. (2009), DM is not an easy process. They view it as a time consuming and challenging endeavour. An endeavour is perceived to have no guarantees of discovering any exciting or potentially useful patterns without having several large data models to work with and use. DM is found in industries whose business model is motivated and driven by customer focus. Some businesses that use DM include Healthcare, Education, Manufacturing/ Engineering, CRM (Customer Relationship Management), Market Research (Customer Segmentation), Financial Banking, Research Analysis, Research Analysis and Bioinformatics. Mamcenko and Beleviciute (2007) are of the view that DM is more suitable in organisations that store and preserve significant amounts of data, such as bank transactions or similar.

When DM is implemented, there is an expectation that it will sift through the large data volumes and discover patterns of relevant and interesting information in the following four phases: (1) Filtering data, (2) Selection of variables, (3) Extracting knowledge, (4) Interpretation and evaluation (Manjarres, Sandoval & Suárez, 2018). Programming methods can be used through DM to identify patterns among data objects. To organise knowledge so it is structured and understandable, knowledge discovery uses an ontology that highlights the relationships between processes or actions; and how their relationship ties and connects into KD and KM. Žáková, Křemen, Železný and Lavrač (2011) discuss a knowledge ontology as defined connections between the ingredients of knowledge discovery framework, both declarative (varying knowledge representation) and algorithmic. The authors provide details that highlight the different building blocks of the ontology, which are:

- Models and algorithms
- Pattern sets
- Boolean language
- Simple data knowledge

- Complex non-logical knowledge
- Logical knowledge
- Classified and Dataset
- Language bias

Dennis, Marsland and Cockett (2001) assert that data warehousing is another theme that is a commonality in establishing the build-up of the knowledge in the organisation. Application of data warehousing in many business processes is essential, so it becomes natural that it is part of the DM process. The intention behind data warehouses is for the company-wide information to be centralised, to design and fulfil the necessary analytical environment (for knowledge discovery and DM) to respond to changing business needs (Chowdhury, 2009). It is beneficial in that trend analyses can be drawn and bring about new perspectives to prepare for or readiness for service delivery. Some features of data warehouses include being subject-oriented, integrating diverse databases, its time-variant aspect, and the non-volatility, making it understandable why it is included as a step in the DM process. There are three steps to using DM, which are:

- Exploration
- Pattern identification
- Deployment

Bharati and Ramageri (2010) define the steps as:

Exploration: Data exploration's first step is for data cleaning to happen and converted into another style and variables considered essential, and determining the nature of data based on the problem.

Pattern Identification: Once the exploration of the data takes place, it is filtered and characterised according to the specific variables, and then forming pattern identification is the second phase. Then the patterns which make the best prediction are identified and chosen.

Deployment: Deployment of patterns on different cases or scenarios for the desired outcome, such as applying some DM techniques to support KM, by using neural networks and association rules in marketing modelling as highlighted by Zekić-Sušac and Has (2015).

To model a platform for DM and KM to cross paths, a conducive environment, such as a business with a dedicated environment for practice and action of KM and DM must be in place. That means that companies need to be operating so business and competitive intelligence are part of the processes and part of the businesses' day to day operations (Wang, Lin & Chu, 2011). Business intelligence is referred to or known as software applications that can analyse vast amounts of raw data. Negash and Gray (2008) highlight that to produce input to a decision process, business intelligence (BI) as a data driven DSS integrates data gathering, data storage and KM with analysis. Competitive intelligence means understanding your organisation's field and who the competitors are in the environment and using that information to make better business decisions. Collecting, analysing, delivering and using publicly available information on activities outside the organisation's walls forms part of the competitive intelligence. Kokubo (1993), states that competitive intelligence that is designed into the business system of any world-class organisation has the advantage of improving R&D effectiveness, supporting the management decisions. Consequently, the company's business can pick up early warnings of threats. Effectively combining these two elements can thrust the organisation ahead of the competitive environment.

There have been discussions as to how much of an impact DM has had and continues to have on business, such as the argument that DM does not contribute to companies on a large scale, based on what Pechenizkiy, Puuronen and Tsymbal (2005) argue, but this does not negate the potential of the impact that DM may be having or could have on KM. Some discussions have highlighted significant perceptions of DM as one action in the stages of the knowledge creation and capture phase of KM. To get to meaningful information and knowledge, sifting through enormous amounts of data must occur to understand what the data represents through interpretation tools.

KM has a lot to do with human subjective knowledge, where data is the main focus of DM. The idea behind the connection between the two is the extension and transformation of human knowledge (Dawei, 2011). Through applying DM as a tool for that extension, human knowledge goes through an act of creative and effective codification so that data mining can mine the collection. DM is an essential component of KM tools (Selamat et al., 2020; Silwattananusarn & Tuamsuk, 2012). Therefore, using algorithms as part of the process can support transforming knowledge into other forms of knowledge. DM can find five types of knowledge, "general knowledge, associated knowledge, classification, forecasting, and deviation knowledge" (Dawei, 2011). The different kinds of knowledge can support an organisation respond in varied ways to its challenges.

Due to the growing need and collection of data in organisations, the same data needs to be accessible for further processing, and it should be on a platform such as a data warehouse. Typically, as Apte et al. (2002) describe, data warehouses are constructed when a business relies heavily on data driven analysis for decision-making. The warehouse is there to capture as much information as their systems will allow about their customers. As part of supporting the decision-making process of an organisation, the data within the warehouse can develop decision alternatives by applying models to the data.

2.8 Challenges of DM implementation

There are different ways DM can be used, and when adding its combination with many other different systems and services, there is a complex aspect to it because then many require their own distinct, often conflicting standards. Implementing DM and integrating it into KM does not always translate into systems working well together, even if the KM approach has a "perfect" design. Just like KM, DM has its challenges with applications (Sharma, 2014), ranging from:

- Poor data quality (noisy or incomplete data) .
- Concerns over security and privacy of the data.
- Data mining algorithms can be challenging to apply to bulk data and massive datasets.
- Searching in heterogeneous databases and global information systems can be challenging, due to the vast amounts of data stored.
- Adjustment of models to the changing data.
- Large, complex and unstructured data processed into structured formats.

The knowledge discovery process has its challenges when it comes to data. Especially for those in the social sciences because there is not enough robust data, the same as smaller businesses that generate little data. The environment can make it challenging to have an efficient working DM system. Data availability and collection are a big task for some organisations. Data comes in different formats, and some of those formats make it possible for automatic data analysis to take place (Rygielski et al., 2002). This could be a problem because the examination assumes that the data is organised according to variables with predictive values. The process does not consider the potential of other data that may appear in different formats, making analysis difficult, especially when data interpretation becomes an issue. DM intends to extract complicated and diverse types of data, including data that may be challenging to mine due to their structure, such as text data, multimedia data, and the World Wide Web (Han & Kamber, 2001, p. 18). Text data is a much easier data type that can be investigated and used through text DM, as it is viewed as being closely linked to or an extension of DM because it adopts textual databases, and therefore, considered a knowledge discovery process (Usai, Pironti, Mital & Aouina, 2018).

Since text mining is an interdisciplinary concept related to DM, it can tackle the issue of quickly extracting information from large databases by bringing together DM techniques, machine learning, linguistics, pattern recognition, information retrieval, databases, statistics and visualisation (Zanasi, 2007). There are four approaches to text mining, where data assessment and analysis arise due to the prominence of the action. These approaches are Document Categorisation, Document Clustering and Similarity, Document Visualisation and User Profiling. According to Davies, Sure, Grobelnik and Mladenić (2005), define the approaches as follows:

- Document Categorisation - when there is a provision of a set of predefined categories and the function is to categorise new, or in some cases documents previously hidden, by designating each document one or more content categories, then this phase applies.
- Document Clustering and Similarity- this approach is built into the algorithm, general data clustering, which is adopted for text data. By viewing each document represented as a word vector (known as the "bag of words" characterisation), whereby each word's weight is determined by the proportion to the frequency number of the word. The cosine-similarity between the word vector characterisation of the documents is commonly used to measure the similarity of the two documents.
- Document Visualisation - to secure measures early of data quality content and distribution, the method of visualisation of data, and the textual material of a document set used.

- User Profiling - user profiling (or user modelling), web and in-text mining have many applications. The main aim is information filtering, either collaborative or content-based filtering. It is used with the user in mind, to decide what information is potentially interesting to them.

<u>Different sized data</u>

Applying DM to huge amounts of data comes with its challenges. Smith (2020) suggests that when variables are discovered, they can appear useful even when they are immaterial and can cause true variables to be overlooked and discarded. According to Neaga and Liu (2014), BI and analytics are essential in dealing with the enormity and consequences of data driven complications and with solutions using new patterns, information and knowledge uncovered through the DM application. Silwattananusarn and Tuamsuk (2012) argue that challenges will arise when looking to discover valuable information in the pools of data and attempting to translate the data into relevant insights.

Additionally, Grossman, Hornick and Meyer (2002) postulate that the lack of standardisation in DM, is due to challenges, such as model inputs being complex to put into practice, which can be categorised into two issues identified as an agreed upon general standard to clean, transform, and preparation of data for DM; and agreeing on a method for working with remote and distributed data using a common set of web services. In addition, Hirji (2001) asserts that the challenges associated with the application of DM are primarily statistical, "that is, to infer patterns or models from data". Madni, Anwar and Shah (2017) assert that the more progress there is in DM, the more disadvantages are exposed, which are related to issues like misuse of information, security issues, privacy issues, use of inaccurate information, and risk of data loss.

There is always the challenge of mining sufficient data in small or medium organisations. This is due to the nature of their business model, structure, and clientele, which makes it difficult to gather large amounts of data for business decision-making. The organisations may already have a KM process in place. However, it does not ensure sufficient data. If it is the case that companies are short of data, they would need to investigate new ways to find and collect more data, depending on their requirements or need (Zain & Rahman, 2017). Data can come from various places, such as social media, consumer behaviour, business transactions, market data, and censuses (Gupta & Chandra, 2020). These can be data-rich. For example, in consumer behaviour, a business can look at purchasing patterns, consumer responses to products, services and marketing, and identifying competition, which is available in a customer database. Here, the different companies should focus on their market and who their customers are. They need to cater to customer-specific needs to retain them by using CRM tools. Shaw, Subramaniam, Tan and Welge (2001) highlight the marketing environment for consumers and what could be generated using CRM. They also highlight that, through DM tasks, a diversified group of customers and marketing knowledge can be created, forming the core of the KMP. Shaw et al. (2001) describe the steps and processes of having DM contribute to knowledge creation. The knowledge collected can be classified by indexing the knowledge components and filtered based on content and building linkages and relationships. That primarily highlights the importance of having defined knowledge elements.

Some of the literature noted that it could be assumed that knowledge creation will occur through Big Data and that organisations can increase their performance by taking advantage of that. The article by Sumbal, Tsui and See-to (2017) highlights the concept of Big Data based KM. The idea focuses on the knowledge creation that happens through the processing of data but also looks at the potential that tacit knowledge could have in the generation of knowledge with Big Data analytics. It also states that even in the oil and gas industry, it is much more challenging to have a precise model of generating new knowledge.

The web has volumes of data and information categorised to create valuable information but is challenging to manage because of the vastness of the data. The KM technologies, coupled with DM principles, expose the explicit and tacit knowledge in the organisation. Big Data has become a field that has grown in use. You find its presence, especially in different social media platforms, such as LinkedIn and Twitter, other multimedia and the Internet. Those platforms are mines of varieties and large volumes of data that can be analysed and used as BI. Big Data has the potential to delve deeper into gathering data that some tools and techniques may not easily extract.

Literature has highlighted the possibility of integration in some fields, such as marketing, business and higher education focusing on students. Marketing is a field increasingly viewed through the lens of CRM. Information, such as customer buying behaviour, opinions, interest in products and just the general life cycle of customer buying patterns is investigated and analysed. To show the potentiality of integrating DM and KM, applying association rules (AR) and neural network (NN) together paint that picture. Zekić-Sušac and Has (2015) highlight the potential of that view, where they developed a model that could best display the process. They developed a model to show the case of integration, which is made up of four phases. The model focuses on using AR for elements, such as the customers buying patterns, product choices and applying NNs to identify the buyer profile.

The second phase investigates the knowledge generated, and the third phase is where it comes from; the innovation the organisation can develop. Last, the marketing strategies the organisation can come up with from the investigation. This model supports the idea that organisations build their marketing strategies, and the overall approach based on what the data has revealed, and solutions derived from the process. The model consists of the steps generally found when seeking solutions. In some cases, the challenge may arise when it comes to the scope of CRM.

There are different scenarios found in the size data can come in, and that depends on the organisation's overall operations. Data can fall within the smaller end of the scale if a particular research area or industry has limited customer engagement (Dawei, 2011). The level of engagement can lead to more data, and subsequently, the discovery of knowledge. The circumstance can change or be adjusted based on how far and wide the investigation of factors, such as the customer experience can go. An example of a scenario illustrating the increasing of data would be, for instance, an investigation of customer behaviour of a particular line of shops, instead of focusing on just one of the chain shops, to attempt to increase the data sets and variables to further research.

Solomon, Ketikidis and Choudhary (2012), highlight that through the knowledge discovery process, using DM in industries, such as supply chain risk management can bring about the necessary knowledge. Business activities in such an industry depend on knowledge acquisition to gain much needed knowledge, which requires validation to form part of the risk knowledge base. Thus, the need for knowledge consolidation and utilisation requires a dedicated model to support their activities. They further argue that there can be a method that flows within databases, from the choice of a database and planning, DM, knowledge evaluation identified and utilisation of that knowledge to bring the activities to life. This can be achieved by using an alternative knowledge discovery model that focuses on DM and its processes, from which usage and decision-making are enabled in the early stages of the cycle. The literature supports the approach that each industry can adopt, using different models of knowledge discovery or the different steps in the cycle, to determine their actions to solve problems within the organisation (Dawei, 2011; Gupta & Chandra, 2020).

<u>Techniques must be uniquely applied depending on the question</u>

The literature suggests that depending on which industry uses DM to find the necessary knowledge, there will always be a technique applied accordingly. Discovery driven data has two categories in which you find supervised and unsupervised data analysis. According to Gupta and Chandra (2020), supervised and unsupervised data analysis (knowledge discovery) are two techniques used in DM. Supervised, predictive or directed, is used when you have a specific target value you would like to predict. This technique uses classification, regressions and anomaly detection as part of the process. Unsupervised, descriptive or undirected, does not focus on prearranged traits, nor are target values predicted and finds hidden structure and relation between data. The technique uses clustering, association and feature extraction as part of the process (Bharara et al., 2017). Classification can be used when working and looking for specific patterns or values. In the article by Jantan, Hamdan and Othman (2012), they highlight that the knowledge discovery process can be enhanced by applying the classification rule and using DSS for predictions based on previous information, outcomes or goals.

According to Silwattananusarn and Tuamsuk (2012) study, you can assign specific techniques to specific industries, such as working with financial resources. Types of knowledge include knowledge sets made up of models, parameters, reports, and other data. The knowledge-sharing process to corporate bond classification where the DM techniques used are classification and clustering. These same techniques can further be broken down into specific models and software, such as the self-organising feature map, short for SOFM, a type of NN and/or the learning vector quantisation, which is short for LVQ that can be used in bond rating. After that an ontology of KM and knowledge sharing and a financial KM system can be applied. When building a knowledge base or a field, such as entrepreneurial science, the techniques commonly used are classification, which can be dissected further into a knowledge-based system, knowledge and information network, and the use of reasoning and pattern recognition.

Based on the argument by Sassenberg, Weber, Fathi and Montino (2009), the combination of feature selection, which is an automatic selection of traits in data, and NN, which is a computer system modelled after the nervous system. KM is supported by quick solutions enabled in scenarios with technical problems. The combination works when trying to establish solutions to a specific issue as it may have fewer features to work with to allow accuracy in the algorithm of the process. In the study conducted by Gröger, Schwarz and Mitschang (2014), in developing a manufacturing knowledge repository, they came up with a model called Insight model, which includes DM as one of the main components and the approach is based on the model using and centring on structured data, guiding the system design and how it will function.

The study conducted by Stanley (2008) looked at the application of supervised and unsupervised knowledge discovery where the data sets can be broken down into smaller sets, leading to predictions in the higher education sector. According to Mamcenko and Beleviciute (2007), DM "can be performed in a highly automated fashion on single documents or archives of related documents, using both expert systems and conceptual clustering," where things like curriculum planners, instructors, and learners can be used to perform an in-depth search for subject matter related to those aspects.

It is possible to omit some steps in either the knowledge discovery or DM processes. The point is to fit the challenges of specific goals into the plan that the organisation wants to develop or grow. This view supports the position that organisations set the boundaries of system use or engagement. The system does not dictate how the business conducts its activities. Kishore, Kumar, Grover and Maloo (2014) show their views on the use of DM in KM by reasoning for applying four DM techniques. These four techniques are classification, clustering, dependency modelling and summarisation. These were applied in healthcare, the financial industry, construction industry, retail, and small and medium businesses. Each sector has its own set of combinations of approaches to assess what type of data was mined and what type of knowledge can be discovered using the different techniques. The literature trends show that hybrid models will emerge from DM techniques in the future (Gupta &

Chandra, 2020) because of how flexible and complex they are. The techniques involve a combination of either clustering and association rule or clustering and classification due to complicated scenarios needing complex solutions.

Expertise is required

Other industries also face serious challenges with the action of generating knowledge. This is an issue, especially if the knowledge is mostly tacit from the relevant and experienced staff in the organisation. The matter highlighted by Yoshioka, Tomioka, Hara and Fukui (2010) reflects on attempting to give novice engineers support through the creation of a record information system. Where there can be an understanding of tacit knowledge exchange, learning from previous engineers and their experiences is usually difficult to achieve. The mining and analysis of patterns of different experiences recorded by project team members, or just the day to day activities, should be carried out to create new knowledge for up and coming people in the environment and or future decisions.

Learning from previous experiences is a useful practice of gathering information about how to make improvements and changes. Maier, Kalus, Wolff, Kalko, Roca, Marin de Mas, Turan, Cascante, Falciani, Hernandez, Villà-Freixa and Losko (2011) developed a KM framework that is used as a basis to planning that allows inputting specific knowledge from clinical and experimental data and using text mining to build a knowledge base for users. Such developments make it possible for firms to design models with the capacity to analyse previously untapped data. This action establishes new knowledge for the business that can strengthen their KM approach. Some organisations face challenges when pursuing a sound KM system. The problem is not because of a lack of knowledge assets but because there is insufficient understanding of what it does. Other possible challenges include the lack of buy-in from employees, which leads to little interest in documenting their experiences (tacit and explicit knowledge) or improving internal knowledge by accessing and retrieving it from organisational storage.

The team involved in the efficiency of DM, and KM are another aspect in the challenge of DM implementation. Wang and Wang (2008a) suggest that during integrating DM and KM, teamwork is necessary while planning for the processes to pursue an effective DM in the context of BI. The knowledge workers, data miners, and business insiders need to collaborate to build a system that considers all the facets involved in creating and using knowledge. They further reflect that the business insider centred knowledge development cycle to the DM cycle, and the knowledge sharing system enables collaboration between the three parties (Kishore et al., 2014). The different groups can share and make decisions based on their knowledge. The argument for knowledge workers to be part of the DM process and build the KM framework with greater integration improves the competitive advantage (Heinrichs & Lim, 2003).

Another challenge with DM implementation is that some processes are harder to automate and where there are parts that cannot be done automatically through DM. The expert knowledge can fill in the gaps. The filling in of the gaps comes in the form of expert understanding of the domain being transformed into queries and models needed for the steps that follow (data understanding, preparation, modelling and evaluation) (Martínez-Plumed et al., 2021). DM still requires some level of interaction with knowledge and data workers, as it is not an automated activity (Wang & Wang, 2008a).

<u>The technologies required to drive DM are very diverse</u>

The literature speaks on DM to discover knowledge processes from the available data stored in company databases or repositories. KM is a strategy used to establish which information and knowledge are necessary for day to day decisions by using different tools and techniques. Most KM and DM techniques require the existing data to reflect learning patterns and information, and are therefore, constructed upon the foundation of machine learning and AI (Tsai, 2013).

AI has proven to be efficient, and it is a factor that is crucial during the process of generating new knowledge from existing knowledge. Nemati et al. (2002) make a case for AI offering a step in creating new knowledge. AI is a DM tool, as it can be used to find patterns. This technology reveals important facts and relations at much faster rates than traditional ways of analysing data. The other argument presented is the use of ontologies in DM and its contribution to KM. That is seen in the discussion by Wang and Wang (2008b). They attempt to reflect that the ontologies of DM can be beneficial for KM where a common understanding of the context and techniques of DM is shared among the data miners. The idea is to have an ontology-driven DM system, where ontologies are structured, used and understood by people who would be mining the data. Li and Lu (2004) argue for an ontology-based generic grid model that allows knowledge integration and discovery on the grid (Universal Knowledge Grid).

The tasks and methods, such as text mining or text analytics and natural language processing, make it possible to transform or translate essential items into understandable content but require prior knowledge, so they can be implemented. As much as, DM is seen as a solution to many organisational data challenges, in some cases, it has been realised that it may not contribute meaningfully to business, as argued by Pechenizkiy et al. (2005). The article explains that DM systems and solutions do not add value on a large scale compared to Database Management Systems (DBMS). A DBMS is a system known as a software package designed to explain, influence, recover and organise data in a database. This means that to get more out of DM, insight beyond just the basic system alone is required. A greater understanding of the different branches and techniques available makes a significant difference.

An entire DM and KM process that is interlinked can be created based on organisational frameworks and goals. The need to find trends in different places has grown over the years that DM evolved and gave rise to technologies like AI, which uses different techniques.

<u>Highly dependent on customer focused industries</u>

The application of DM in industries with a substantial customer focus has shown that businesses can discover many valuable patterns in their data. Using DM techniques allows customer segmentation-based preferences and choices, which can be used to predict or forecast the customer's future preferences, and market basket analysis to better understand the customers through their activities. This can happen in the presence of a customer data warehouse, where analysis takes place. Therefore, it supports the view that decision-making processes for the business executive become easy, with information related to trends available and made accessible by DM techniques and tools (Jayanthi & Vishal, 2011).

This also highlighted that an organisation with CRMs uses DM techniques and tools to integrate KM. This means that organisations that can implement such systems focus on overall customer activities to generate such large volumes that DM can be applied (Ranjan & Bhatnagar, 2011). The integration in such organisations intensifies their

KM, meaning that firms could establish which potential customers will spend more money and time on the business and encourage loyalty to the brand, products and services. This way, they can custom make products and services and target advertising for such customers. This creates a back and forth the development of data and knowledge generation.

Neaga and Liu (2014), highlight the gap that exists where a KM framework can be defined with data analytics. That deals with the challenge of integrating people’s knowledge and expertise through computational tools that explore Big Data (Sumbal et al., 2017). Such tools include sentiment analysis and opinion mining, focusing on platforms such as Facebook and Twitter. These tools can analyse the emotions of people and their opinions about services and products. In such instances, unstructured data and text could be analysed to look at the patterns, and a framework would determine what was worth being identified and used.

Mehmood and Maurer (2013) discuss integrating web images to generate and incorporate them as information sources, even though they are from different sources. The information integration model makes it easier to search for these images through an information integration system, where the mapping of these images takes place. The model highlights the potential of DM task and tool application in Big Data, where identification of various patterns occurs (Sumbal et al., 2017).

The literature highlights a key challenge for DM and KM to occur within the same process or establish a clear link. The organisation needs to know what the objectives of the process are, what information and knowledge are required, and what decision. It is difficult to activate the systems without mapping the needs. It is essential to establish frameworks to have a solid foundation.

2.9 Overview of literature

The current state of DM practice shows a rapid change, which has supported many more organisations' activities. How businesses conduct interactions had transformed, but many organisations have kept their key goals and definitions. The most common and primary goal of many businesses is to position themselves ahead of the competition and provide better services to customers. The significant change over time with business operations has been the explosion of different customers with different behaviours and the technology developed that makes it easy and quick for everyone to reach different markets. Organisations have had to re-evaluate themselves and redesign how they run their businesses. With the fast-paced change taking place, it is becoming harder and more challenging for companies to predict the future. Lawal, Odeniyi and Kayode (2015) state that the market environment is changing so much that the customers hold power as they become the competitors due to the variety of options for product or service providers. Competitors that have been in the business for a long time are partnering because their services have become similar and improved over time, and the relationships with customers are developing. The factors contribute to emerging businesses.

The rapid change illustrates the importance of tacit and explicit knowledge. In the book by Becerra-Fernandez and Sabherwal (2010, p. 194), they discuss that socialisation can be a mechanism to discover knowledge through activities shared among masters and apprentices or between researchers at a conference specifically catering to what they do. This informal exchange takes place between people who share workspaces and environments. According to Nissen and Bordetsky (2011), they support this level of tacit knowledge flow, where learning takes place, and the individual masters apply the new knowledge via trial and error. Sharing the tacit knowledge can also occur through mentoring or master and apprentice relationships with a few colleagues in a group who participate in mentoring other compact groups, such as in various communities of practice. The organisation has an alternative option of formalising the tacit knowledge, such as creating a classroom course to capture that knowledge.

They further discuss that creating new explicit knowledge can be done through technologies, such as DM techniques. The collection of knowledge forms part of the organisational memory; the memory collected from the individuals, and the knowledge shared in social interactions within the organisations. According to Walsh and Ungson (1991), the different aspects of organisational memory extend beyond individuals' memory to include other components. The components include corporate culture, change in work procedures and production processes, the formal composition of the organisational roles (structure), the work setting of the environment (ecology), and information records (both internal to the organisation and external to it). This idea has not changed over the years, even as the environment changed. This shows that there are opportunities to harvest both tacit and explicit knowledge to benefit the business. This is a double-edged sword, especially with the vast amounts of data collected and stored, both intentionally and unintentionally. Studies have shown that not all data is useful for decision-making and knowledge creation just because knowledge is discoverable. It always depends on the type of system applied to the data and the types of knowledge necessary for the business. According to Khanbabaei, Alborzi, Sobhani and Radfar (2019), organisations can use DM to harvest knowledge in knowledge-intensive processes and to be competitive. These processes can have large amounts of structured (Customer Data, Sales Data, Financial Data, Demographic Data), semi-structured (Action Plans, Policies, Procedures, Cases) and unstructured data (Emails, Documents, Presentations, Videos), which can be mined to support the performance against their competitors.

Many organisations have moved on from only focusing on the traditional DM applications and have ventured into mining engines, such as Web data mining and Data Visualisation. An area that has received much attention over the years has been CRM. According to Mithas, Krishnan and Fornell (2005), the use of CRM application takes centre stage as it is associated positively with enhanced customer knowledge and enhanced customer satisfaction. CRM has become the heartbeat of monitoring the client group of businesses. Customer interactions have also evolved with changing technologies, which has provided more platforms for customers to voice and share their opinions, experiences and preferences of products. Customers communicate in natural human language, and firms realised that web DM and text mining would provide more opportunities to tap into the textual data and make sense of it. The tools make it possible to interpret the customer's voice by applying linguistic analysis or natural language processing (NLP). According to Becerra-Fernandez and Sabherwal (2010, p. 217-218), there are three uses:

- Web Structure Mining - looks at the form of web documents and attempts to highlight the underlying model link structures of the web.
- Web Usage Mining - looks at the identification of user navigation patterns through the different web pages.
- Web Content Mining - makes it possible to discover what the web page is about and how new knowledge can be unearthed.

The focus can help reduce operating costs for things like CRM and KM by using DM engines, such as text mining, semantic mining and sentiment analysis. Using these tools helps to understand customer opinions, as highlighted by Yamanishi and Morinaga (2005). The progression in development shows additional ways to identify old and new data on different kinds of platforms. This reflects the expansion of different mining methods available. With this growth, challenges emerge when attempting to achieve some advantage in incomplete or imbalanced data usage.

Research has shown that industries have been redeveloping their practices by moving towards more semantic technologies and machine “processable” data to support the analyses of large quantities of data. It is evident with the dynamic that Big Data brings. The analysis of the data is more challenging and requires complete automation due to its volume, velocity, variety and veracity (Labrinidis & Jagadish, 2012). To prevent possible strain or challenges with bulky data, developments in Big Data technology have emerged in response to deal with those scenarios. These technological tools led to less human involvement and sped up analysing data and quicker knowledge discovery. Mirza, Mittal and Zaman (2016), and Silwattananusarn and Tuamsuk (2012) discuss trends that help organisations predict the future behaviour of customers or products by using the same traditional DM techniques. Maksood and Achuthan (2016) argue that DM has seen successful applications in several industries, such as retail and services, medical and healthcare, and higher education. There has been a rise in activities in the medical and healthcare services industry moving towards using these systems because of the data collected and stored, such as clinical trials, health surveys, claims data, and patient and disease registries. These industries can be observed and their cases easily documented, but not without issues, as the adoption of technologies can determine how quickly the industries take advantage of the developments.

Lawal et al. (2015) argue that data capture and accumulation capabilities have increased at a rate higher than organisations can make the most of, because of the technological advancement. It is evident how imperative it is for the burgeoning of new tools and computational theories to support humans in creating and unearthing knowledge that is useful from the rapidly growing bulk of digital data. What each business needs is a well thought out KM framework in place. A framework supported by the whole organisation, where technology, the leadership, the people and the culture have a shared vision and an understanding of the goals and expectations. When established, internal investigations to discover what data and knowledge assets exist and how that can support the business's longevity to be set out. The outcome can support the blueprint of business operations.

Based on the literature, trends show an uptick in adopting DM and KM systems and technologies, with more industries showing interest. There is a growing interest in using the systems in smart cities with significant populations, information service businesses, crime prevention institutions, target marketing communication and sentiment analyses. This supported growth by emerging technologies is discussed by Cios and Kurgan (2005) as:

- XML (Extensible Markup Language) - has the capacity to describe and store the relationships of structured or semi-structured data. The standardisation of communication between diverse databases and DM tools becomes possible.
- XML-RPC (XML-Remote Procedure Call) - outlined to simplify processing, transmittal and return of complex data structures.
- PMML (Predictive Model Markup Language) - for defining models, the vendor-independent method is defined. The conflicting issues among applications and proprietary formats are removed.
- SOAP (Simple Object Access Protocol) - accessing objects, services and servers on the Internet using the XML/HTTP based protocol.
- UDDI (Universal Description Discovery and Integration) - can help to resolve challenges like identifying the right service from millions currently available or interfacing with a service using Web Services Description Language (WSDL).
- OLE BD-DM (OLE DB for data mining) is an SQL query language extension where users instruct and test DM models.

These technologies are available or built within commercial products, such as SPSS Modeler, Oracle Data Mining, Microsoft SharePoint and IBM Cognos that organisations can purchase to integrate into their systems. The popular technologies during this systematic review include data warehousing, OLAP, high performance DMKD systems, AR, applications of data mining, and visualisation techniques. The popularity of these technologies may have slowed due to other technologies arising and providing the necessary solutions to issues previously faced. The attention shifted towards new areas and applications. The literature shows that capacitated systems perform data handling and not people. Still, people need to understand the overall activities to input the correct language and required outcomes in the technologies. Organisations have more options to choose from in the market to purchase different technologies to support their systems.

2.10 Chapter summary

The literature provided a comprehensive background for the DM and KM processes to provide context. The chapter started by providing background literature into KM, its application and then challenges of implementation thereof. DM was also discussed, and its application and challenges. The chapter then offers an overview of the literature, which highlights trends and developments within both processes.

Chapter Three

3.1 Methodology

This chapter presents the methodology, and the different stages followed in the study. The different stages include the research design, data collection, data analysis and the reliability, validity and trustworthiness of the study.

3.2 Introduction

The study adopts a systematic review methodology, with the illustrated step by step flow of the process to highlight the motivation for the steps taken to identify publications for the study. The criteria used to determine the articles included as a unit of analysis are highlighted. Using a systematic review methodology, the study identified and analysed the research articles informed by the research objectives. The study also reflects on the connection of DM to KM through commonly found themes and the understanding thereof by applying thematic analysis. Motivation is provided for the decisions adopted for this chapter.

3.3 Research Design

A systematic review is widely understood as a literature review designed to discover, evaluate, and synthesise the best evidence to provide enlightening and evidence-based answers relating to a specific research question (Boland, Cherry & Dickson, 2014). The goal of systematic reviews is to provide comprehensive overviews of the existing evidence on a specific research question by attempting to gather all the empirical evidence that fits pre-determined qualified criteria (Drucker, Fleming & Chan, 2016; Higgins, Thomas, Chandler, Cumpston, Li, Page, Welch and Cochrane Collaboration, 2019). The definition is further explained by Needleman (2002), who posits that systematic and explicit methods should be used to identify, select, critical appraisal, and summarise relevant research. Using a systematic review makes it possible to consolidate and collate large quantities of information (Petticrew & Roberts, 2006, p. 2), in this context, the unit of analysis, and diligently summarise factors for DM in KM. Transparency and reproducibility of search methodology are a requirement for a systematic review to limit systematic error (bias). A systematic review is guided by a set of scientific methods whose explicit aim is to counter bias (Kocher &

Riegelman, 2018; Petticrew & Roberts, 2006, p. 2). This is especially important because the study includes primary studies, with different study designs as a unit of analysis, requiring a method inclusive of case, exploratory, conceptual, and empirical studies. Therefore, the methodology is appropriate for this current study. A synthesis of findings of studies with different research designs can contribute to a broad understanding of the subject, unlike if the synthesis was of either qualitative or quantitative studies (Pluye & Hong, 2014).

In addition, Mulrow (1994) suggest that the logic for embarking on a systematic review includes decision-makers use systematic reviews as part of their estimation of the variables and outcomes included in their evaluations; a review is a scientific technique that is efficient, especially with cost and time; in a review, it is easier to establish the generalisability of scientific findings; the consistency of relationships can be assessed; it can explain data inconsistencies and conflicts in data; it has an advantage of increasing power through statistical use and increased precision, and provides an improved reflection of reality. The logic supports the application of systematic reviews in the sphere of the social sciences because it is an area that depends on evidence to respond to “unmet needs and the processes, implementation issues, and impacts of social programmes” (Littell, 2006). The medical and health areas have seen the benefits of systematic reviews because it is a method commonly used for investigations (Mallett et al., 2012). The social sciences are still using systematic reviews since the method is proving to be exceptional in summarising different topics and supporting decision-making based on the outcome of the reviews.

Systematic reviews are illustrious because it is a method that can pave the way for future research by highlighting predominant themes in research. This assertion is supported by Higgins et al. (2019), who posit that before embarking on new primary research, a systematic review should be conducted first. They further suggest that limitations are revealed in how previous studies were conducted in a systematic review, and that can be addressed in a new study or studies. Systematic Reviews are not perfect on any measure because they are subject to systematic and random error because of their retrospective nature and being observational studies (Dybå, Dingsøyr & Hanssen, 2007). Thus reviews must establish a protocol that is clear and understandable. A protocol ensures transparency and accountability for the research (Kocher & Riegelman, 2018).

There are four activities described in a systematic review that must take place, which are (1) shedding light on the question being asked; (2) establishing and reporting on the applicable research ('mapping' the research); (3) systematically and critically assessing of the research reports, so it brings the findings into a logical statement, known as synthesis together; and (4) set up what claims can be made from the research evidence (Gough, Oliver & Thomas, 2017, p. 4). With these activities, transparency becomes key in keeping the research aligned with the objectives set. Figure 3.1 shows the activities from the 2017 edition "*An introduction to systematic reviews.*"

The flow of the activities from 2017 can be illustrated as follows:

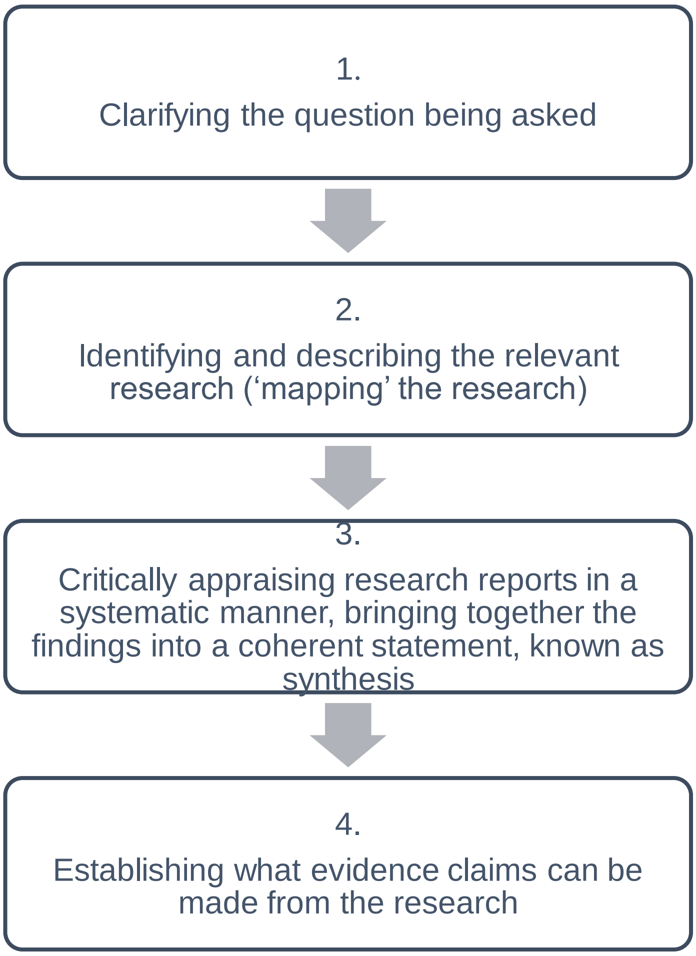

Figure 3. 1: Flow of activities

Source: Gough et al., 2017, p. 16

The current study follows the steps stated in the flow of activities, where the research questions can be answered in a logical manner. A systematic review protocol was followed to guide the process of the review and minimise bias by using the background and rationale of the study, the review question(s), the inclusion criteria, search strategy, critical appraisal, data collection, and data synthesis as guidelines (Pluye & Hong, 2014). According to Moher et al. (2015), to increase the transparency of the research process and reliability of published papers, reporting and publishing protocols is an important step. The review followed 18 applicable items from the PRISMA-P checklist to ensure the study does not deviate from the protocol and improve quality, completeness, and consistency. The PRISMA-P checklist includes 27 items that form part of the protocol. The items on the checklist are intended for the review to adhere to the guidelines. The current study adopted the PRISMA-P guideline for structure, transparency and thorough reporting (Appendix A) (Moher et al., 2015).

3.4 Data Collection

3.4.1 Sampling from Databases

In a systematic review, a unit of analysis is the previous studies (Littell & Corcoran, 2010), and for this current study, the unit of analysis are selected from databases. The number of databases depends on topic, time and resources but a minimum of two or more databases are necessary for a comprehensive search (Petticrew & Roberts, 2006, p. 102-103). After a comprehensive electronic search, four relevant databases were identified, namely ACM Digital library, Emerald Insight, Proquest and Science Direct. To ensure transparency and replicability, information retrieval specialists or librarians can help document the search process for researchers (Patrick, Demiris, Folk, Moxley, Mitchell & Tao, 2004). The selection process was arrived at in consultation with the librarian who specialises in data mining and knowledge management subjects (Littell & Corcoran, 2010). The four databases were carefully selected and used to balance the specific search and limit bias, where the aim of the study was to find published articles. The search strategy was developed in accordance with item 9 of the PRISMA-P checklist (Appendix A).

Databases offer excellent sources of academic and peer-reviewed research that is multi and interdisciplinary in coverage. In many instances, the indexing of the sources can date to the nineteen hundreds. The extensive coverage of the publications over the years makes a database more appealing for searching for articles. Many databases have criteria and guidelines for each content item, such as chapter, article, and report and ensure quality assurance. The databases also offer additional tools to make navigation, searching, and retrieval easier through built-in capabilities, such as the delimiters such as publication period, publication type, etc. The delimiters make it possible to specify the content that can be retrieved, such as using time frame, company or scholarly journals and specifying a language. Databases have the advantage of having built-in features that make searching efficient. The resources available in the databases reflect trends and developments in the various fields through graphs and offering related material or resources to what you are searching for. One of the most important reasons for selecting the specific databases was based on access, as they are provided through the university's licensing agreements. The databases needed to go through a screening stage, where they are first filtered by their subject coverage, especially in the fields of DM and KM. Additional criteria considered for selection was the high number of indexed journals that cover multi and interdisciplinary topics and fields, as this increases the retrieval of relevant content for the study (Petticrew & Roberts, 2006, p. 105). Table 3.1 illustrates the list of the electronic databases selected for searching for the unit of analysis.

Table 3. 1: List of electronic databases

Database	Details
ACM Digital library	The full name is Association of Computing Machinery and offers literature that covers computing and IT, in which there are over two million records available.
Emerald Insight	A scholarly publisher of academic journals and books in the fields of management, business, education, library studies, health care, and engineering.
Proquest	It is the largest full-text database that contains multidisciplinary and interdisciplinary literature, such as business, engineering, psychology, and technology, and has six billion digital pages.
Science Direct	A database with access to volumes of journals that cover subjects such as engineering, computer science, medicine, and chemistry and with over twelve million records.

Four selected electronic databases were in accordance with item 9 of the PRISMA-P checklist (Appendix A) (Moher et al., 2015).

3.4.2 Eligibility Criteria

In the current study, the electronic searches included publications between the timeframe of 2000 and 2017. By employing a systematic review, the researcher ensured that a thorough and complete search was conducted, with the time frame spanning over 17 years. As part of the search strategy, the researcher must formulate the eligibility criteria that inform the studies that will be included or excluded (Littell & Corcoran, 2010). The inclusion and exclusion criteria are specified in the systematic review protocol (item 8 of the PRISMA-P checklist in Appendix A). Criteria guide the researcher in identifying suitable studies. Still, it ensures that the selection is based on the criteria consistently, avoiding selection bias and not on the researcher's personal preferences for selection (Leeflang, 2014). Clearly defined inclusion and exclusion criteria provide clear limits so others can replicate or extend the review (Littell & Corcoran, 2010; Stern, Jordan & McArthur, 2014). Google Scholar was consulted as part of the preliminary and initial scouting of publication patterns on the topic. The platform could acclimate the researcher with publication trends and databases on the topic. In carrying out the review, the PRISMA Model was applied to reflect the unit of analysis used. PRISMA stands for Preferred Reporting Items for Systematic Reviews and Meta-Analyses, and it is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses and can provide the basis for reporting templates (Gough et al., 2017). The inclusion and exclusion for this review specified the individual criteria for each category below.

Inclusion Criteria:

<u>Publication date >2000 <2017</u>

The study's goal is to be comprehensive and highlight the theory, understanding, and adoption of DM in KM. The timeframe provides such an overview, and detailed information can be collected across time and within publications.

<u>Authors clearly stated</u>

The current study aims to identify the unit of analysis to make sure that every publication must have a clear and identifiable author to give the analysis credibility. It also enables detailed information about the publication to be identified, which builds towards capturing the data extraction process (Wang, Brito, Tsapas, Griebeler, Alahdab & Murad, 2015).

<u>Publication is in English</u>

Publications that were not in English were excluded from the group, which the researcher acknowledges that bias could have occurred by exclusively focusing on

English (Littell & Corcoran, 2010; Wang et al., 2015). The decision to enforce language criteria has much to do with resource availability.

Research question or purpose clearly stated

Clearly stated research questions or purpose gives a snapshot of how the study is carried out and how it can speak to the current study's research objective based on its goal for the outcome since the outcomes are imperative for theoretical or practical reasons (Littell & Corcoran, 2010). The variety of the different publications provided an opportunity to identify and evaluate the varying information.

Primary studies

Primary studies offer primary goals of a new investigation and provide rich information to reviews and establish a new and overall perspective. Especially so because the current study looked at the different study designs and includes different lengths of studies since "there is no single set of design criteria that reviewers can use for every systematic review" (Littell & Corcoran, 2010).

DM Techniques or tasks identifiable

Identifiable DM techniques and tasks inform the activity forming part of the outcome and can indicate which combination of techniques works.

KM processes or activities are evident

The evidence of KM processes and activities highlight which part of KM is focused on and to highlight to which task or technique in DM is this part of KM linked to.

The methodology stated within the study

A clearly stated methodology of the publications for a unit of analysis informs the activities to reach a particular outcome of the publications, which is necessary for the data extraction and provides rich details to form part of the data extraction and analysis.

Includes data mining contributing or relationship or integration to knowledge management.

The study tries to investigate DM in KM and its relationship, so every publication must reflect such in its content to support the development of answers to the objectives.

Availability of Full-text

The full-text availability allows for a thorough synthesis of each publication and for coding of the full-text, for developing coding and theme development.

The study can be properly cited

The full citation enables the data extraction to take place.

Exclusion Criteria:

The exclusion criteria depended on the publisher not being identifiable, the publication date is unknown, author(s) are anonymous, and it is not a primary study but a review. In addition to the criteria, DM and KM relationship are not identifiable in the publication.

Table 3.2 shows the inclusion and exclusion criteria applied to the publications to reach the final evidence for the systematic review.

Table 3. 2: Inclusion and Exclusion Criteria

Inclusion	Exclusion
Publication date >2000 <2017	Publisher not identified
Authors clearly stated	Publication date is unknown
Publication is in English	Author(s) are anonymous
Research question or purpose stated	It is a review
Primary studies	The study is not in the English language
DM Techniques and tasks identifiable	DM and KM relationship are not identifiable
KM processes or activities are evident	
The methodology stated within the study	
Interest includes DM contributing or relationship or integration to knowledge management	
Availability of Full-text articles	
The study can be properly cited	

The criteria ensured that the search was aligned throughout the process of searching.

3.4.2 Search Strategy

Comprehensive electronic searches were carried out on four databases selected as the sources for the unit of analysis, ACM Digital library, Emerald Insight, Proquest and Science Direct. The four databases helped highlight the suitability of the search strategy and helped with the specificity of the search (Leeflang, 2014). When framing the search query, the essential keywords in the research objectives were developed and extracted to connect DM and KM in an applicable context when searching in an electronic database. The following key search terms were included in the query:

“Data Mining” AND “Knowledge Management” AND (Integration OR Inclusion OR Incorporation) AND (Relationship OR Connection) AND Development.

To further get clarity on the keywords, they are defined as a part of making sense of their meaning. The Online Oxford English Dictionary defined the keywords for the topic and search query as:

- *Relationship* - “The state or fact of being related; the way in which two things are connected; a connection, an association.”
- *Developments* - “The action or process of bringing something to a fuller or more advanced condition; the explanation or elaboration of an idea, theory.”
- *Integration* - “The making up or composition of a whole by adding together or combining the separate parts or elements; combination into an integral whole: a making whole or entire.”
- *Contribution* - “The action of contributing or giving as one's part to a common fund or stock; the action of lending aid or agency to bring about a result.”

When searching for the publications to be reviewed, search techniques and strings were employed, including synonyms and using the Boolean Operators (AND, OR and NOT), then combining the terms, which retrieved the most relevant studies (Petticrew & Roberts, 2006, p. 81). The search strategy was in accordance with item 10 of the PRISMA-P checklist (Appendix A) (Moher et al., 2015).

A combination of key search terms was applied consistently across all four databases. The search scope had two stages of searching, where a preliminary search was conducted to scope out what was available in the databases. Then a secondary search level was performed to limit and identify specific publications. The scope only focused on research articles, highlighting the appropriateness of databases offered by the university and their wide range and broad reach of sources. While searching the databases, the main keywords used were “*data mining and knowledge management and relationship, or integration, or developments,"* and the time frame was set between 2000 and 2017.

When developing a search strategy, it should be broad enough to allow as much available evidence to be found (Leeflang, 2014). In line with broadening the search, the review was not specific to any industry to give way for as many multi and interdisciplinary studies as possible to be retrieved. In broadening the search in terms of excluding industry, it gave way for the evidence to show that the processes can be applied in many scenarios and different organisational settings.

A total of 5 222 hits were obtained based on the specifications of the search strategy. Based on the individual retrievals, Emerald Insight: 500; ProQuest: 2779; Science Direct: 1102; and ACM Digital Library: 841. Figure 3.2 illustrates the hits from each one of the specific databases.

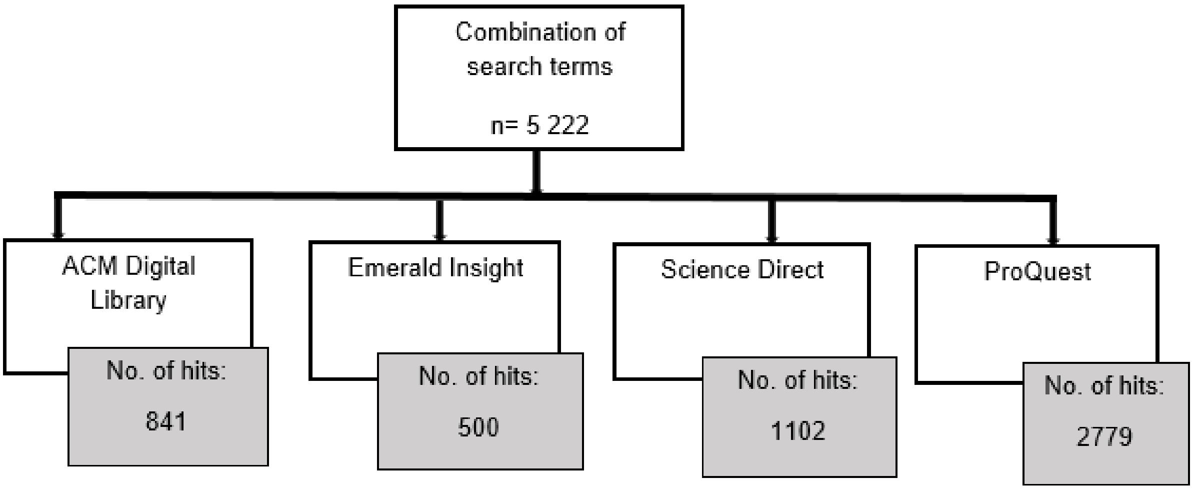

Figure 3. 2: The number of publications obtained from databases

The retrieved references were exported into EndNote 20, which is a reference management software. The EndNote library created with the 5 222 references was then transferred into the Distiller SR software to conduct the study selection process. The Distiller SR software is discussed in the Data Management section of this chapter.

3.4.3 Study Selection

While determining which studies were relevant for the final review, they were appraised on three levels using the PRISMA-P statement. The identification of the final review articles followed the stages of the study selection phase in the steps for a systematic review. The stages are as follows: Stage 1 - Identify relevant studies – search databases and conference proceedings; Stage 2 - Exclude studies based on titles; Stage 3 - Exclude studies based on abstracts; Stage 4 - Obtain primary papers and critically appraise studies. Each of the four databases was searched, and each study was scanned; then meaningful data was extracted, which was informed by the PRISMA-P guideline (Moher et al., 2015).

Level 1: Before the screening of each study commenced, duplicates were identified and discarded n= 125. The first level involved screening the studies based on the title, abstract, the English language and with identifiable authors (n=5097). The studies that were unrelated because of the title and abstract were excluded in this level (n=4754).

Level 2: The remaining studies (n=343) were moved to level 2, where they were screened based on the availability of full-texts to determine the appropriateness for inclusion (Moher et al., 2015). The studies that did not have full-texts, were not in the English language, the author was not clearly stated, and could not be properly cited were excluded from the review (n=90), in accordance with the PRISMA-P statement.

Level 3: The remaining studies were further assessed on level 3 using a set of questions that a researcher can manually add into a filtering form on Distiller SR to determine further consideration for inclusion. The questions were based on the criteria developed for inclusion (Table 3), which required thorough scanning of the publications. From level 3, additional studies were excluded (n=199), and the remaining studies included in the review were n=54.

The article selection process of this systematic review is summerised using the PRISMA flow diagram generated by the DistillerSR software. The PRISMA flow diagram allows for the documenting and managing of the searches systematically. The diagram provides visualisation of the steps taken to reach the final selection of articles used in the review. Figure 3.3 illustrates the PRISMA flow diagram showing the process of identifying articles for review, adapted from PRISMA-P statement and Evidence Partners.

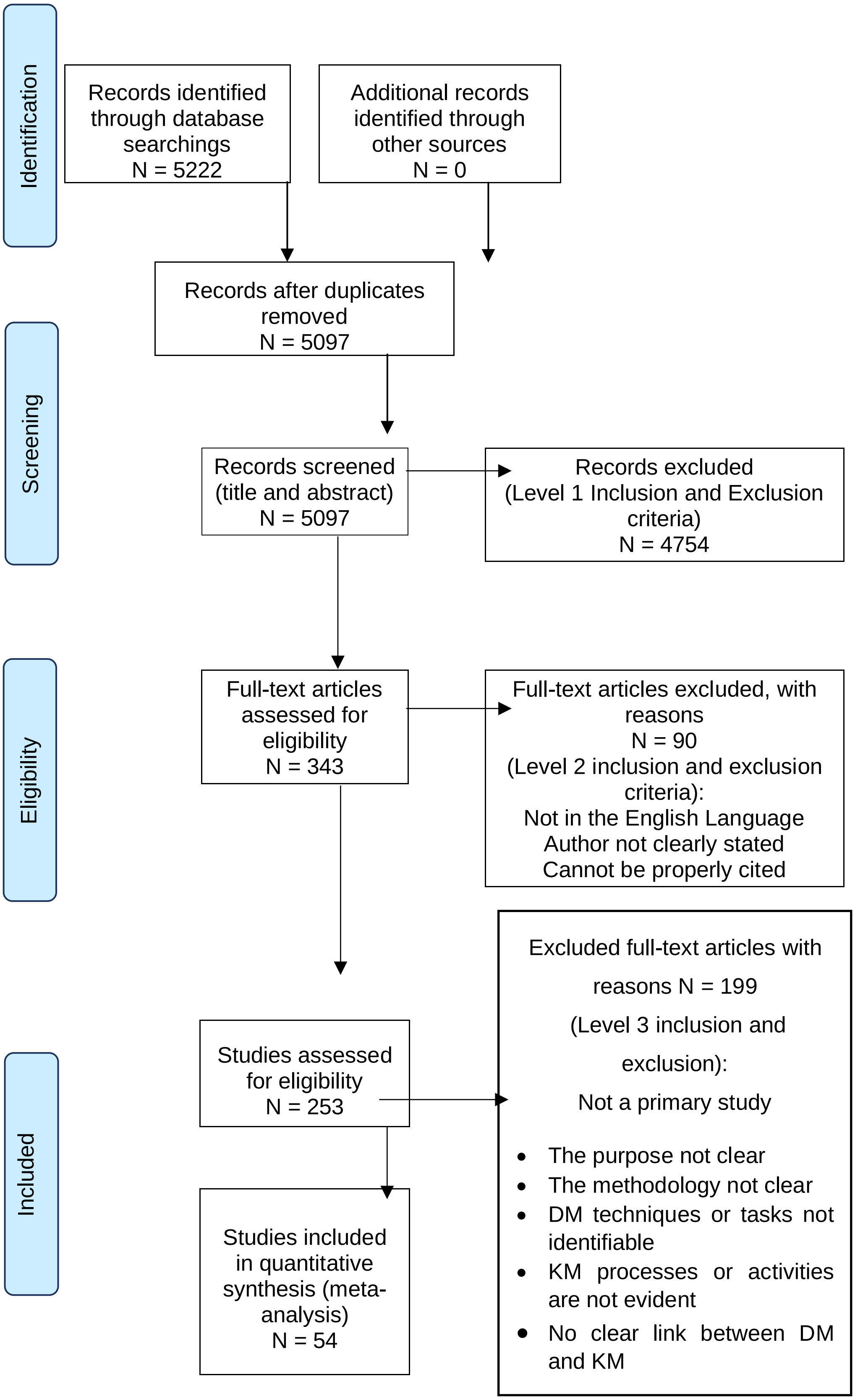

Figure 3. 3: PRISMA Flow Diagram Showing The Process of Identifying Articles for Review

Source: adapted from PRISMA-P Statement and Evidence Partners

The process revealed 54 relevant articles eligible for inclusion as the unit of analysis for the study. The study selection was in accordance with item 11b of the PRISMA-P checklist (Moher et al., 2015). The 54 references were exported into EndNote 20, from the DistillerSR programme.

3.4.4 Data Management

The systematic review makes it possible to review the publication gathering process at each stage. During the process, the results from the four specific databases were collated through the software DistillerSR, whose sole goal supports systematic reviews and literature reviews. The software is among the most popular software in the world, with over 250 members, organisations and higher education institutions using it. DistillerSR improves accuracy and transparency for all review sizes (Evidence Partners, 2021). The software further assesses the risk of bias and extraction of study aspects and outcomes by following its project management process of searching, screening, full-text retrieval, data extraction, and reporting and auditing. The process involved transferring selected references from databases into EndNote 20, then into Distiller SR. In Distiller SR, the references were evaluated on three levels until the final batches of a unit of analysis for the study were identified. The Data Management was in accordance with item 11a of the PRISMA-P checklist (Moher et al., 2015).

3.4.5 Data Extraction and Risk of bias

3.4.5.1 Data Extraction

Primary studies can be filled with bias, from which systematic reviews may also be susceptible, so each needs to be critically appraised (Drucker et al., 2016). According to Harrison, Jones, Gardner and Lawton (2021), no appraisal tool is available for mixed- and multi-methods work with study heterogeneity. They further argue that a dual approach does make it difficult for the appraisal and creates challenges in appraising aspects of multi-methods work. The articles included as a unit of analysis were critically appraised using the TAPUPAS framework. According to Ma, Wang, Yang, Huang, Weng and Zeng (2020), assessment tools for qualitative studies are fewer compared to the tools used in quantitative studies, and there are three types of tools that can be used, scales, checklist and items. For this systematic review, the studies were included because they are primary studies. Therefore, an item list was identified as suitable (Zeng et al., 2015).

According to Littell and Corcoran (2010), data extraction forms need not be created from scratch. However, researchers can adapt existing forms used in other systematic reviews by tailoring one or more to fit the topics to be addressed in a new review, the specific purposes and questions. Therefore, the current review settled on the TAPUPAS framework because of its application and use in social work (Pawson, Boaz, Grayson, Long & Barnes, 2003), which can be adapted to the current study a set of key questions by Orme and Shemmings (2010). The two frameworks were consulted in framing the appropriate set of boundaries and acceptance of the unit of analysis based on the objective of the study. The TAPUPAS framework devised by Pawson et al. (2003) provide the thinking space to assess the quality of the articles based on the different criteria in the framework:

- Transparency – can it be scrutinised and challenged?
- Accuracy – is the research grounded?
- Purposivity – does it meet its purpose?
- Utility – can it be used?
- Propriety – was it carried out ethically and legally?
- Accessibility – is it understandable?
- Specificity – are the source-specific standards met?

Further support in developing the criteria is the guideline by Orme and Shemmings (2010), where a set of 'key questions' are suggested that should be asked as part of a critical appraisal of social work research. The questions are as follows:

- How relevant is the study to the review question?
- How much contribution is there from the information?
- Are the findings trustworthy?
- Are the findings generalisable?
- Was it ethically conducted?

The two frameworks as item lists are powerful in contextualising the author's point through their selection of publications. The item list allows for quality assurance of the studies by using a framework with seven criteria and questions and configured to be judged by just a tick as a symbol for acknowledging an item presence. The questions by Orme and Shemmings (2010) were used to develop the DistillerSR assessment form in identifying eligible articles for a unit of analysis and providing preliminary quality checks before the appraisal item list, where each question was judged by "Yes," "No" or "Unclear." Articles that had “Yes” were automatically included, the ones with a “No” were excluded, and the “Unclear” articles were sanctioned for assessment — the criteria assisted in assessing the trustworthiness, relevance and results of published papers. In addition, the checklist aids in the transparency and reproducibility of the process followed in the study (Aromataris & Pearson, 2014). The primary goal of the review was to retain publications reporting on DM in KM. The method made it possible to further assess whether the publications selected in the selection process were suitable for the study or not. The checklist provides support in establishing the quality of the articles that contribute to the study. Table 3.3 depicts the item list used to appraise the articles, in which a tick can indicate which quality is present in the article.

Table 3. 3: Description of appraisal checklist

Article	Transparency	Accuracy	Purposivity	Utility	Propriety	Accessibility	Specificity
D1	√	√	√	√	√	√	√
D2	√	√	√	√	√	√	√

The 54 publications were appraised using the list, and no inconsistency was recorded because of the preliminary quality checks applied in the DistillerSR programme. According to Littell and Corcoran (2010), there are no stable “quality” or bias measures. Therefore, including and keeping the design quality separate is important. The extracted data was exported into MS Excel and summarised (Appendix B). The summarisation of the data was in accordance with items 11 c and 12 of the PRISMA-P checklist (Appendix A) (Moher et al., 2015). It included the following: the author, publication title, year of publication, design (qualitative, quantitative or mixed methods), DM task, KM activity, and findings.

3.4.5.2 Risk of bias

Bias refers to the selective communication of research findings based on the preferred result instead of their true value (Drucker etal., 2016). Bias can lead to underestimating or over exaggerating the results of a study (Aromataris & Pearson, 2014). Therefore, as part of minimising bias, it is necessary to put in place precautions that respond to such possibilities. The risk of bias was reduced by developing and applying a systematic review protocol, which guided the entire review. Having the protocol in place deters bias since it contextualises the research question, including the participants and outcomes, and describes the methodology thoroughly to allow replication by others (Drucker et al., 2016).

A protocol developed after the PRISMA-P statement provides key content, such as describing the review objectives and questions, the search strategies and data sources, eligibility criteria, study screening method and selection, outcomes, data extraction, and analyses (Shamseer et al., 2015). The key content was defined in the study by using the PRISMA-P checklist (Appendix A) (Moher et al., 2015). Four databases were searched, utilising multiple search terms, to obtain an expansive number of articles to reduce publication bias (Drucker et al., 2016). With the inclusion and exclusion criteria available, it prevented the researcher from selecting articles based on their preferences, thus reducing selection bias. To evaluate the risk of bias every article selected for the review was individually analysed. This study included various studies with different quality levels, designs, and outcomes to reduce bias. The risk of bias was in accordance with item 14 of the PRISMA-P checklist (Appendix A) (Moher et al., 2015).

3.4.5.3 Publication Bundling List of Literature Sources:

The review articles were grouped to identify their publishers to identify initial trends. The list shows the number of publications from the different publishers. Table 3.4 illustrates publication by the publisher.

The articles were further grouped in a list with their titles, authors, source title and year of publication. Table 3.5 shows the list of individual articles listed in the table.

Table 3. 4: List of journals used in the study

Journal Title	Number of journal articles in study
Industrial Management & Data Systems	4
Internet Research	2
Journal of Money Laundering Control	1
Social Responsibility Journal	1
The Learning Organisation	1
Business Systems Research	1
Communications of the Association for Information Systems	1
Computer Science	1
Decision Support Systems	4
E-Service Journal	1
Expert Systems with Applications	3
IBM Systems Journal	1
Industrial Management & Data Systems	3
Information Technology & Tourism	1
International Journal of Computer Science Issues (IJCSI)	1
International Journal of Innovation and Knowledge Management in the Middle East and North Africa	1
Journal of Computer Information Systems	1
Journal of Economic Development, Management, IT, Finance and Marketing	1
Journal of Enterprise Information Management	1
Journal of Hospitality and Tourism Technology	1
Journal of International Technology and Information Management	1
Journal of Knowledge Management	10
Journal of Manufacturing Technology Management	1
Journal of Systems and Information Technology	2
Knowledge and Process Management	1
Knowledge-Based Systems	1
Procedia Computer Science	1
The Journal of Strategic Information Systems	1
The Knowledge Engineering Review	1
VINE	4

The study found 54 publications within 30 journal titles, covering the fields of management sciences, technology, hospitality and tourism and computer sciences.

Table 3. 5: List of articles with journal titles and year of publication

Article Title	Author	Journal Title	Year Published
A framework of intelligent decision support system for Indian police	Gupta M., Chandra B., Gupta M.P.	Journal of Enterprise Information Management	2014
A knowledge management approach to data mining process for business intelligence	Wang, H. and Wang, S.	Industrial Management & Data Systems	2008
A knowledge management framework using business intelligence solutions	Gadu, M. and El-Khameesy, N.	International Journal of Computer Science Issues (IJCSI)	2014
A model of Information Technology opportunities for facilitating the practice of knowledge management	Wild, R. and Griggs, K.	VINE	2008
An enterprise-wide knowledge management system infrastructure	Lee, S.M. and Hong, S.	Industrial Management & Data Systems	2002
An integrated electronic medical record system (iEMRS) with decision support capability in medical prescription	Ting, S.L., Ip, W.H., Tsang, A.H.C. and Ho, G.T.S.	Journal of Systems and Information Technology	2012
An Intelligent Decision-Making Architecture for Banks: Business Intelligence And Knowledge Management Systems Integration	Rao, G. K. and Dey, S.	Journal of Economic Development, Management, IT, Finance and Marketing	2012
An ontology-based business intelligence application in a financial knowledge management system	Cheng, H., Lu, Y-C and Sheu, C.	Expert Systems with Applications	2009
Application of knowledge management technology in customer relationship management	Bose, R. and Sugumaran, V.	Knowledge and Process Management	2003
Artificial Intelligence in Knowledge Management: Overview and Trends	Birzniece, Ilze	Computer Science	2011
Automated knowledge discovery in advanced knowledge management	Grobelnik, M. and Mladenić, D.	Journal of Knowledge Management	2005
Automating knowledge acquisition for constraint-based product configuration	Huang, Y., Liu, H., Keong Ng, W., Lu, W., Song, B. and Li, X.	Journal of Manufacturing Technology Management	2008

Article Title	Author	Journal Title	Year Published
Big data systems: knowledge transfer or intelligence insights?	Rothberg, H.N. and Erickson, G.S.	Journal of Knowledge Management	2017
Business intelligence for cross-process knowledge extraction at tourism destinations	Höpken, W., Fuchs, M., Keil, D. and Lexhagen, M.	Information Technology & Tourism	2015
Converting computer-integrated manufacturing into an intelligent information system by combining CIM with concurrent engineering and knowledge management	Prasad, B.	Industrial Management & Data Systems	2000
Data Mining as Support to Knowledge Management in Marketing	Zekić-Sušac, M. and Has, A.	Business Systems Research	2015
Data mining for building knowledge bases: Techniques, architectures and applications	Krzywicki, A., Wobcke, W., Bain, M., Calvo Martinez, J., & Compton, P.	The Knowledge Engineering Review	2016
Data mining in anti-money laundering field	Yasaka, N.	Journal of Money Laundering Control	2017
Data mining's capabilities for knowledge creation in the GCC counties	Wahab, A. and Rasha, S.	International Journal of Innovation and Knowledge Management in the Middle East and North Africa	2012
Data-driven decision-making for the enterprise: an overview of business intelligence applications	Hedgebeth, D.	VINE	2007
Decision Support System and Knowledge-based Strategic Management	Alyoubi, Bader A.	Procedia Computer Science	2015
Development of an intelligent e-healthcare system for the domestic care industry	Wong, B., Ho, G.T.S. and Tsui, E.	Industrial Management & Data Systems	2017
Eight questions for customer knowledge management in e-business	Rowley, J.	Journal of Knowledge Management	2002
Enhancing business networks using social network based virtual communities	Lea, B., Yu, W., Maguluru, N. and Nichols, M.	Industrial Management & Data Systems	2006

Article Title	Author	Journal Title	Year Published
Extensible markup language and knowledge management	Otto, J.R., Cook, J.H. and Chung, Q.B.	Journal of Knowledge Management	2001
Facilitating knowledge management through information technology in hospitality organisations	Okumus, F.	Journal of Hospitality and Tourism Technology	2013
Gaining customer knowledge through analytical CRM	Xu, M. and Walton, J.	Industrial Management & Data Systems	2005
How the Internet of Things can help knowledge management: a case study from the automotive domain	Uden, L. and He, W.	Journal of Knowledge Management	2017
Improving Knowledge Management by Integrating Hei Process and Data Models	Natek, S. and Lesjak, D.	Journal of Computer Information Systems	2013
Information and reformation in KM systems: big data and strategic decision-making	Intezari, A. and Gressel, S.	Journal of Knowledge Management	2017
Integrating web-based data mining tools with business models for knowledge management	Heinrichs, John H. and Lim, J-S.	Decision Support Systems	2003
Intelligent Knowledge Beyond Data Mining: Influences of Habitual Domains	Yu, Xiaodan; Shi, Yong; Zhang, Lingling; Nie, Guangli; and Huang, Anqiang	Communications of the Association for Information Systems	2014
Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector	Sumbal, M.S., Tsui, E. and See-to, E.W.K.	Journal of Knowledge Management	2017
Knowledge elicitation and mapping in the design of a decision support system for the evaluation of suppliers' competencies	Cannavacciuolo, L., Iandoli, L., Ponsiglione, C. and Zollo, G.	VINE	2015
Knowledge elicitation approach in enhancing tacit knowledge sharing	Ting, S.L., Wang, W.M., Tse, Y.K. and Ip, W.H.	Industrial Management & Data Systems	2011
Knowledge integration in organisations: an empirical assessment	Kenney, J.L. and Gudergan, S.P.	Journal of Knowledge Management	2006
Knowledge management and data mining for marketing	Shaw, M. J., Subramaniam, C., Tan, Gek Woo and Welge, M. E.	Decision Support Systems	2001

Article Title	Author	Journal Title	Year Published
Knowledge management and its link to artificial intelligence	Liebowitz, J.	Expert Systems with Applications	2001
Knowledge management for the analysis of complex experimentation	Maule, R., Schacher, G. and Gallup, S.	Internet Research	2002
Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing	Nemati, H. R., Steiger, D. M., Iyer, L. S. and Herschel, R. T.	Decision Support Systems	2002
Knowledge4Scrum, a novel knowledge management tool for agile distributed teams	Sungkur, K. R. and Ramasawmy, M.	VINE	2014
Knowledge-collector agents: Applying intelligent agents in marketing decisions with knowledge management approach	Moradi, M., Aghaie, A. and Hosseini, M.	Knowledge-Based Systems	2013
Linking key figures and internet business news for personalised management information	Meier, M. and Mertens, P.	Journal of Systems and Information Technology	2000
Managing extracted knowledge from big social media data for business decision-making	He, W., Wang, F.-K. and Akula, V.	Journal of Knowledge Management	2017
Metadata as a knowledge management tool: supporting intelligent agent and end user access to spatial data	West, L. A. and Hess, T. J.	Decision Support Systems	2002
Project-based knowledge maps: combining project mining and XML-enabled topic maps	Liu, D. and Hsu, C.	Internet Research	2004
Relevance of data mining for accounting: social implications	Mraović, B.	Social Responsibility Journal	2008
Role of knowledge management and analytical CRM in business: data mining based framework	Ranjan, J. and Bhatnagar, V.	The Learning Organisation	2011
Student data mining solution–knowledge management system related to higher education institutions	Natek, S. and Zwilling, M.	Expert Systems with Applications	2014
The concepts of big data applied in personal knowledge management	Liu, C.-H., Wang, J.S. and Lin, C.-W.	Journal of Knowledge Management	2017

Article Title	Author	Journal Title	Year Published
The integration of business intelligence and knowledge management	Cody, W. F., Kreulen, J. T., Krishna, V. and Spangler, W. S.	IBM Systems Journal	2002
The Interaction of Business Intelligence and Knowledge Management in Organisational Decision-Making	Vinekar, V., Teng, J.T.C. and Chennamaneni, A.	Journal of International Technology and Information Management	2009
The role of AI-based technology in support of the knowledge management value activity cycle	Fowler, A.	The Journal of Strategic Information Systems	2000
Web Service for Knowledge Management in E-Marketplaces	Singh, R., Iyer, L., & Salam, A. F.	E-Service Journal	2003

The articles were then tabled, reflecting the knowledge that different studies focused on during their investigation. The table also illustrates the types of DM techniques and models applied to the cases. This aided in understanding which techniques worked for the kind of knowledge identification and showed the techniques used in finding patterns in the rich data collected throughout the activities in an organisation. Data cannot be "understood" without making sense of it first. It needs analysis through models and algorithms to create knowledge and then to understand the environment (Smith, 2020). Table 3.6 shows an excerpt from the list of the types of knowledge and the techniques used, where applicable (Appendix B).

Table 3. 6: Excerpt of types of knowledge and techniques applied

Title	Author	Types of knowledge and Techniques
A framework of intelligent decision support system for Indian police	Gupta M., Chandra B., Gupta M.P. (2014)	• Knowledge Acquisition; • Clustering, Classification and Association Rules.
A knowledge management approach to data mining process for business intelligence	Wang, H. and Wang, S. (2008)	• Human Knowledge • DM Centred Life Cycle
A knowledge management framework using business intelligence solutions	Gadu, M. and El-Khameesy, N.(2014)	• Knowledge sharing • Predictive and descriptive DM

The publications were categorised according to their title, author, journal title and date of publication, as presented in Table 3.6. The extensive form with the necessary information about each article is available in Appendix B. The purpose was to identify the pattern of interest in the field. The articles were further arranged according to how many publications were published each year. Table 3.7 reflects the publication rates per year.

Table 3. 7: Publication Rate between 2000 and 2017

Year	Number of publications
2000	3
2001	3
2002	6
2003	3
2004	1
2005	2
2006	2
2007	1
2008	4
2009	2
2011	3
2012	3
2013	3
2014	5
2015	4
2016	1
2017	8

The publications were added to an excel spreadsheet to generate a trend graph across the study period. The graph in Figure 3.4 reveals the rate of publication, with a rise in publications.

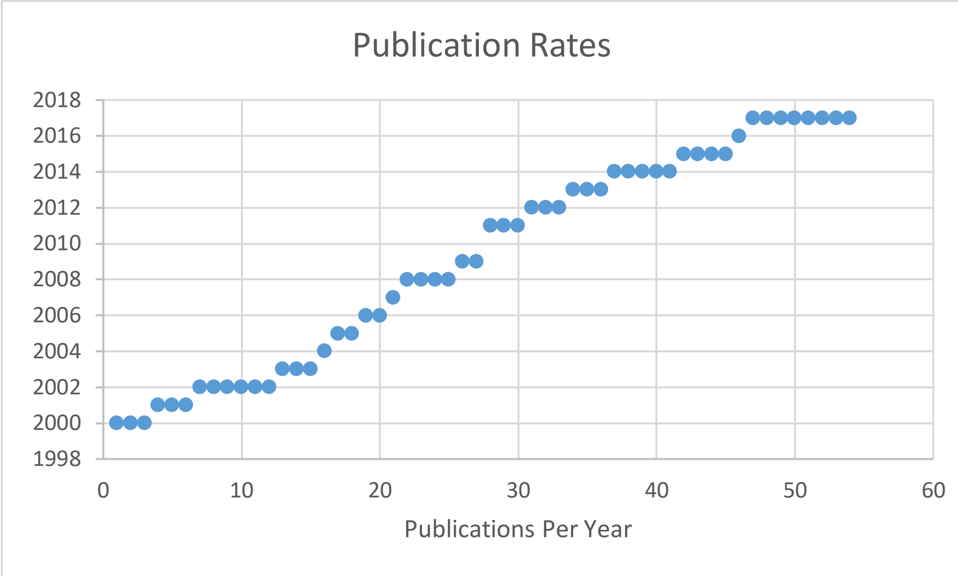

Figure 3. 4: Graph of the publication years

3.5 Data Analysis

Once the articles were appraised, thematic analysis was employed to analyse the data and group the patterns found regarding DM in KM. Thematic analysis is often referred to as 'data analysis techniques in the social sciences.' It is a method used in qualitative data to identify themes and have an approach flexible to coding and theme development (Terry, Hayfield, Clarke & Braun, 2017, p. 17). Thematic analysis can support the extraction, organisation and synthesis of data from different study designs (qualitative, quantitative and mixed methods). It can identify the information by identifying specific words and word patterns in the titles and abstracts of the articles. Thematic analysis can provide insight by identifying recognisable reoccurring topics and ideas or patterns within the data and throughout a variety of texts (Allen, 2017, p. 1; Hawkins, 2017).

The 54 articles went through a process of individual assessment, and critical themes emerged from the publications. Analytical themes were identified from the critical themes by applying thematic analysis. According to Braun and Clarke (2006), there are six steps to be followed when using thematic analysis, including:

- Familiarising yourself with your data.
- Generating initial codes.
- Searching for themes.

- Reviewing themes.
- Defining and naming themes.
- Producing the report.

The thematic analysis approach was used in the study to analyse the articles found as part of the review, to highlight and reflect on DM in KM. The analysis made it possible to code each article to provide a broader view of what and how each article approached their investigation and generated themes from the codes. For the organisation and interpretation of a large body of data, most qualitative researchers use computer software such as Atlas.ti, Nvivo, MAXQDA, etc. (Leedy, Ormrod & Johnson, 2021, p. 392; Silver & Lewins, 2014; Zakaria & Zakari, 2016). For this study, the software Atlas.ti 9 was employed to go through each article on the study to contextualise their investigations, results, and coding the data. This programme is used mostly for the qualitative analysis of large bodies of textual, graphical, audio and video data, making it highly suitable for contextualising the evidence (Scientific Software Development GmbH, 2021). When using Atlas.Ti 9 to code, sometimes the data highlighted for analyses do not contain explicit meaning. When that happens, "the researcher must interpret or go beyond the stated words to the contextual, or latent, meaning" (Zakaria & Zakari, 2016).

The researcher acknowledges that because there was only one coder, the researcher, without another coder, has introduced bias in inter-rater reliability. The literature shows that "the value of independent double selection over selection by one author has not been investigated" (Leeflang, 2014). When only one coder conducts data coding, the coder is susceptible to bias and data coding can be influenced thereof (Leedy et al., 2021, p. 383). For transparency and to counter bias, the coding of data and the development of codes were agreed upon and approved by the expert. In the context of this study, the expert was the supervisor. When there were ambiguities about coding the data, motivation and definitions were discussed with the expert for clarity. The lists of codes and corresponding groups are attached in Appendix C.

According to Gough et al. (2017), Thematic Synthesis has three stages: the first is Coding Text; stage 2 is developing descriptive themes, and stage 3 is generating analytical themes. Stage 1 analysed and scrutinised each article, line by line, to create codes; stage 2 involved creating and developing the descriptive themes; and stage 3 involved developing analytical themes. According to Terry et al. (2017, p. 17), when coding, most researchers begin with a fundamental descriptive level and systematically work upwards towards a more interpretative level. This occurs when following the three stages of coding and when employing the inductive approach.

Stage 1

Each article went through inductive analyses and scrutiny. The researcher used the Atlas.Ti 9 programme to go through it line by line to identify possible codes. Inductive coding can lead to identifying and developing new concepts and theories (Silver & Lewins, 2014). Coding refers to the systematic application of markers, words or short phrases that represent and summarise critical features of studies included in a review (Gough et al., 2017, p. 124). To find the 'evidence' for the themes, coding must first occur (Terry et al., 2017, p. 19).

The programme highlights the section of interest for coding and the surrounding texts on that code (Silver & Lewins, 2014; Zakaria & Zakari, 2016). This provides context to the codes generated and can provide an understanding of why and where that code is used in the article. Figure 3.5 provides an example of a coded section of one article used in the study.

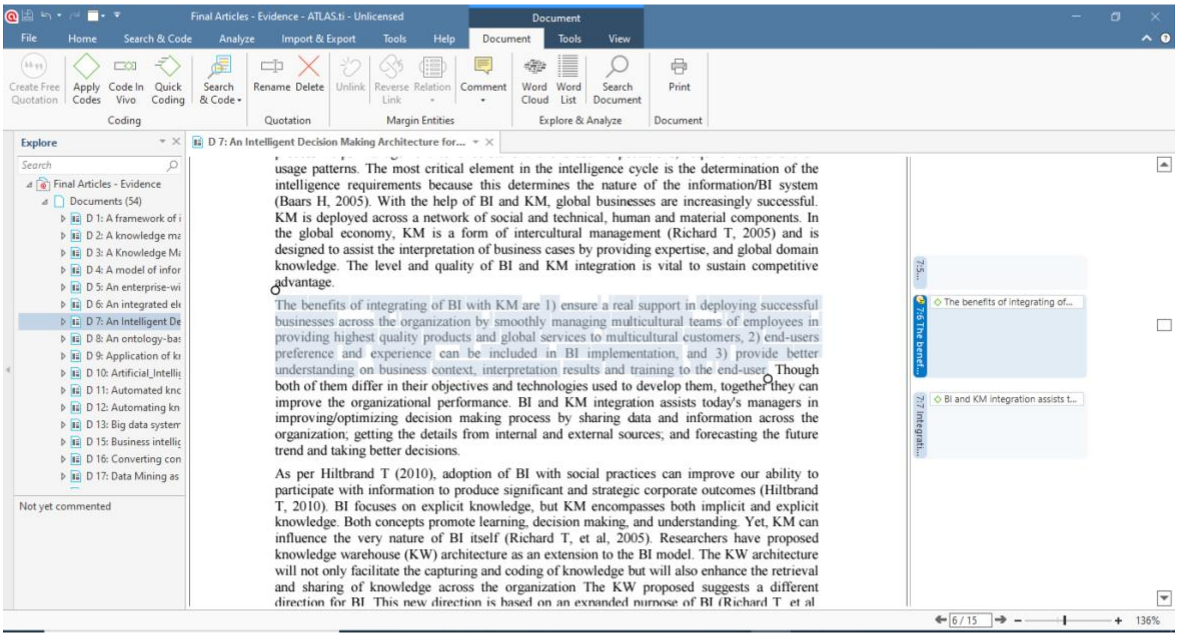

Figure 3. 5: Thematic coding of an article text

The approach used in coding the articles of this study was based on the open coding strategy, which allowed the researcher to code the data with an open mind to identify potential meaning and categories and then adopt axial coding (Leedy et al., 2021, p. 386). They describe axial coding as core categories becoming apparent and central to the phenomenon being investigated, which reflects the context of the core category, enabling and influencing conditions, strategies and consequences (Silver & Lewins, 2014; Zakaria & Zakari, 2016). The coding process included creating codes grounded in the data, defining what part of the text is to be coded, paying attention to co-occurring codes to see what patterns appear, defining and listing the meaning, scope and application of codes. The process was followed on the backdrop of the following strategies (Leedy et al., 2021, p. 386; Silver & Lewins, 2014):

- Following a structure reflecting the theoretical context as identified by the literature review.
- Labelling relevant segments of interest from an inclusive and open perspective and create consistency, and reliability in the categorisation of the data.
- Considering overlapping ideas and how they can be merged into a single code.

- Precaution against being too item-specific and considering themes that appear to be important in the overall data set.

Coding the articles was guided by the article's perspective on DM and KM activities, the approach and motivation and the findings and discussions. The codes were generated based on the concepts, themes and terminology. Throughout the use of Atlas.ti 9, codes were generated using the *in vivo* coding tool, which creates codes by using the entire highlighted quotation as a code name itself (Zakaria & Zakari, 2016). The tool was intuitive and easy to follow, and the researcher could rename the codes when appropriate. Atlas.ti 9 enables generating reports from the different activities during coding, which can further help in refining the codes.

Each of the 54 articles was carefully assessed and analysed line by line, and coding of the entire data sets in completion occurred (Terry et al., 2017, p. 23). After the coding process, 187 codes were generated. The research question was kept in mind during the article assessment to ensure that the codes highlighted and reflected the objectives of the current study. The generated list of codes adequately reflects the diverse and varied patterning of the meaning in the set of data. The list was shared and discussed with the expert.

Stage 2

In thematic analysis, there is systematic coding of the identified data. They are arranged based on their characteristics, then into different categories that may lead to some structure of codes that can provide themes and sub-themes (Nowell, Norris, White & Moules, 2017). Themes are not conjured up in the minds of researchers, nor are they anticipated early on, as they do not drive analytic direction. They are developed from coding and working with both the data and codes, rather than "pre-existing the coding process" (Terry et al., 2017, p. 20). They come from the analytic process and are assumed to be subjective and interpretative processes.

In using Atlas.ti 9, there were 187 codes generated, which can be challenging administratively. Several challenges come with using a software coding tool, such as Atlas.ti. These challenges can range from seeing coding as a replacement analysis instead of the researcher conducting the analysis, too many codes being generated, not discerning when to stop coding, and not defining the codes well enough (Silver & Lewins, 2014). As a precaution, the current study conducted a coding cleanup, which involved either deleting, merging and/or renaming codes for an organisation to refine and strengthen the codes (Leedy et al., 2021, p. 383).

The 187 codes were individually assessed, and the differences and similarities between the codes were identified and grouped. New codes were created from the group of initial codes to highlight the meaning (Thomas & Harden, 2008). Once the coding was completed, the next step was to cluster the codes into themes or narratives that require a relevant theory for interpretation (O'Connor & Joffe, 2020). The codes were grouped before developing descriptive and analytical themes. Group codes were generated based on the surrounding text or quotation to determine the meaning, topic and theme. The grouping of the codes was developed according to categories that represented an umbrella concept. Figure 3.6 shows an excerpt of the codes from Stage 2 grouped by the software based on the groups created (Appendix C).

	Code	Code Groups
○	A data warehouse should always provide its users with accurate, consi	Data warehouse
○	A DM result might not trigger an action, but can be learned by busine	DM result to knowledge
○	A knowledge repository of previous data mining efforts could be used	Knowledge repository
○	Aggregation of human knowledge and feedback into KB construction seems	Automation
○	AI-based technology, on its own, does not provide a unique solution t	Knowledge acquisition for Expertise
○	Amarket basket analysisB w x 3 and gives us the relationship between	Customer relationship management
○	An action must have its outcome. An action outcome is the assessment	Measurement
○	An enterprise-wide knowledge management system infrastructure	Enterprise wide knowledge
○	are typically non-technology oriented.	Non technology driven
○	As the other drawbacks of ES towards KM the following issues are nam	Expert system
○	As thousands of knowledge sets gradually piled up in the knowledge da	Knowledge base
○	as with any KM-related activity, support from top management is neces	Management support
○	Association rules There are condition-outcome relationships among att	DM tasks
○	Besides the signiffcant of the approach in product conffguration the	Design knowledge

Figure 3. 6: Codes and Code Groups

From the grouping process, 26 groups were developed from the code groups. The groups on Atlas.ti 9 code grouping were networked and highlighted within the group to show the relationships between the ideas to contextualise the reason for grouping them.

There is a general acceptance that different approaches are used by different analysts, where they apply different ways to organise codes into themes (Armstrong, Gosling, Weinman & Marteau, 1997). There is no one way of developing themes, and just because your data set appears consistently, it does not make it a considerable theme (Braun & Clarke, 2006). For this study, themes were developed using the prevalence of the codes because the study wants to understand the common themes found in cases where DM in KM is applied. An inductive approach was used, and for this reason, the themes were noted based on the pervasiveness of data sets because the themes are grounded in the data. In addition, the researcher utilised the network tool in Atlas.ti 9 to generate a network to visualise the relationships between codes and themes (Nowell et al., 2017). Figure 3.7 illustrates the visualisation tool used in Atlast.ti 9 to support the theory building.

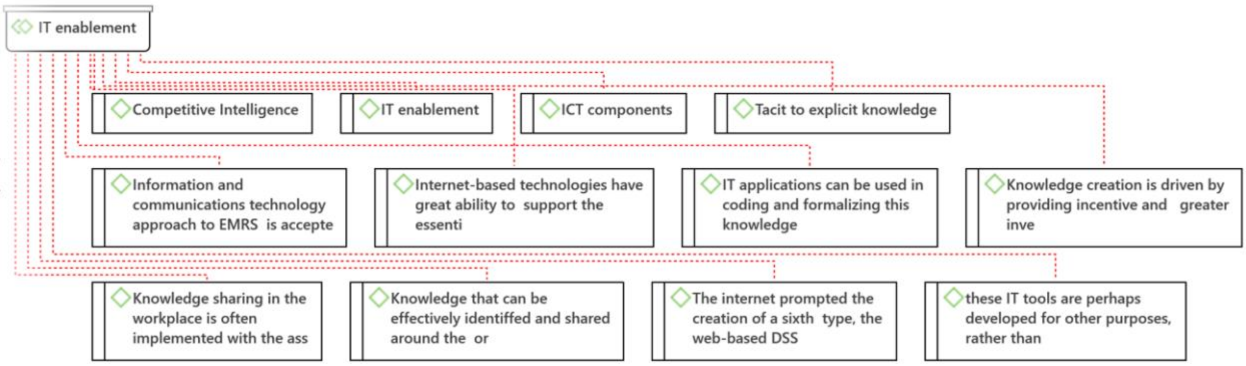

Figure 3. 7: Coding to Theory Building

Stage 2 revealed descriptive themes evaluated for relevance to the research question; 12 themes were identified and developed. Table 3.8 shows the list of descriptive themes used in the study.

Table 3. 8: Descriptive Themes emerging from the systematic review

	Descriptive Themes	Thematic Code
1	Improving operations means applying technologies to make appropriate changes.	Availability of IT
2	Establishing agility in an environment where other organisations are vying for the same niche forces organisations to change the way they offer services.	Competitive Advantage
3	CRM has become the concept that generates significant data, and organisations can use it to support their customers and improve their offerings.	Customer Relationship Management (CRM)
4	To perform at the highest level, decision-making activities need to respond to the organisational needs.	Decision Support
5	Lack of experience and skills in managing organisational data, information, and knowledge can lead to gaps in planning.	Knowledge activities for expertise
6	Big Data is proving to push organisations to relook their approach in identifying and capturing data to understand patterns and find solutions.	Big Data
7	Confusion about whether the organisation has valuable knowledge and what type of knowledge it is makes it difficult to conceptualise how to transform and use it.	Tacit vs Explicit knowledge acquisition
8	Organisations tend not to know where their knowledge assets are, how and where they flow, and if they are usable.	Enterprise wide knowledge
9	Knowledge base and Preserving and harnessing the organisational knowledge from employees.	Knowledge base
10	Understanding of data, information and knowledge assets of the organisation.	Intangible assets
11	Organisational performance dependent on a variety of capabilities.	Organisational Performance
12	Organisational data mining efforts need to apply specific techniques to search for particular details in their data. Some of the data that exists in organisations are used for different things.	Appropriate task matching during DM application

Stage 3

The selected descriptive themes went through further refinement, so they are specific enough to reflect a set of ideas contained in numerous text segments (Nowell et al., 2017). The refinement process included some themes not having enough supporting data, or the data diversity was too much, and consequently having to collapse some themes into each other while others had to be separated to create more themes (Braun & Clarke, 2006). A clear distinction between the themes was established, and the analysis and refinement in this stage developed analytical themes, which were individually named (Braun & Clarke, 2006). Figure 3.8 illustrates the process of developing analytical themes.

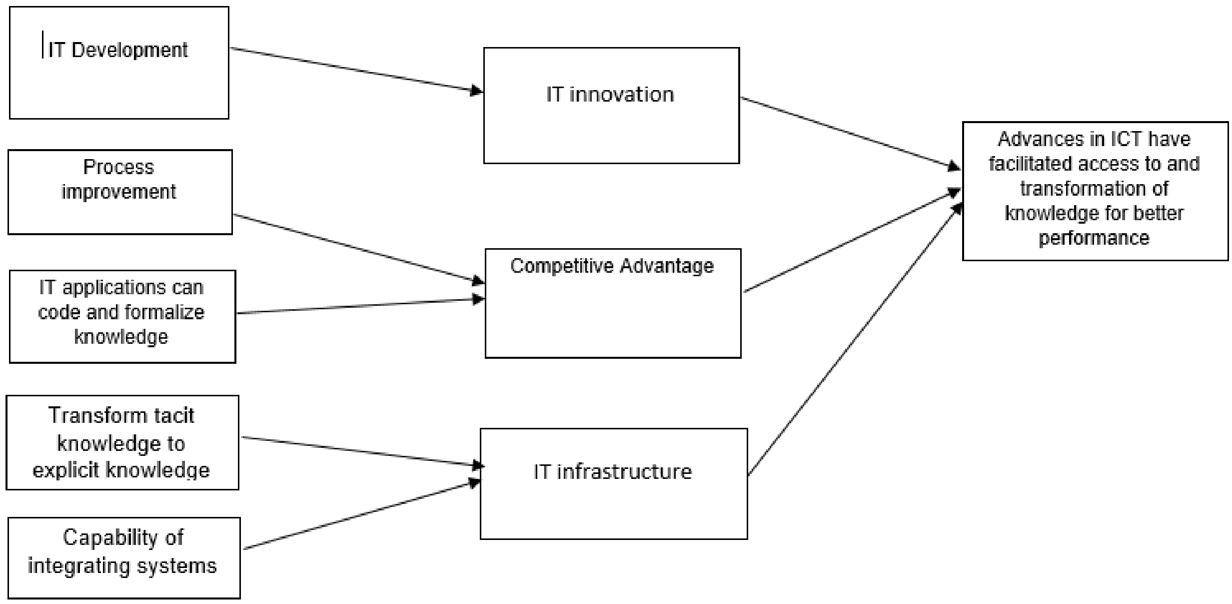

Figure 3. 8: Example of formulation of analytical themes

Analytical themes offer new conceptualisation and explanations because they go beyond the content of the original studies, especially with the descriptive themes as a backdrop of formulation (Thomas & Harden, 2008). The analytical themes were generated from descriptive themes, analysing statements with the research question in mind. Seven analytical themes were then developed by thematically analysing 54 articles related to the review objectives. Table 3.9 shows the seven analytical themes that emerged from the thematic analysis.

Table 3. 9: Stage 3- Analytical Themes

Analytical Theme	Description
AT- 1	Advances in ICT facilitate access to and transformation of knowledge for better performance.
AT- 2	A Knowledge base as a tool for improved competitive advantage and product and service development.
AT- 3	The emergence of Big Data analytics for Customer Relationship Management (CRM).
AT- 4	Understanding data, information and knowledge assets as interlinked capabilities for Decision Support.
AT- 5	A dual approach of automation and human expertise for efficiency.
AT- 6	The specificity of a KM task determines the effectiveness of the application of DM techniques.

The final analysis developed the analytical themes that could provide an overview of the story about what they each reveal about the topic (Nowell et al., 2017).

From the codes extracted from phase 1, the categorised codes reflected themes, and therefore, descriptive themes. Descriptive themes make it possible to show the relationship and similarities between the themes. This stage allowed for the grouping of the codes. Stage 1 provided a list of codes generated using Atlas.ti 9, where inductive coding was applied, the code and the surrounding text were highlighted to give the context of the codes in the articles. The list of codes generated went through an evaluation, and the researcher and the supervisor reached a consensus on the list of codes. The code was further evaluated to group the codes. It should be noted that the grouping was carried out by the individual researcher, which can pose the risk of bias.

3.6 Reliability, Validity and Trustworthiness

Consideration was given to the reliability, validity, and trustworthiness of the systematic review. Reliability is generally defined as the consistency of a measure or the degree to which scores approximate each other across multiple assessments of an instrument or multiple ratings of the same event (Cook & Campbell, 1979, p. 88). Validity is the measure of accurate depiction related to something that has been studied (Frey, 2018).

According to Lewis-Beck, Bryman & Futing Liao (2004, p. 1145), the criteria for trustworthiness are defined as follows:

Credibility is the persuasiveness of the report, and if it is a true representation of the experiences discussed;

Transferability refers to the usability of the report in settings similar to those of the original report;

Dependability looks at the methodological process and decisions employed and follows a process audit that offers sound assessment; and

Confirmability verifies that all the data collected and reported can be traced back to the original sources.

The trustworthiness criteria are important in developing a study that is acceptable and useful to other researchers. The review was conducted through recording, systematising, and being transparent about the methods of data collection and data analysis by providing comprehensive step by step details to allow the reader to make determinations on the credibility of the process (Nowell et al., 2017).

The integrity of the data obtained was ensured by:

- By adopting and adhering to the PRISMA-P statement checklist (Moher et al., 2015). In fulfilling a systematic review, the steps from the checklist were followed, ensuring a complete systematic review.
- Four databases were employed combined with different search terms to categorise the literature. Primary studies of different designs and lengths with differing outcomes were considered for the review, thus expanding the unit of analysis and providing comprehensive analysis.
- Synthesising and analysing relevant publications according to the inclusion and exclusion criteria.
- The researcher acknowledges that the lack of inter-coder reliability, as discussed in the Data Analysis of the section, may question the credibility of the study but the consultation and discussion with the expert concerning the code list and availability of the list of codes in Appendix D are in efforts to create transparency and replicability for future studies.

3.7 Chapter Summary

The systematic review is a very rigorous methodology because it provides a thorough and unbiased synthesis of a large quantity of information or relevant studies in one document. The chapter gives a background to the method, provides the justifications and activities for the research design employed, data collection, data analysis, and the methods' reliability, validity, and trustworthiness.

Chapter Four

4.1 Findings of the study

In Chapter 3, the design and methods for the systematic review was described. Through a process of screening and reading titles and abstracts, articles were included or excluded for the systematic review based on the inclusion and exclusion criteria.

4.2 Chapter Introduction

In this chapter, the findings from the 54 articles were synthesised through a thematic analysis approach. The findings are addressed in terms of the characteristics of the studies and followed by the synthesis in the form of analytical themes identified in the methodology section. The outcome of the data synthesis process is discussed below.

4.3 Study General Characteristics

The 54 articles appeared in 30 journals whose focus is DM and KM, covering the fields of management sciences, natural sciences, technology, hospitality and tourism and computer sciences. The publication rate across the time frame 2000 – 2017 shows that KM and DM are still interesting enough to investigate, with a steady climb of publications (Figure 3.4). The studies themselves were of a mix of qualitative and quantitative studies, where 33 (61%) of them were qualitative, and 21 (39%) were quantitative (Appendix B). The study designs assessed in the articles are discussed below.

<u>Qualitative</u>

The qualitative studies employed interviews and case studies and presented theoretical models, frameworks, and concepts related to using analytics and DM tasks to enable a knowledge-based or KM approach in improving performance in organisations. The different knowledge types are mentioned consistently in the articles, and some researchers go to the extent of associating the types with specific KM activities. The articles highlighted knowledge creation and knowledge acquisition as key steps of the KM life cycle and to further competitive advantage. Since the majority of the reviewed articles fall within the category of qualitative research, the researchers offered a general view of what theoretical application requires to come to reality. They further laid out relationships and linkages between the DM and KM activities and their importance.

Quantitative

The quantitative studies employed surveys and real-life application of tasks and techniques in DM and KM. They reflected more complicated scenarios in applying DM algorithms and models in generating knowledge that could be used in decision-making and performance improvement. Knowledge creation and acquisition stages in the KM life cycle are closely linked to the application of DM, with specific DM tasks. The quantitative studies reflected a diverse group of factors evident during application in an organisation:

- A specific issue is faced in an organisation that requires a specific answer or solution to it. Organisations presented in the quantitative studies presented their scenarios by classifying the problem, defining the problem, and what the answer to the problem must satisfy.
- There is much emphasis on data size since DM has shown to work better with much bigger data. The more data is available, the more comprehensive the DM activities.
- Technologies are a commonality in enabling a more robust and responsive organisation. Technologies such as IT show how they enhance systems available in the organisation.
- Data attributes are a significant consideration in determining the application of DM, and knowing which algorithms and model approach is appropriate in either identifying patterns or determining predictions.
- Articles showed that DM activities are closely related to one or two of the KM activities and not particularly the entire life cycle.

In the below presentation of analytical themes and subsequent sections, the articles were assigned identifiers (e.g. D1-D54 for all 54 articles), used to describe the supporting articles. The identifier per article can be seen in Appendix B.

4.4 The Analytical Themes Identified

Six analytical themes emerged from the systematic review, thematic analysis and synthesis.

4.4.1 AT- 1: Advances in ICT facilitate access to and transformation of knowledge for better performance.

The overall views throughout the 54 articles were that ICT enables better access to knowledge in the organisation. It is further highlighted that ICT supports the integration of other available systems and enhances the collaboration and interaction between those in an organisation. The studies illustrate this by D4, D13, D19, D20, D26 D28, D34 and D52, who discussed the potential of technologies and how they can support the organisation's performance. The following statements support the views:

"Many existing IT systems facilitate organising, formalising and distributing information that may be converted to knowledge using current components of an organisation's IT infrastructure such as databases and communication technologies. For example, the Internet, corporate extranets, and company portals primarily support the objective of making it easy to find and re-use sources of know how and expertise" (D4).

"Sensors and embedded technology now enable the transmission of real-time data from wireless networks, which will lead to the co-creation of new real-time knowledge among customers and vendors" (D28).

4.4.2 AT- 2: A Knowledge base as a tool for improved competitive advantage and product and service development.

The articles discussed the availability and accessibility of a platform, such as a knowledge base to improve the expertise of staff to solve problems and support the enhanced development of their products and service that provides a competitive edge for the organisation. D12, D15, D17, D23, D27, D28 and D53 emphasised that a knowledge base system supports product customisation, rapid innovation in the development and product configuration. This is indicated in the following quotations:

"...It allows an effective and efficient processing of knowledge transactions during product realisation" (D15).

"Organisations can use the availability of customer behaviour data to build on their development processes" D23.

4.4.3 AT- 3: The emergence of big data analytics for Customer Relationship Management (CRM).

The CRM theme was evident in the synthesis and was highlighted in different aspects. Appropriate CRM means (1) presenting an image of unity of the organisation; (2) knowing who your customers are and what their preferences and dislikes are; (3) pre-empting the needs of the customer and proactively addressing them; and (4) acknowledging customers dissatisfaction and taking corrective action (Dennis et al., 2001). D5, D24, D28, D38, and D49 highlight the importance of having a CRM enabled environment with the necessary tools and technologies. The following quotation supports the statement:

“The segmentation enables the company to provide more personalised and, therefore, more attractive product and service offerings to individual customer groups. Criteria for segmenting customers include customer profitability score, retention score, satisfaction and loyalty score, response to promotion” (D28).

In addition, D18, D23, D29, D34, D45, D48 and D51, highlighted the emerging technologies that generate data about customer behaviour, sales, buying patterns, including data that can develop marketing strategies. They further discuss how organisations must adopt technologies to capture, interpret and provide the necessary answers to strategically respond to the markets. The following quotation supports this:

“Big Data technology is used as a solution to analyse the big social media data related to the organisations and their competitors, and to visualise and benchmark comparisons among competitors across events, products, issues and any other areas that may affect business performance” (D45).

4.4.4 AT- 4: Understanding data, information and knowledge assets as interlinked capabilities for Decision Support.

The theme of understanding data, information and knowledge assets is represented by the overall studies that were thematically analysed. The articles highlighted the importance of knowing what intangible assets are available in the organisation, especially understanding what knowledge assets are available and accessible. They further indicate that knowledge as an asset, and a resource is important to an organisation’s daily operations and decision-making because the knowledge exists within the firm and its employees. They see data, information and knowledge assets as the catalysts for decision-making in an organisation. The following statement illustrates this:

"If we can put together some interesting metrics, understand what they mean in terms of the contribution of different intangibles (big data/information, knowledge, intelligence) to competitiveness and then allow specific firms in specific industries to evaluate their own circumstances, we can better advise practitioners on what intangibles to apply in what manner to what situation. Indeed, another core contribution from KM/IC is matching the right tools (e.g. communities of practice for tacit-to-tacit exchange) to the right circumstance, and this approach expands that capability" (D13).

4.4.5 AT- 5: A dual approach of automation and human expertise for efficiency.

The theme of automation and human intervention based on their expertise is presented in D10, D11, D14, D17, D28, D32 and D53. They present the factors that illustrate the challenges of attempting to advance either automation only or human expertise only. Additionally, they offer what they consider a more viable view of integrating the two approaches and the importance of using a dual approach. The following statements represent this:

"One also needs an unusual combination of skills and knowledge to transform data into actionable decisions. Not even the most sophisticated data mining software can obviate the need for a high degree of human skill and experience in the successful analysis and use of business data. A deep understanding of how the data are produced and transformed often comes only from experience"(D28).

"Therefore, given unusual data mining tasks or unfamiliar data mining algorithms, it is important for the data mining project teams to choose team members with diverse educational background and data mining experience so that the team can make an optimal decision on choosing a data mining method"(D32).

"Intelligence amplification by machine and mind can outperform a mind-imitating AI system working by itself. The combination of formal, explicit knowledge in the machine, and the non-formal, tacit knowledge of the users, can thereby result in problem-solving capabilities, which surpass either one of these components acting alone"(D53).

4.4.6 AT- 6: The specificity of a KM task determines the effectiveness of the application of DM techniques.

Throughout the 54 articles, the common theme evident in all articles was the application of specific DM techniques to specific knowledge types or a specific KM task to unearth possible solutions to specific challenges faced by the organisation. The consistently applied techniques identified in the reviewed articles were Fuzzy L Regression Analysis, Classification, Artificial Neural Network, Cluster Analysis, Prediction, Sequential P and Association Rules. With the combination of text mining, AI, XML and OLAP. In addition, they indicate that the size of available data determines the DM application to find patterns and knowledge, as illustrated by the following statement:

"But no one data mining algorithm has proved to outperform other algorithms in all tasks. Therefore, in the real-world data mining projects, data mining teams have to compare more than one data mining method carefully and choose one that has the best functioning performance" (D32).

"The selection of data mining algorithms, hypotheses formation, model evaluation and refinement are key components of this discovery process. Because it takes cycles of trials and errors to progressively produce the most useful knowledge from the data mining, a learning by experimentation approach can be useful to ensure that the process can eventually discover the useful knowledge" (D37).

The themes and the KM Process

In analysing the themes, different elements can be placed within the various stages of the KM process to highlight which stage of the process commonly occurs theoretically. Figure 4.1 illustrates four stages of KM placed in quadrants with vital elements found in the themes associated with them.

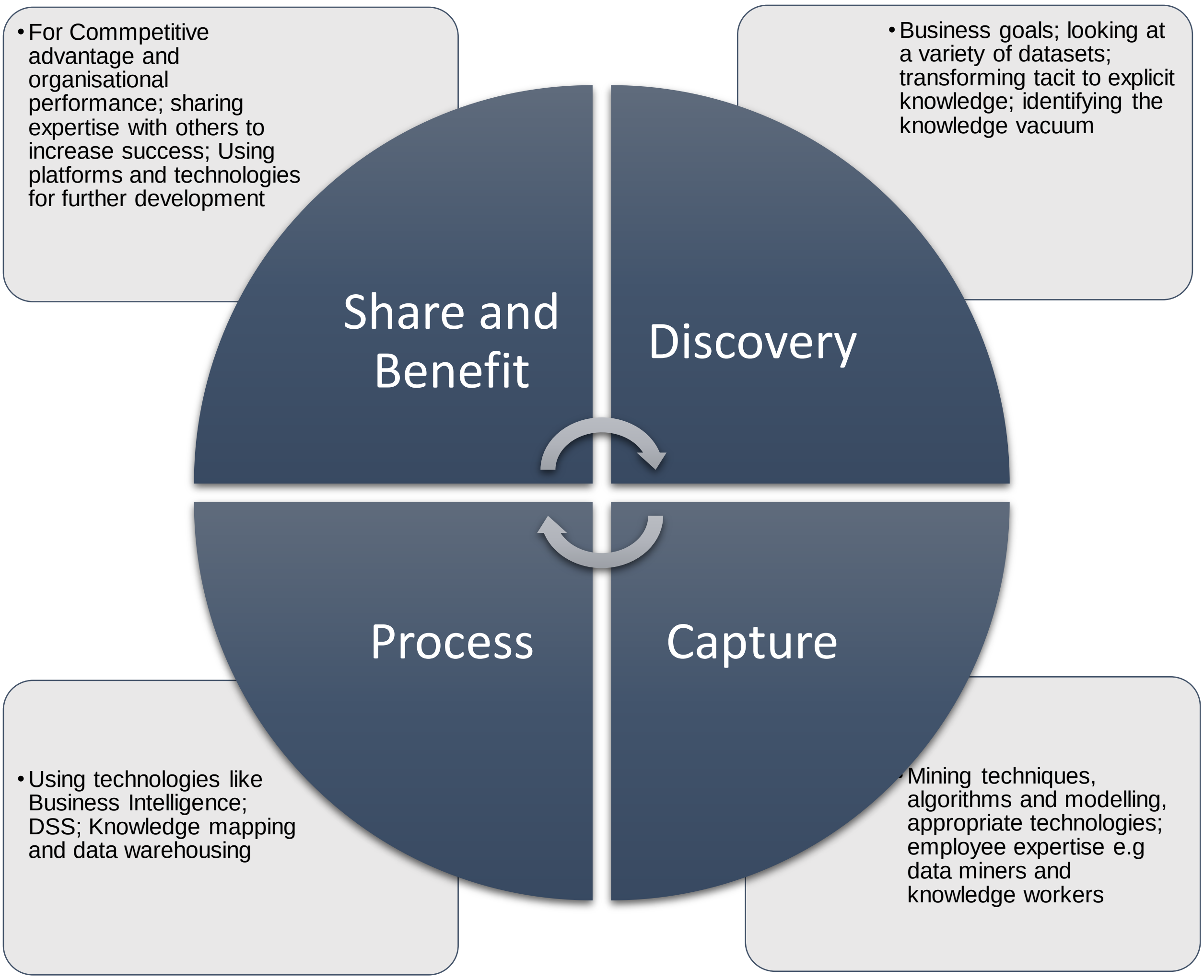

Figure 4. 1: Activity placement in the KM Process

Placing the common elements in the themes reflects the discovery stage, including the organisation's specific goals and processes of achieving them by experiencing the discovery process. The capture stage consists of DM techniques, algorithms, and models based on the goals set out while using technologies and expertise to support this part of the process. The process stage uses technologies, such as a knowledge base and DSS to access knowledge for better decision-making from continuous data analysis. The share and benefits stage allows the contributors to process, learn and develop abilities to make predictions and contribute to the firm's performance. Thus, reflecting on the articles' findings, the themes can be interlinked to the activities of DM and KM processes.

4.5 Chapter Summary

The chapter provided the findings from the thematic analysis of the 54 systematically reviewed literature articles. They were reported and presented through the six analytical themes in the methodology section and placed within the four quadrants of a KM cycle for context.

Chapter Five

5.1 Chapter Introduction

The synthesis in Chapter 3 laid out the analytical themes identified, with their findings from the 54 articles reviewed. Chapter 5 discusses the findings and how they answer the research objectives and analyse the six themes highlighted from the synthesis.

5.2 Discussion

The aim of the study is to provide answers to the following objectives derived from the topic:

- What themes are prevalent in the study's unit of analysis?
- From the patterns discovered and trends highlighted, what do they illustrate about the nature of the relationship link between DM and KM?
- Is the process of discovering knowledge an automated activity, or is it an action that requires integration into the process of discovering new knowledge?

The study analysed 54 research articles from four databases published between 2000 and 2017. The purpose of the study was to provide perspectives on data mining in KM. The findings identified the following:

- **What themes are prevalent in the study's unit of analysis?**

The first objective was to identify the prevalent themes found in the reviewed article. Six main themes were identified, and they were:

1. Advances in ICT facilitate access to and transformation of knowledge for better performance.
2. A knowledge base as a tool for improved competitive advantage and product and service development.
3. The emergence of Big Data analytics for CRM.
4. Understanding data, information and knowledge assets as interlinked capabilities for Decision Support.
5. A dual approach of automation and human expertise for efficiency.
6. The specificity of a KM task determines the effectiveness of the application of DM techniques.

The themes are discussed below:

AT- 1: Advances in ICT facilitate access to and transformation of knowledge for better performance.

Fast-tracking performance means organisations are adopting resilience for measuring the company's adaptive, flexibility, and learning capacity. The articles reflected the importance of having high performance as a foundation for responding to the organisation's needs. Performance depends on several aspects dependent on one another for survival in a highly competitive environment. It is a deciding factor as to whether the firm will survive or not in the changing environment. One of the key aspects highlighted by the reviewed articles was the existence of ICT technologies and how they can create operational efficiencies and enable comprehensive access to the organisation’s assets. The goal of information technologies is facilitation, leveraging, and sharing of knowledge. Technology has made it possible to identify and collect large amounts of data. This data is obtained from various places, whether intentionally stored or not, databases and social media platforms. There are large amounts of codified knowledge among the stored data. Technologies enabling the discovery of knowledge in organisations can be beneficial. Support for new technological adoption depends on the accessibility of internal resources, such as human resources, capital, equipment, and organisational time to implement. Understanding what technologies exist supports the structuring of business, considering that they are already part of different operations.

ICT tools have revolutionised how organisations and knowledge workers interact with the available knowledge, from access, analysis and retrieval. Considering the large amounts of data available in different places or on different platforms within an organisation, implementing and applying the right technology can bring the most significant advantages. This depends on the strategic decision to implement systems at the beginning of the IT solution development to ensure continuity in developing new and older systems. A competitive advantage that is knowledge-based and lasting can be achieved and sustained if organisations are actively and continuously involved in knowledge creation. This theme is in line with the research by Agarwal (2014); Crawford et al. (2011); Mishra et al. (2018); and Melville et al. (2004), who believe that ICT facilitates structured and cross-organisational collaborations between systems, where data can be transformed and create effective KM.

AT- 2: A Knowledge base as a tool for improved competitive advantage and product and service development.

The second theme identified in the articles was that for products and services to be ahead of others, an organisation needs to develop a well-designed and functioning knowledge base. Data and knowledge can be captured to support the staff and equip them with the expertise to develop products and services and provide an organisation with a competitive advantage. A knowledge-based approach (data analysis and reasoning) can solve problems, as it combines DM tools and KM framework to leverage hidden knowledge. It is the organisation’s responsibility to anticipate and prepare, adapt, and respond to the environment and competitive market changes. Competitive advantage is a common goal that differentiates organisations from others. The intention of a knowledge base must be inspired by knowledge as “a factor of production.” Knowledge elicitation and Knowledge-based engineering (KBE) have become the key to endless innovation due to organisations wanting to be different in product and service development and offering.

The articles pointed to organisations' pursuing innovative actions to maintain a competitive advantage. Competitive Advantage has three known aspects: the benefit that the business offers, the target audience of the business, and knowing who the competitors are in the environment. These aspects direct and help to measure the progress and goals set and reached. Ultimately, this indicates that the goal is sustainable. Using a knowledge base can take advantage of the resources available in the organisation by using them to predict the future and redesign the business model. Knowledge bases support collecting, retrieving, organising and sharing knowledge.

There are two known primary forms of knowledge bases: machine-readable, where the data needs to be structured, and humans can easily read human-readable base data. Some articles asserted that a knowledge base supports the identification of organisational knowledge, employee knowledge, competitor knowledge, operation's knowledge, industry knowledge and supplier knowledge by integrating a KM approach. This theme confirms the views of Abusweilem and Abualoush (2019); Obeidat et al. (2017); Abualoush et al. (2018), that the value of knowledge as an asset for production is unparalleled and stands to differentiate an organisation from others. Making informed decisions and developing innovative products results directly from having adequate knowledge capability. A knowledge base can play a role and keep all the types of knowledge from which solutions to problems can be found.

AT- 3: The emergence of big data analytics for Customer Relationship Management (CRM)

The third theme looked at the rise of Big Data. This is the ability to take advantage of the CRM application in an organisation means there is the availability of data that can answer the knowledge needs of the firms and those of the customers. This will contribute to the growth and survival of the firm. CRM can be taken a step further by applying analytics to focus on customer data, their purchasing patterns of the different products offered over time, to understand the trends and possibly customise future offerings. This means that the organisation needs to interrogate the data from these customers to determine if the organisation is reaching the right market and if the market is responding to their efforts. This can be done through feedback from the customers, where an implemented feedback loop can support securing the voice of the customer. Customer comments, sentiments or experiences resulting from introducing new products to the market are collected and processed to understand their reactions and make necessary changes.

The articles further reflected on the enormous explosion of data from elements, such as 'sentiments' or opinions available from social media platforms, such as Facebook, Twitter, blogs, message boards, user forums, to sources of data such as purchasing patterns and (on-line) customer history, financial data, supply chain data mainly collected through RFID (Radio-Frequency Identification), and communications data (from phone call records to social media content) among others. Significant data is needed to be interrogated in making informed business decisions to support executives, managers and corporate end-users. Such data is not easily processed using regular processing technologies but requires many dedicated technologies to analyse it, confirming what Manjarres et al. (2018) discuss in their study about finding the appropriate technology for analysing large quantities of data. Within the synthesised studies, articles highlighted the development of marketing strategies associated with CRM, which benefit from the availability of Big Data.

AT- 4: Understanding data, information and knowledge assets as interlinked capabilities for Decision Support.

Throughout the synthesis, the common belief in all the articles is the existence and availability of data and whether organisations understand its use. Understanding the use means understanding how it can be turned into information and knowledge for decision-making. Organisations absorb, collect, store, and process information from the environment, which functionally resembles information-processing systems. It is ultimately essential for the organisation to know which data is needed for which problem and to interpret it effectively, and into a format understood by the whole organisation. The studies highlighted the complexities of identifying intangible assets and accepting the role they play when making decisions. The data collected can be varying in sets; it can be relational, complex and sometimes incomplete, structured, semi-structured and unstructured data, where some can be static or dynamic, which is due to the way, it is collected. While the information can be easier to identify, it can be challenging to find a use for all of it.

Knowledge is complicated because it categorises tacit or explicit and determines which category is suitable for what scenario, confirming the perspectives of Selamat et al., 2020; Davenport et al. (1998); and Faniel and Majchrzakb (2007). They stated that tacit and explicit knowledge was important but explicit knowledge is easily classified and explored, while tacit can provide more in-depth details but require approaches and technologies to codify. Capitalising on the knowledge worker's experience and skills can give the organisation a competitive advantage. Experts' knowledge (in decision support) and knowledge extracted from data (in DM) can be combined to increase the quality of decisions. Technologies with AI technologies can enhance the creation of knowledge, the storage thereof, knowledge dissemination and management processes for better decision-making.

The nature of decisions varies, whether natural or delicate. Therefore, data can be used as part of investigations to respond to needs within the organisation and to make informed decisions by collecting data that reflect your business or scientific activities. Decision-making must be rooted in quality for operations, management and planning through the collection, organisation, and analyses of business data. The articles went further to offer a perspective that a computer-based application, such as a DSS can help a firm collect or identify data from many sources; such as raw data from the organisation itself, documents compiled, firm staff knowledge, such as personal experiences, and expertise such as employees, managers, executives, and from the business models.

The articles highlighted that DSS could work in conjunction with other DM techniques to sift through massive amounts of data reflecting performance. Decision-Making processes consider the different intentions and goals of the organisation. With the availability of data, analysis can occur, which will lead to better decision-making.

AT- 5: A dual approach of automation and human expertise for efficiency

The fifth theme identified is the dual approach of human intervention and some automation. Decisions about automation are more feasible, and organisations must consider which decisions can be made by people, and which can be computerised. The articles provided a view that data alone cannot directly answer organisational questions, but it needs integration with other resources and tools to respond to different challenges. Discovering this data and knowledge requires the right people for the job, who have experience and knowledge. Knowing the data and its application allows for preparation to know the data and develop learning through that interaction to develop expertise. Automated DM happens when algorithms are "trained" on patterns and structures of data that may be valuable. As much as it is automation, a human element is still essential to feed the systems with activities to retrieve the outputs. The expertise can be in the form of problems already solved by previous workers that can help current and future workers to apply to new situations. Having automation with human intervention improves the accuracy and timeliness of decision-making.

Decision-making tasks do not work in isolation and depend on experiences, judgements, human knowledge and preferences. This highlights that the system works under the assumption that the decision-maker is knowledgeable with the type of problem to be solved, giving complete control to the user regarding the acquisition of the information, evaluation thereof and making the final decision. The notion puts the spotlight and pressure on the expertise of the workers because they would need to be well versed in every aspect of the firm. The other side of this is that if the workers do not know or understand the nature of the problems faced, wrong decisions are taken, with potential consequences. The theme confirmed the views and linked to findings by Larose and Larose (2014); Liao et al., 2012; and Wang and Wang, 2008a), that it is a fallacy to believe that human intervention is not a requirement when running a system and that the system can self-correct.

AT- 6: The specificity of a KM task determines the effectiveness of the application of DM techniques.

The sixth and last theme that emerged and was predominant in the articles was how specific KM activity is. The more specific the application of a DM technique will be. DM can extract tacit knowledge relevant through data correlation sourced from the information system assets.

The studies consistently highlighted that a specific DM technique or task must be developed and applied for each problem. Sometimes it takes more than one technique or more than one KM task to unearth the right answer. The objective of DM techniques is to explore tomorrow's success signals. DMs objective is the development of predictive models. Through observation, a model is constructed that allows the classifications or predictions for new observations to be actioned by an analyst. The DM method and datasets are domain-dependent, and breaking extensive data into smaller segments makes it easier to manage and apply the specific technique. Each DM technique can find specific patterns within the firm's data. Meaning they should be applied for a particular need or to answer specific questions that have risen.

DM requires considerable amounts of data to effectively find patterns and knowledge. The algorithms are applicable only on large amounts of data and not necessarily on the smaller sized data. Therefore, there is a need for specific DM systems designed for particular data types. In addition, the articles discussed DM playing a dominant role in the stages of the knowledge creation and capture phase of the KM life cycle. DM can evaluate vast amounts of data in the organisation that can help answer specific questions, since it can reveal particular patterns in the data and information used. The theme is linked to the findings from Liao et al., (2012); Kriegel, et al. (2007); and Rygielski et al. (2002), who highlighted that for DM, one size does not fit all, but it requires several possible scenarios.

- **From the patterns discovered and trends highlighted, what do they illustrate about the nature of the relationship between DM and KM?**

DM contributes towards knowledge discovery and overall KM only when an organisation establishes their KM strategy first, are acquainted with their data and then applies DM to help solve their problems. The contribution of DM to KM is based on the business intentionality towards making the most of the available data by using qualified and experienced workers who can sift through the data and formulate solutions for the decision-makers.

The articles reflected upon an integrative process between DM and KM, where DM is a crucial driver to the KM system. When integrated, it broadens the possibilities of diverse outcomes. Such outcomes can be in the form of an increase in sales, targeted and proactive marketing strategies, being more customer focused, and higher profitability to gain a competitive advantage. This confirms what Bharara et al. (2017), Neaga and Liu (2014), and Selamat et al. (2020) discussed in their investigations, that DM is an essential component of KM and sometimes as a stage in the KM process. The capability of having the right strategy in terms of knowledge discovery is in line with Silwattananusarn and Tuamsuk (2012). They assert that the right knowledge can be discovered by appropriately using DM techniques and algorithms.

KM is just as susceptible to faults of DM as other systems because it is not a perfect process. DM does not always reveal relevant knowledge because data changes or accelerates in volumes at volatile rates. Such volatility requires that mining models adjust accordingly and be up to date to give accurate interpretations. The articles further revealed that DM techniques are difficult to compare in terms of applicability and choosing the most appropriate one because some methods have complex computations. This links to the views of Sharma (2014), who highlighted the challenges found with the application of DM. There are scenarios of DM techniques unearthing knowledge that are not relevant but may be useful as the business needs evolve or circumstances change over time.

When planning for DM and KM implementation and application, strategies need to reflect organisational goals. The implementation should align with the business goals to leverage benefits for performance while meeting objectives and priorities. KM technologies, coupled with DM principles, can effectively manage explicit and implicit knowledge especially when transforming implicit knowledge to explicit knowledge. The success of KM relies on the integration within the organisation's overall strategic vision, so the true value may not be realised.

Common elements in a Data Mining process for Knowledge Management:

Some very particular elements need to be in place for knowledge to be discovered using DM in KM. The analysis of the articles showed these elements should be present for application, and they are:

Variety in datasets **-** For DM to take place, there is a need for different types of data in the various databases in the organisation, such as relational, transactional and multimedia databases. It can be collected, accessed, transformed, codified and used by various stakeholders.

Employee expertise in the field **-** tacit to explicit knowledge will always assist the firm in knowing which people have experience and in which area. The correct workers will interpret data into high-quality information because they know what the organisation's decision-makers are looking for from the data.

Appropriate Algorithms and Models for the right datasets **-** Since there is no specific technique to answer all questions, suitable techniques need to be applied to a situation trying to answer a specific question.

Client-centred process - where structured and unstructured data can enhance the customer's experience, identify shopping patterns, improve customer satisfaction, and increase brand loyalty. Customer's habits become a focal point and a rich source of data.

Built-in systems such as a knowledge base to support decision support for organisational performance - businesses need to survive in the competitive environment, so keeping up to date with trends and planning for the future is vital. Using knowledge or databases as part of the DSS would open doors to growth and innovation. This tool with DM can streamline processes and provide a greater understanding of the business.

For understanding the role DM plays in KM, it is necessary to understand the organisational strategy. To do that is to ask strategic questions, such as what the core business processes are, which key decisions within the processes require support from analytic insights, what does the organisation consider important in terms of information, and would enable better business performance and propel them to be better than their competitors. Combined with the emergence of various data, knowledge forms and sources generated in the environment and within organisations, strategic decision-making has become much more complex and not one size fits all.

- **Is the process of discovering knowledge an automated activity or is it an action that requires integration into the process of discovering new knowledge?**

DM is not an automated activity but rather one that requires integration because of its capacity to be configured to suit a particular situation. The articles put into focus that integration of technologies, infrastructure, strategies, operations and talent across the business and geographical boundaries can bring in challenges that companies cannot manage due to capacity. Through the process of organisational learning, firms develop and grow their strategies and implement incremental changes because of the awareness created throughout the process. Whether DM is an automated process depends on the organisation and the system design. The organisation can answer whether the process needs to be part of the strategies, technologies, and people to advance a more automated approach in line with the business strategies and goals.

Thus, from the initial strategic planning, decision-makers must decide which systems are suitable for the challenges or solutions they seek, figuring out how the system must respond and to what degree the response needs to reach. Reflecting how systems and technologies require human intervention for efficiency and improvement. This is in line with the findings of Bach and Alessa (2014); Martínez-Plumed et al. (2021); and Larose and Larose (2014), who highlighted that DM processes are harder to automate and where there are parts that cannot be done automatically, human intelligence and intervention must support the phases that cannot be automated.

5.3 Limitations of the study

The limitations of the study are threefold. First, the study did not employ an independent coder during the coding phase of the study, which may have led to some bias since an independent coder could have introduced different perspectives from their coding method. Second, there is a lack of theory development in DM for KM. Including qualitative and quantitative studies provided a skewed picture of how DM plays a role in KM theoretically because the application and theory varied with applying appropriate algorithms and techniques to real-life scenarios. The third and last limitation, KM processes across the studies, were defined and understood differently, which created inconsistencies in developing an overview of KM and its application.

5.4 Recommendation for future studies

This study reflected on the common themes found with DM in KM for 2000 to 2017, and themes were illustrative of the activities during the rise and steadiness of publications in the areas that provide an overview of theory and some level of application. There are two recommendations, the first is that future research should include perspectives from real-life data driven organisations with KM systems already in place so that more in-depth and reality-based scenarios can be investigated. In addition, future studies can adopt a mixed-method approach, with a combination of a case study and quantitative method to have a large size of data to work from.

The second recommendation is future research focused on the concept and construct changes of DM and KM through a systematic review over two years. This can speak to the commonly understood or misunderstood definitions of DM and KM and the emerging change in the names.

5.5 Conclusion of the study

The study was guided by its aim to understand whether knowledge discovery is automated or integrated through a systematic review of DM in KM. The aim was guided by three objectives, which were what themes are prevalent in the study's unit of analysis; the nature of the relationship link between DM and KM based on the patterns discovered and trends highlighted; and whether the process of discovering knowledge an automated activity or whether it is an action that requires integration into the process of discovering new knowledge. A systematic review was conducted to identify themes found in studies that covered the topic.

The current study has found themes to represent varying factors in application but with similar meaning in definition, and the themes were evident across different disciplines. Conclusions drawn from the findings indicate that DM and KM work where there is the availability of data. Industries with high levels of both knowledge and competitive intelligence have high data capabilities. The nature of competitive intelligence makes it possible to further developments of capturing different knowledge types. Collecting and analysing data leads to competitive insights making it necessary for the identification of intangible assets, such as data, information, knowledge and intelligence. KM issues are related to the quality, relevancy and usability of knowledge being adopted. Thus, comprehensive knowledge discovery requires efficient KM and DM integration. The integrative models of DM and KM are still not investigated enough, especially in practice. It is a dynamic process that requires and depends on several other pieces within an organisation to come to life. KM needs an intertwined network of systems to enhance an organisation's performance.

Bibliography

Abualoush, S. H., Obeidat, A. M., Tarhini, A., Masa'deh, R. e., & Al-Badi, A. (2018). The role of employees' empowerment as an intermediary variable between knowledge management and information systems on employees' performance. *VINE Journal of Information and Knowledge Management Systems, 48*(2), 217-237. doi:10.1108/VJIKMS-08-2017-0050

Abualoush, S., Masa'deh, R. e., Bataineh, K., & Alrowwad, A. (2018). The role of knowledge management process and intellectual capital as intermediary variables between knowledge management infrastructure and organization performance. *Interdisciplinary Journal of Information, Knowledge, and Management, 13*, 279-309.

Abubakar, A. M., Elrehail, H., Alatailat, M. A., & Elçi, A. (2019). Knowledge management, decision-making style and organizational performance. *Journal of Innovation & Knowledge, 4*(2), 104-114. doi:https://doi.org/10.1016/j.jik.2017.07.003

Abusweilem, M., & Abualoush, S. (2019). The impact of knowledge management process and business intelligence on organizational performance. *Management Science Letters, 9*(12), 2143-2156.

Agarwal, P. (2014). Benefits and issues surrounding data mining and its application in the retail industry. *International Journal of Scientific and Research Publications, 4*(7), 1-5.

Alasadi, S. A., & Bhaya, W. S. (2017). Review of data preprocessing techniques in data mining. *Journal of Engineering and Applied Sciences, 12*(16), 4102-4107.

Alavi, M., & Leidner, D. E. (2001a). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*, 107-136.

Alavi, M., & Leidner, D. E. (2001b). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS quarterly, 25*(1), 107-136. doi:10.2307/3250961

Al-Emran, M., Mezhuyev, V., Kamaludin, A., & Shaalan, K. (2018). The impact of knowledge management processes on information systems: A systematic review. *International Journal of Information Management, 43*, 173-187. doi:https://doi.org/10.1016/j.ijinfomgt.2018.08.001

Allen, M. (2017). The SAGE Encyclopedia of Communication Research Methods [4]. doi:10.4135/9781483381411

Alyoubi, B. A. (2015). Decision Support System and Knowledge-based Strategic Management. *Procedia Computer Science, 65*, 278-284. doi:https://doi.org/10.1016/j.procs.2015.09.079

Amin, M. S., Chiam, Y. K., & Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics, 36*, 82-93. doi:https://doi.org/10.1016/j.tele.2018.11.007

Apte, C., Liu, B., Pednault, E. P., & Smyth, P. (2002). Business applications of data mining. *Communications of the ACM, 45*(8), 49-53.

Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1997). The Place of Inter-Rater Reliability in Qualitative Research: An Empirical Study. *Sociology, 31*(3), 597-606. doi:10.1177/0038038597031003015

Aromataris, E., & Pearson, A. (2014). The Systematic Review: An Overview. *AJN The American Journal of Nursing, 114*(3), 53-58. doi:10.1097/01.NAJ.0000444496.24228.2c

Bach, C., & Alessa, A. (2014). Data Mining and Knowledge Management for Marketing. *International Journal of Innovation and Scientific Research, 2*, 321-328.

Bahra, N. (2001). *Competitive knowledge management.* Houndmills, Basingstoke, Hampshire: Palgrave.

Becerra-Fernandez, I., & Sabherwal, R. (2010). *Knowledge management: systems and Processes.* New Delhi: PHI Learning Private Limited.

Berente, N., Vandenbosch, B., & Aubert, B. (2009). Information flows and business process integration. *Business Process Management Journal*, 15(1):119-141.

Bharara, S., Sabitha, A. S., & Bansal, A. (2017). *A review on knowledge extraction for Business operations using data mining.* Paper presented at the 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence.

Bharati, M., & Ramageri, B. (2010). Data mining techniques and applications. Indian Journal Of Computer Science And Engineering 1(4):301-305.

Birzniece, I. (2011). Artificial intelligence in knowledge management: Overview and trends. *Computer Science (1407-7493), 46*.

Boland, A., Cherry, M. G., & Dickson, R. (2014). *Doing a systematic review: A student's guide.* London: SAGE.

Bolisani, E., Scarso, E., & Zieba, M. (2015). Emergent Versus Deliberate Knowledge Management Strategy: Literature Review and Case Study Analysis. In (pp. 153-160). Kidmore End: Academic Conferences International Limited.

Bolisani, E., Scarso, E., & Zięba, M. (2016). How to deal with knowledge in small companies? Defining emergent KM approach. *International Journal of Learning and Intellectual Capital, 13*(2-3), 104-118. doi:10.1504/ijlic.2016.075701

Bose, R., & Sugumaran, V. (2003). Application of knowledge management technology in customer relationship management. *Knowledge and Process Management, 10*(1), 3-17. doi:https://doi.org/10.1002/kpm.163

Bosnjak, Z., Grljevic, O., & Bosnjak, S. (2009). CRISP-DM as a framework for discovering knowledge in small and medium sized enterprises' data.509-514.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3(2):77-101.

Cannavacciuolo, L., Iandoli, L., Ponsiglione, C., & Zollo, G. (2015). Knowledge elicitation and mapping in the design of a decision support system for the evaluation of suppliers' competencies. *VINE, 45*(4), 530-550. doi:10.1108/VINE-01-2015-0011

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. *SPSS inc*, 9:13.

Chen, M., Ebert, D., Hagen, H., Laramee, R. S., Van Liere, R., Ma, K.-L., Ribarsky, W., Scheuermann, G., and Silver, D. (2008). Data, information, and knowledge in visualization. *IEEE computer graphics and applications, 29*(1), 12-19.

Cheng, H., Lu, Y.-C., & Sheu, C. (2009). An ontology-based business intelligence application in a financial knowledge management system. *Expert Systems with Applications, 36*(2, Part 2), 3614-3622. doi:https://doi.org/10.1016/j.eswa.2008.02.047

Chowdhury, S. I. (2009). A conceptual framework for data mining and knowledge management. In: Rahman, H. (ed.). *Social and political implications of data mining: Knowledge management in e-government: Knowledge management in e-government*. Hershey, PA: Information Science Reference.

Cios, K. J., & Kurgan, L. A. (2005). Trends in data mining and knowledge discovery. In: Pal, N.R., & Jain, L. (eds.). *Advanced Techniques in Knowledge Discovery and Data Mining*. London: Springer London.

Coakes, E., Amar, A. D., & Luisa Granados, M. (2010). Knowledge management, strategy, and technology: a global snapshot. *Journal of Enterprise Information Management, 23*(3), 282-304. doi:10.1108/17410391011036076

Cody, W. F., Kreulen, J. T., Krishna, V., & Spangler, W. S. (2002). The integration of business intelligence and knowledge management. *IBM Systems Journal, 41*(4), 697-713. doi:10.1147/sj.414.0697

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation : design & analysis issues for field settings*. Boston: Houghton Mifflin.

Crawford, J., Leonard, L. N. K., & Jones, K. (2011). The human resource's influence in shaping IT competence. *Industrial Management & Data Systems, 111*(2), 164-183. doi:10.1108/02635571111115128

Dang, Y., & Yuan, Z. (2011). *The Application and Research of Knowledge Warehouse in the Management of Enterprise Resources.* Paper presented at the 2011 International Conference on Computational and Information Sciences.

Dastyar, B., Kazemnejad, H., Sereshgi, A. A., & Jabalameli, M.A. (2017). Using data mining techniques to develop knowledge management in organizations: A review. *Journal of Engineering, Project, and Production Management,* 7(2):80-89.

Davenport, T. H., De Long, D. W., & Beers, M. C. (1998). Successful knowledge management projects. *Sloan management review, 39*(2), 43-57.

Davies, J., Sure, Y., Grobelnik, M., & Mladenić, D. (2005). Automated knowledge discovery in advanced knowledge management. *Journal of Knowledge Management,* 9(5):132-149.

Dawei, J. (2011). *The Application of Date Mining in Knowledge Management.* Paper presented at the 2011 Fifth International Conference on Management of e-Commerce and e-Government.

Demarest, M. (1997). Understanding knowledge management. *Long Range Planning*, 30(3):321-384.

Denizhan Kalkan, V. (2008). An overall view of knowledge management challenges for global business. *Business Process Management Journal, 14*(3), 390-400. doi:10.1108/14637150810876689

Dennis, C., Marsland, D., & Cockett, T. (2001). Data mining for shopping centres–customer knowledge-management framework. *Journal of Knowledge Management,* 5(4):368-374.

Desouza, K. C., & Awazu, Y. (2006). Knowledge management at SMEs: five peculiarities. *Journal of Knowledge Management, 10*(1), 32-43. doi:10.1108/13673270610650085

Devedzic, V. (2001). Knowledge modeling – State of the art. *Integrated Computer-Aided Engineering, 8*, 257-281. doi:10.3233/ICA-2001-8307

Drucker, A. M., Fleming, P., & Chan, A.-W. (2016). Research Techniques Made Simple: Assessing Risk of Bias in Systematic Reviews. *Journal of Investigative Dermatology, 136*(11), e109-e114. doi:https://doi.org/10.1016/j.jid.2016.08.021

Dybå, T., Dingsoyr, T., & Hanssen, G. K. (2007). *Applying Systematic Reviews to Diverse Study Types: An Experience Report.* Paper presented at the First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007).

Evidence Partners. (2021). DistillerSR. Retrieved from https://www.evidencepartners.com/

Faniel, I. M., & Majchrzak, A. (2007). Innovating by accessing knowledge across departments. *Decision Support Systems, 43*(4), 1684-1691. doi:https://doi.org/10.1016/j.dss.2006.09.005

Fowler, A. (2000). The role of AI-based technology in support of the knowledge management value activity cycle. *The Journal of Strategic Information Systems, 9*(2), 107-128. doi:https://doi.org/10.1016/S0963-8687(00)00041-X

Frey, B. B. (2018). The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation [4]. doi:10.4135/9781506326139

Gadu, M., & El-Khameesy, N. (2014). A knowledge management framework using business intelligence solutions. *International Journal of Computer Science Issues (IJCSI), 11*(5), 102.

Geist, I. (2002). *A framework for data mining and KDD*. Paper presented at the Proceedings of the 2002 ACM symposium on Applied computing, Madrid, Spain. https://doi.org/10.1145/508791.508887

GmbH, S. S. D. (2021). ATLAS.ti. Retrieved from https://atlasti.com/product/what-is-atlas-ti/

Gold, A. H., Malhotra, A., & Segars, A.H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of management information systems,* 18(1):185-214.

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews.* 2nd edition. ed. Los Angeles: SAGE.

Grant, R.M. (1991). The resource-based theory of competitive advantage: Implications for strategy formulation. *California Management Review*, 33(3):114-135.

Grey, D. (1996). *What is knowledge management?* [Online] Available from: http://www.km-forum.org/t000008.htm [Accessed: 10 February 2019].

Grobelnik, M., & Mladenić, D. (2005). Automated knowledge discovery in advanced knowledge management. *Journal of Knowledge Management, 9*(5), 132-149. doi:10.1108/13673270510622500

Gröger, C., Schwarz, H., & Mitschang, B. (2014). The manufacturing knowledge repository consolidating knowledge to enable holistic process knowledge management in manufacturing. *ICEIS 2014 - Proceedings of the 16th International Conference on Enterprise Information Systems*, Lisbon:39-51. [Online] Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84902354934&partnerID=40&md5=996028059b9f9264850b953eb36637bd [Accessed: 27 April 2014 through 30 April 2014].

Grossman, R. L., Hornick, M. F., & Meyer, G. (2002). Data mining standards initiatives. *Communications of the ACM, 45*(8), 59-61.

Gupta, M. K., & Chandra, P. (2020). A comprehensive survey of data mining. *International Journal of Information Technology*, 1-15.

Gupta, M., Chandra, B., & Gupta, M. P. (2014). A framework of intelligent decision support system for Indian police. *Journal of Enterprise Information Management, 27*(5), 512-540. doi:10.1108/JEIM-10-2012-0073

Han, J., & Kamber, M. (2001). *Data mining : Concepts and techniques.* San Francisco: Morgan Kaufmann Publishers.

Han, J., Kamber, M. & Pei, J. (2012). Data mining : concepts and techniques. 3rd ed. Amsterdam: Elsevier/Morgan Kaufmann.

Harrison, R., Jones, B., Gardner, P., & Lawton, R. (2021). Quality assessment with diverse studies (QuADS): an appraisal tool for methodological and reporting quality in systematic reviews of mixed- or multi-method studies. *BMC Health Services Research, 21*(1), 144. doi:10.1186/s12913-021-06122-y

Hawkins, J.M. (2017). Thematic Analysis. In: Allen, M. (ed.). *The SAGE Encyclopedia of Communication Research Methods*. Thousand Oaks, California: SAGE Publications.

He, W., Wang, F.-K., & Akula, V. (2017). Managing extracted knowledge from big social media data for business decision making. *Journal of Knowledge Management, 21*(2), 275-294. doi:10.1108/JKM-07-2015-0296

Hedgebeth, D. (2007). Data-driven decision making for the enterprise: an overview of business intelligence applications. *VINE, 37*(4), 414-420. doi:10.1108/03055720710838498

Heinrichs, J. H., & Lim, J.-S. (2003). Integrating web-based data mining tools with business models for knowledge management. *Decision Support Systems, 35*(1), 103-112. doi:https://doi.org/10.1016/S0167-9236(02)00098-2

Heisig, P. (2009). Harmonisation of knowledge management – comparing 160 KM frameworks around the globe. *Journal of Knowledge Management, 13*(4), 4-31. doi:10.1108/13673270910971798

Hema, R., & Malik, N. (2010). Data mining and business intelligence. In Proceedings of the 4th National Conference.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., Welch, V.A. and Cochrane, C. (2019). *Cochrane handbook for systematic reviews of interventions* [1 online resource (xxviii, 694 pages) : illustrations](Second edition. ed.). Retrieved from https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=2259855 [Accessed: 2021-05-25]

Hirji, K. K. (2001). Exploring data mining implementation. *Commun. ACM, 44*(7), 87–93. doi:10.1145/379300.379323

Hitt, M. A., Ireland, R. D., & Hoskisson, R. E. (2017). *Strategic management : Competitiveness & globalization Concepts* (12e. ed.). Australia ;: Cengage Learning.

Höpken, W., Fuchs, M., Keil, D., & Lexhagen, M. (2015). Business intelligence for cross-process knowledge extraction at tourism destinations. *Information Technology & Tourism, 15*(2), 101-130. doi:10.1007/s40558-015-0023-2

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC385300/

Huang, M. J., Chen, M.Y. & Lee, S.C. (2007). Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis. *Expert Systems with Applications,* 32(3):856-867.

Huang, Y., Liu, H., Keong Ng, W., Lu, W., Song, B., & Li, X. (2008). Automating knowledge acquisition for constraint-based product configuration. *Journal of*

Manufacturing Technology Management, 19(6), 744-754. doi:10.1108/17410380810888120

Hussinki, H., Ritala, P., Vanhala, M., & Kianto, A. (2017). Intellectual capital, knowledge management practices and firm performance. *Journal of Intellectual Capital, 18*(4), 904-922. doi:10.1108/JIC-11-2016-0116

Intezari, A., & Gressel, S. (2017). Information and reformation in KM systems: big data and strategic decision-making. *Journal of Knowledge Management, 21*(1), 71-91. doi:10.1108/JKM-07-2015-0293

Jantan, H., Hamdan, A.R., & Othman, Z.A. (2012). *Intelligent DSS for talent management: a proposed architecture using knowledge discovery approach*. Paper presented at Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication, Kuala Lumpur, Malaysia:1-7.

Jayanthi, R., & Vishal, B. (2011). Role of knowledge management and analytical CRM in business: data mining based framework. *The Learning Organization,* 18(2):131-148.

Kalkan, D. V. (2008). An overall view of knowledge management challenges for global business. *Business Process Management Journal, 14*(3), 390-400. doi:10.1108/14637150810876689

Karami, M., Alvani, S. M., Zare, H., & Kheirandish, M. (2015). Determination of Critical Success Factors for Knowledge Management Implementation, Using Qualitative and Quantitative Tools (Case study: Bahman Automobile Industry). *Iranian Journal of Management Studies, 8*(2), 181-201. doi:10.22059/ijms.2015.52588

Kenney, J. L., & Gudergan, S. P. (2006). Knowledge integration in organizations: an empirical assessment. *Journal of Knowledge Management, 10*(4), 43-58. doi:10.1108/13673270610679354

Khanbabaei, M., Alborzi, M., Sobhani, F.M., & Radfar, R. (2019). Applying clustering and classification data mining techniques for competitive and knowledge-intensive processes improvement. *Knowledge and Process Management,* 26(2):123-139.

King, E. (1999). Data warehousing and data mining : implementing strategic knowledge management. 1st ed. ed. Charleston, SC: Computer Technology Research Co.

Kishore, B., Kumar, R., Grover, R., & Maloo, A. (2014). Data mining usages in knowledge management: A Review Study. *International Journal of Scientific & Engineering Research,* 5(4).

Koç, T., Kurt, K., & Akbıyık, A. (2019). A Brief Summary of Knowledge Management Domain: 10-Year History of the Journal of Knowledge Management. *Procedia Computer Science, 158*, 891-898. doi:https://doi.org/10.1016/j.procs.2019.09.128

Kocher, M., & Riegelman, A. (2018). Systematic reviews and evidence synthesis: Resources beyond the health sciences. *2018, 79*(5). doi:10.5860/crln.79.5.248

Kokubo, A. (1993). Competitive intelligence. *IEEE Spectrum,* 30(8):44-46.

Kriegel, H.-P., Borgwardt, K. M., Kröger, P., Pryakhin, A., Schubert, M., & Zimek, A. (2007). Future trends in data mining. *Data Mining and Knowledge Discovery, 15*(1), 87-97. doi:10.1007/s10618-007-0067-9

Krzywicki, A., Wobcke, W., Bain, M., Calvo Martinez, J., & Compton, P. (2016). Data mining for building knowledge bases: techniques, architectures and applications. *The Knowledge Engineering Review, 31*(2), 97-123. doi:10.1017/S0269888916000047

Kumar, D., & Bhardwaj, D. (2011). Rise of data mining: current and future application areas. *International Journal of Computer Science Issues (IJCSI), 8*(5), 256.

Kvassov, V., & Madeira, S.C. (2004). Using data mining techniques for knowledge management: An empirical study.

Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proc. VLDB Endow., 5*(12), 2032–2033. doi:10.14778/2367502.2367572

Larose, D. T.,& Larose, C. D. (2014). Discovering knowledge in data: an introduction to data mining. John Wiley & Sons.

Lawal, N., Odeniyi, O., & Kayode, A. (2015). Application of data mining and knowledge management for business improvement: An exploratory study. *International Journal of Applied Information Systems,* 8:13-19.

Lea, B. R., Yu, W. B., Maguluru, N., & Nichols, M. (2006). Enhancing business networks using social network based virtual communities. *Industrial Management & Data Systems, 106*(1), 121-138. doi:10.1108/02635570610641022

Lee, S. M., & Hong, S. (2002). An enterprise-wide knowledge management system infrastructure. *Industrial Management & Data Systems, 102*(1), 17-25. doi:10.1108/02635570210414622

Leedy, P. D., Ormrod, J. E., & Johnson, L. R. (2021). *Practical research : planning and design* (Twelfth edition, Global edition. ed.). Harlow, England: Pearson.

Leeflang, M. M. G. (2014). Systematic reviews and meta-analyses of diagnostic test accuracy. *Clinical Microbiology and Infection, 20*(2), 105-113. doi:https://doi.org/10.1111/1469-0691.12474

Lewis-Beck, M. S., Bryman, A., & Futing Liao, T. (2004). The SAGE Encyclopedia of Social Science Research Methods. doi:10.4135/9781412950589

Li, Y., & Lu, Z. (2004). Ontology-based universal knowledge grid: Enabling knowledge discovery and integration on the grid. In: Zhang, L.J., Li, M., Sheth, A.P. & Jeffery, K.G. (eds.). *Proceedings - 2004 IEEE International Conference on Services Computing, SCC 2004*, Shanghai:557-560. [Online] Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-4644246800&partnerID=40&md5=b77466ddf451d3fee4972a0cf3b8dd67 [Accessed: 15 September 2004 through 18 September 2004].

Liao, S.-h. (2003). Knowledge management technologies and applications- literature review from 1995 to 2002. *Expert Systems with Applications,* 25(2):155-164.

Liao, S.-H., Chu, P.-H., & Hsiao, P.-Y. (2012). Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications, 39*(12), 11303-11311. doi:https://doi.org/10.1016/j.eswa.2012.02.063

Liebowitz, J. (2001). Knowledge management and its link to artificial intelligence. *Expert Systems with Applications, 20*(1), 1-6. doi:https://doi.org/10.1016/S0957-4174(00)00044-0

Liew, A. (2007). Understanding data, information, knowledge and their inter-relationships. *Journal of knowledge management practice, 8*(2), 1-16.

Liew, A. (2013). DIKIW: Data, information, knowledge, intelligence, wisdom and their interrelationships. *Business Management Dynamics, 2*(10), 49.

Littell, J. H. (2006). Systematic reviews in the social sciences: a review.

Littell, J. H. and Corcoran, J. (2010). Systematic Reviews. In B. A. Thyer (Ed.), *The Handbook of Social Work Research Methods* (Second Edition ed.). doi:10.4135/9781544364902

Liu, C.-H., Wang, J. S., & Lin, C.-W. (2017). The concepts of big data applied in personal knowledge management. *Journal of Knowledge Management, 21*(1), 213-230. doi:10.1108/JKM-07-2015-0298

Liu, D. R., & Hsu, C. (2004). Project-based knowledge maps: combining project mining and XML-enabled topic maps. *Internet Research, 14*(3), 254-266. doi:10.1108/10662240410542689

Luan, J. (2002). Data mining and knowledge management in higher education-Potential applications. *Workshop associate of institutional research international conference, Toronto.*

Lundvall, B.-Å. (1999). National business systems and national systems of innovation. *International Studies of Management & Organization*, 29(2):60-77.

Luo, Q. (2008). *Advancing Knowledge Discovery and Data Mining.* Paper presented at the First International Workshop on Knowledge Discovery and Data Mining (WKDD 2008).

Ma, L.-L., Wang, Y.-Y., Yang, Z.-H., Huang, D., Weng, H., & Zeng, X.-T. (2020). Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? *Military Medical Research, 7*(1), 7. doi:10.1186/s40779-020-00238-8

Madni, H. A., Anwar, Z., & Shah, M. A. (2017). *Data mining techniques and applications—a decade review.* Paper presented at the 2017 23rd International Conference on Automation and Computing (ICAC).

Maier, D., Kalus, W., Wolff, M., Kalko, S. G., Roca, J., Marin de Mas, I., Turan, N., Cascante, M., Falciani, F., Hernandez, M., Villà-Freixa, J. & Losko, S. (2011). Knowledge management for systems biology a general and visually driven framework applied to translational medicine. *BMC Systems Biology,* 5.

Maksood, F. Z., & Achuthan, G. (2016). Analysis of data mining techniques and its applications. *International Journal of Computer Applications,* 140(3):6-14.

Mallett, R., Hagen-Zanker, J., Slater, R., & Duvendack, M. (2012). The benefits and challenges of using systematic reviews in international development research. *Journal of Development Effectiveness, 4*(3), 445-455. doi:10.1080/19439342.2012.711342

Mamcenko, J., & Beleviciute, I. (2007). Data mining for knowledge management in technology enhanced learning. In Proceedings of the 6th WSEAS International Conference on Applications of Electrical Engineering:115-119.

Manjarres, A. V., Sandoval, L. G. M., & Suárez, M. S. (2018). Data mining techniques applied in educational environments: Literature review. *Digital Education Review*(33), 235-266.

Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., Quintana, M.J.R. and Flach, P. (2021). CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Transactions on Knowledge and Data Engineering, 33*(8), 3048-3061. doi:10.1109/TKDE.2019.2962680

Maule, R., Schacher, G., & Gallup, S. (2002). Knowledge management for the analysis of complex experimentation. *Internet Research, 12*(5), 427-435. doi:10.1108/10662240210447173

Mehmood, R., & Maurer, H. (2013). *Towards the integration of images on the Web*. Paper presented at Proceedings of International Conference on Information Integration and Web-based Applications; Services, Vienna, Austria:580-584.

Meier, M., & Mertens, P. (2000). Linking key figures and internet business news for personalized management information. *Journal of Systems and Information Technology, 4*(2), 13-32. doi:10.1108/13287260080000753

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information Technology and Organizational Performance: An Integrative Model of IT Business Value. *MIS quarterly, 28*(2), 283-322. doi:10.2307/25148636

Microsoft. (2018). *Analysis services documentation: Data mining algorithms (Data Mining)*. [Online] Available from: https://docs.microsoft.com/en-us/analysis-services/data-mining/data-mining-algorithms-analysis-services-data-mining [Accessed: 2019-10-15].

Mirza, S., Mittal, S., & Zaman, M. (2016). A review of data mining literature. *International Journal of Computer Science and Information Security (IJCSIS),* 14(11):437-442.

Mishra, P. C., Kishore, S., & Shivani, S. (2018). The Role of Information Technology for Knowledge Management: An Empirical Study of the Indian Coal Mining Industry. *Journal of Global Information Technology Management, 21*(3), 208-225. doi:10.1080/1097198X.2018.1498275

Mithas, S., Krishnan, M. S., & Fornell, C. (2005). Why do customer relationship management applications affect customer satisfaction? *Journal of Marketing,* 69(4):201-209.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A., and Group, P.-P. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*(1), 1. doi:10.1186/2046-4053-4-1

Moradi, M., Aghaie, A., & Hosseini, M. (2013). Knowledge-collector agents: Applying intelligent agents in marketing decisions with knowledge management approach. *Knowledge-Based Systems, 52*, 181-193. doi:https://doi.org/10.1016/j.knosys.2013.08.014

Mraović, B. (2008). Relevance of data mining for accounting: social implications. *Social Responsibility Journal, 4*(4), 439-455. doi:10.1108/17471110810909858

Mulrow, C. D. (1994). Systematic reviews: rationale for systematic reviews. *Bmj, 309*(6954), 597-599.

Natek, S., & Zwilling, M. (2014). Student data mining solution–knowledge management system related to higher education institutions. *Expert Systems with Applications,* 41(14):6400-6407.

Natek, S., & Lesjak, D. (2013). Improving Knowledge Management by Integrating Hei Process and Data Models. *Journal of Computer Information Systems, 53*(4), 81-86. doi:10.1080/08874417.2013.11645653

Natek, S., & Zwilling, M. (2014). Student data mining solution–knowledge management system related to higher education institutions. *Expert Systems with Applications, 41*(14), 6400-6407. doi:https://doi.org/10.1016/j.eswa.2014.04.024

Neaga, I., & Liu, S. (2014). The knowledge management context of cloud based big data analytics.1339.

Needleman, I. G. (2002). A guide to systematic reviews. *Journal of Clinical Periodontology, 29*(s3), 6-9. doi:https://doi.org/10.1034/j.1600-051X.29.s3.15.x

Negash, S., & Gray, P. (2008). Business intelligence. In: Burstein, F. & W. Holsapple, C. (eds.). *Handbook on Decision Support Systems 2: Variations*. Berlin, Heidelberg: Springer.

Nemati, H. R., Steiger, D. M., Iyer, L. S., & Herschel, R. T. (2002). Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing. *Decision Support Systems, 33*(2), 143-161. doi:https://doi.org/10.1016/S0167-9236(01)00141-5

Nguyen, N. B. C. (2018). *Data mining in knowledge management processes: developing an implementing framework.* Student thesis. [Online] Available from: http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-149668 [Accessed: 2019-06-25].

Nissen, M. E., & Bordetsky, A. (2011). Leveraging mobile network technologies to accelerate tacit knowledge flows across organisations and distances. In: Trentin, G. (ed.). *Technology and Knowledge Flow: Chandos Publishing.*

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis:Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods, 16*(1), 1609406917733847. doi:10.1177/1609406917733847

Obeidat, B. Y., Tarhini, A., Masa'deh, R. e., & Aqqad, N. O. (2017). The impact of intellectual capital on innovation via the mediating role of knowledge management: a structural equation modelling approach. *International Journal of Knowledge Management Studies, 8*(3-4), 273-298.

O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods, 19*, 1609406919899220. doi:10.1177/1609406919899220

Ode, E., & Ayavoo, R. (2020). The mediating role of knowledge application in the relationship between knowledge management practices and firm innovation. *Journal of Innovation & Knowledge, 5*(3), 210-218. doi:https://doi.org/10.1016/j.jik.2019.08.002

Oktari, R. S., Munadi, K., Idroes, R., & Sofyan, H. (2020). Knowledge management practices in disaster management: Systematic review. *International Journal of Disaster Risk Reduction, 51*, 101881. doi:https://doi.org/10.1016/j.ijdrr.2020.101881

Okumus, F. (2013). Facilitating knowledge management through information technology in hospitality organizations. *Journal of Hospitality and Tourism Technology, 4*(1), 64-80. doi:10.1108/17579881311302356

Orme, J., & Shemmings, D. (2010). *Developing research based social work practice.* Palgrave Macmillan.

Otto, J. R., Cook, J. H., & Chung, Q. B. (2001). Extensible markup language and knowledge management. *Journal of Knowledge Management, 5*(3), 278-285. doi:10.1108/13673270110401248

Parlby, D. (2000). Knowledge management research report 2000. *London: KPMG Consulting. Dixon, NM (2000). Common Knowledge*.

Partners, E. (2021). DistillerSR. Retrieved from https://www.evidencepartners.com/

Patrick, T. B., Demiris, G., Folk, L. C., Moxley, D. E., Mitchell, J. A., & Tao, D. (2004). Evidence-based retrieval in evidence-based medicine. *Journal of the Medical Library Association : JMLA, 92*(2), 196-199. Retrieved from https://pubmed.ncbi.nlm.nih.gov/15098048

Pawson, R., Boaz, A., Grayson, L., Long, A., & Barnes, C. (2003). Types and quality of social care knowledge. Stage two: towards the quality assessment of social care knowledge. *ESRC UK Center for Evidence Based Policy and Practice: Working Paper,* 18.

Pechenizkiy, M., Puuronen, S., & Tsymbal, A. (2005). Why data mining research does not contribute to business? In Proc. of Data Mining for Business Workshop DMBiz (ECML/PKDD'05), Porto, Portugal: 67-71.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences : a practical guide*. Malden, MA ;: Blackwell Pub.

Pires, V., & Trez, G. (2018). Corporate reputation A discussion on construct definition and measurement and its relation to performance. *Revista de Gestão, 25*(1), 47-64. doi:10.1108/REGE-11-2017-005

Pluye, P., & Hong, Q. N. (2014). Combining the Power of Stories and the Power of Numbers: Mixed Methods Research and Mixed Studies Reviews. *Annual Review of Public Health, 35*(1), 29-45. doi:10.1146/annurev-publhealth-032013-182440

Prasad, B. (2000). Converting computer-integrated manufacturing into an intelligent information system by combining CIM with concurrent engineering and knowledge management. *Industrial Management & Data Systems, 100*(7), 301-316. doi:10.1108/02635570010349104

Quan, X., Xiao, H., Ji, Q., & Zhang, J. (2021). Can innovative knowledge management platforms lead to corporate innovation? Evidence from academician workstations in China. *Journal of Knowledge Management, 25*(1), 117-135. doi:10.1108/JKM-12-2019-0684

Ranellucci, J., Poitras, E. G., Bouchet, F., Lajoie, S. P., & Hall, N. (2016). Chapter 5 - Understanding emotional expressions in social media through data mining. In:

Tettegah, S.Y. (ed.). *Emotions, Technology, and Social Media*. San Diego: Academic Press.

Ranjan, J., & Bhatnagar, V. (2011). Role of knowledge management and analytical CRM in business: data mining based framework. *The Learning Organization, 18*(2), 131-148. doi:10.1108/09696471111103731

Rao, G. K., & Dey, S. (2012). An intelligent decision making architecture for banks: Business intelligence and knowledge management systems integration. *Journal of Economic Development, Management, IT, Finance, and Marketing, 4*(1), 49.

Ribiere, V., & Calabrese, F. (2016). Why are companies still struggling to implement knowledge management? Answers from 34 experts in the field. In (pp. 13-34).

Richard, P. J., Devinney, T.M., Yip, G.S., & Johnson, G. (2009). Measuring organizational performance: towards methodological best practice. *Journal of Management,* 35(3):718-804.

Rigby, D., & Bilodeau, B. (2007). Bain's global 2007 management tools and trends survey. *Strategy & Leadership, 35*(5), 9-16. doi:10.1108/10878570710819161

Rigby, D., & Bilodeau, B. (2017). Management tools & trends 2017. *London: Bain & Company*.

Rolf, B. (2004). Two theories of tacit and implicit knowledge. *Blekinge Institute of Technology. Ronneby, Sweden*.

Rothberg, H. N., & Erickson, G. S. (2017). Big data systems: knowledge transfer or intelligence insights? *Journal of Knowledge Management, 21*(1), 92-112. doi:10.1108/JKM-07-2015-0300

Rowley, J. (2002). Eight questions for customer knowledge management in e-business. *Journal of Knowledge Management, 6*(5), 500-511. doi:10.1108/13673270210450441

Rygielski, C., Wang, J.-C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society, 24*(4), 483-502. doi:https://doi.org/10.1016/S0160-791X(02)00038-6

Sangiorgi, D., & Siboni, B. (2017). The disclosure of intellectual capital in Italian universities. *Journal of Intellectual Capital, 18*(2), 354-372. doi:10.1108/JIC-09-2016-0088

Sassenberg, C., Weber, C., Fathi, M., & Montino, R. (2009). A data mining based knowledge management approach for the semiconductor industry. *Proceedings of 2009 IEEE International Conference on Electro/Information Technology, EIT 2009*, Windsor, ON:72-77. [Online] Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-70449339752&doi=10.1109%2fEIT.2009.5189587&partnerID=40&md5=eb35e4469f7abeaf0680b7c1099a6f27 [Accessed: 9 June 2019].

Selamat, S. A. M., Prakoonwit, S., & Khan, W. (2020). A review of data mining in knowledge management: applications/findings for transportation of small and medium enterprises. *SN Applied Sciences, 2*(5), 1-15.

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A. and Group, T. P.-P. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ : British Medical Journal, 349*, g7647. doi:10.1136/bmj.g7647

Sharma, M. (2014). Data mining: A literature survey. International Journal of Emerging Research in Management & Technology, 3(2).

Shaw, M. J., Subramaniam, C., Tan, G. W., & Welge, M. E. (2001). Knowledge management and data mining for marketing. *Decision Support Systems, 31*(1), 127-137. doi:https://doi.org/10.1016/S0167-9236(00)00123-8

Silver, C., & Lewins, A. (2014). *Using Software in Qualitative Research: A Step-by-Step Guide*(Second Edition ed.). doi:10.4135/9781473906907

Silwattananusarn, T., & Tuamsuk, K. (2012). Data mining and its applications for knowledge management: a literature review from 2007 to 2012. *arXiv preprint arXiv:1210.2872*.

Singh, R., Iyer, L., & Salam, A. F. (2003). Web Service for Knowledge Management in E-Marketplaces. *e-Service Journal, 3*(1), 32-52. doi:10.2979/esj.2003.3.1.32

Singh, V., & Kumar, K. (2017). Data Mining and Knowledge Management. *Int. Res. J. Eng. Technol, 4*(3), 1-281.

Smith, E. A. (2001). The role of tacit and explicit knowledge in the workplace. *Journal of Knowledge Management, 5*(4), 311-321. doi:10.1108/13673270110411733

Smith, G. (2020). Data mining fool's gold. *Journal of Information Technology, 35*(3), 182-194.

Solomon, A., Ketikidis, P., & Choudhary, A. (2012). A knowledge based approach for handling supply chain risk management. *ACM International Conference Proceeding Series*, Novi Sad:70-75. [Online] Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84867339296&doi=10.1145%2f2371316.2371330&partnerID=40&md5=491772505faba25eae5a52a693e06bf4 [Accessed: 2019-04-10].

Spencer-Oatey, H., & Dauber, D. (2019). What is integration and why is it important for internationalization? A multidisciplinary review. *Journal of Studies in International Education*, 23(5):515-534.

Stanley, C. J. (2008). *A data mining study of the matriculation of Covenant college applicants*. Paper presented at Proceedings of the 46th Annual Southeast Regional Conference on XX, Auburn, Alabama:209-214.

Stern, C., Jordan, Z., & McArthur, A. (2014). Developing the Review Question and Inclusion Criteria. *AJN The American Journal of Nursing, 114*(4), 53-56. doi:10.1097/01.Naj.0000445689.67800.86

Su, Y.-S., & Wu, S.-Y. (2021). Applying data mining techniques to explore user behaviors and watching video patterns in converged IT environments. *Journal of Ambient Intelligence and Humanized Computing*, 1-8.

Sumbal, M. S., Tsui, E., & See-to, E. W. K. (2017). Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector. *Journal of Knowledge Management, 21*(1), 180-196. doi:10.1108/JKM-07-2016-0262

Sungkur, K. R., & Ramasawmy, M. (2014). Knowledge4Scrum, a novel knowledge management tool for agile distributed teams. *VINE, 44*(3), 394-419. doi:10.1108/VINE-12-2013-0068

Teece, D. J. (2007). Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal, 28*(13), 1319-1350. doi:https://doi.org/10.1002/smj.640

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), 509-533. doi:https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z

Terry, G., Hayfield, N., Clarke, V., and Braun, V. (2017). The SAGE Handbook of Qualitative Research in Psychology. In C. Willig & W. S. Rogers (Eds.). doi:10.4135/9781526405555

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(1), 45. doi:10.1186/1471-2288-8-45

Ting, S. L., Ip, W. H., Tsang, A. H. C., & Ho, G. T. S. (2012). An integrated electronic medical record system (iEMRS) with decision support capability in medical prescription. *Journal of Systems and Information Technology, 14*(3), 236-245. doi:10.1108/13287261211255347

Ting, S. L., Wang, W. M., Tse, Y. K., & Ip, W. H. (2011). Knowledge elicitation approach in enhancing tacit knowledge sharing. *Industrial Management & Data Systems, 111*(7), 1039-1064. doi:10.1108/02635571111161280

Tsai, H.-H. (2013). Knowledge management vs. data mining: research trend, forecast and citation approach. *Expert Systems with Applications,* 40(8):3160-3173.

Turulja, L., & Bajgoric, N. (2018). Information technology, knowledge management and human resource management. *VINE Journal of Information and Knowledge Management Systems, 48*(2), 255-276. doi:10.1108/VJIKMS-06-2017-0035

Uden, L., & He, W. (2017). How the Internet of Things can help knowledge management: a case study from the automotive domain. *Journal of Knowledge Management, 21*(1), 57-70. doi:10.1108/JKM-07-2015-0291

Usai, A., Pironti, M., Mital, M., & Aouina, M. C. (2018). Knowledge discovery out of text data: a systematic review via text mining. *Journal of Knowledge Management*, 22(7):1471-1488.

Victor, L. (2008). Systematic reviewing. *Social research update, 54*(1), 1-4.

Vinekar, V., Teng, J. T. C., & Chennamaneni, A. (2009). The Interaction of Business Intelligence and Knowledge Management in Organizational Decision-Making. *Journal of International Technology and Information Management, 18*(2), 143-159. Retrieved from https://www.proquest.com/scholarly-journals/interaction-business-intelligence-knowledge/docview/205859311/se-2?accountid=14717 [Accessed: 2021-05-05]

Wahab, A. and Rasha, S. (2012). Data mining's capabilities for knowledge creation in the GCC counties. *International Journal of Innovation and Knowledge Management in the Middle East and North Africa, 1*(2), 129-142. Retrieved from https://www.proquest.com/scholarly-journals/data-mining-s-capabilities-knowledge-creation-gcc/docview/2130299836/se-2?accountid=14717

Wahab, W. A., & Rahman, S. A. (2018). A Brief Review on the Knowledge Management and Data Mining for Marketing Decision. *International Journal of Academic Research in Business and Social Sciences, 8*(9).

Walsh, J. P., & Ungson, G.R. (1991). Organizational memory. *Academy of management review,* 16(1):57-91.

Wang, H., & Wang, S. (2008a). A knowledge management approach to data mining process for business intelligence. *Industrial Management & Data Systems,* 108(5):622-634.

Wang, H., & Wang, S. (2008b). Ontology for data mining and its application to mining incomplete data. *Journal of Database Management,* 19(4):81.

Wang, W.-C., Lin, C.-H., & Chu, Y.-C. (2011). Types of competitive advantage and analysis. *International Journal of Business and Management, 6*(5), 100.

Wang, Z., Brito, J. P., Tsapas, A., Griebeler, M. L., Alahdab, F., & Murad, M. H. (2015). Systematic reviews with language restrictions and no author contact have lower overall credibility: a methodology study. *Clinical epidemiology, 7*, 243-247. doi:10.2147/CLEP.S78879

Wee, J. C. N., & Chua, A. Y. K. (2013). The peculiarities of knowledge management processes in SMEs: the case of Singapore. *Journal of Knowledge Management, 17*(6), 958-972. doi:10.1108/JKM-04-2013-0163

West, L. A., & Hess, T. J. (2002). Metadata as a knowledge management tool: supporting intelligent agent and end user access to spatial data. *Decision Support Systems, 32*(3), 247-264. doi:https://doi.org/10.1016/S0167-9236(01)00102-6

Wild, R., & Griggs, K. (2008). A model of information technology opportunities for facilitating the practice of knowledge management. *VINE, 38*(4), 490-506. doi:10.1108/03055720810917732

Wong, B., Ho, G. T. S., & Tsui, E. (2017). Development of an intelligent e-healthcare system for the domestic care industry. *Industrial Management & Data Systems, 117*(7), 1426-1445. doi:10.1108/IMDS-08-2016-0342

Wu, W. W., Lee, Y. T., Tseng, M. L., & Chiang, Y. H. (2010). Data mining for exploring hidden patterns between KM and its performance. *Knowledge-Based Systems,* 23(5):397-401.

Xie, X., Zou, H., & Qi, G. (2018). Knowledge absorptive capacity and innovation performance in high-tech companies: A multi-mediating analysis. *Journal of Business Research, 88*, 289-297. doi:https://doi.org/10.1016/j.jbusres.2018.01.019

Xu, M., & Walton, J. (2005). Gaining customer knowledge through analytical CRM. *Industrial Management & Data Systems, 105*(7), 955-971. doi:10.1108/02635570510616139

Yamanishi, K., & Morinaga, S. (2005). Data mining for knowledge organization. *NEC journal of advanced technology,* 2(2):129-136.

Yasaka, N. (2017). Data mining in anti-money laundering field. *Journal of Money Laundering Control, 20*(3), 301-310. doi:10.1108/JMLC-09-2015-0041

Yoshioka, M., Tomioka, K., Hara, S., & Fukui, T. (2010). *Knowledge exploratory project for nanodevice design and manufacturing*. Paper presented at Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services, Paris, France:871-874.

Yu, X., Shi, Y., Zhang, L., Nie, G., & Huang, A. (2014). Intelligent knowledge beyond data mining: influences of habitual domains. *Communications of the Association for Information Systems, 34*(1), 53.

Zain, M. S. I. M., & Rahman, S. A. (2017). Challenges of Applying Data Mining in Knowledge Management towards Organization. *International Journal of Academic Research in Business and Social Sciences, 7*(12), 405-412.

Zakaria, N. and Zakari, N. (2016). Qualitative Content Analysis: A Paradigm Shift from Manual Coding to Computer-Assisted Coding Using ATLAS.ti. doi:10.4135/978144627305015599170

Žáková, M., Křemen, P., Železný, F., & Lavrač, N. (2011). Automating knowledge discovery workflow composition through ontology-based planning. *IEEE Transactions on Automation Science and Engineering,* 8(2):253-264.

Zanasi, A. (2007). Text mining and its applications to intelligence, CRM and knowledge management. Southampton, UK: WIT Press.

Zekić-Sušac, M., & Has, A. (2015). Data mining as support to knowledge management in marketing. Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy, 6(2), 18-30.

Zeng, X., Zhang, Y., Kwong, J. S., Zhang, C., Li, S., Sun, F., Niu, Y., & Du, L. (2015). The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. *Journal of evidence-based medicine, 8*(1), 2-10.

Zhang, H., & Liang, Y. (2006). A knowledge warehouse system for enterprise resource planning systems. Systems Research and Behavioral Science: The Official Journal of the International Federation for Systems Research, 23(2), 169-176.

List of appendices

Appendix A Evidence of the adapted PRISMA-P Checklist in the review

Section		Checklist Item	Reported in Article
Administrative Information			
Identification	1a	Identify the report as a protocol of a systematic review	Title page
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	N/A
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number	N/A
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	Title page
Contribution	3b	Describe contributions of protocol authors and identify the guarantor of the review	Title page
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	N/A
Sources	5a	Indicate sources of financial or other support for the review	N/A
Sponsor	5b	Provide name for the review funder and/or sponsor	N/A
Role of sponsor/ funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	N/A
Introduction			
Rationale	6	Describe the rationale for the review in the context of what is already known	Chapter 1: Introduction
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	Chapter 1: Purpose of the study
Method			
Eligibility Criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics	Chapter 3: Methodology Eligibility Criteria

Section		Checklist Item	Reported in Article
		(e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	
Information Sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	Chapter 3: Methodology Sampling in Databases
Search Strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	Chapter 3: Methodology Search Strategy
Data Management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	Chapter 3: Methodology Data Management
Selection Process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	Chapter 3: Methodology Study Selection
Data Collection Process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	Chapter 3: Methodology Data Extraction
Data Items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	Chapter 3: Methodology Eligibility Criteria and Data Extraction
Outcomes and Prioritisation	13	List and define all outcomes for which data will be sought, including prioritisation of main and additional outcomes, with rationale	Chapter 3: Methodology Data Analysis
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	Chapter 3: Methodology Data Extract and Risk of bias
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	N/A
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I2, Kendall’s tau)	N/A
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	N/A

Section		Checklist Item	Reported in Article
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	Chapter 3: Methodology Data Analysis
Meta-bias(ses)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	Chapter 3: Methodology Data Extraction and Risk of Bias
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	Chapter 3: Methodology Data Extraction and Risk of Bias; Reliability, Validity and Trustworthiness

Appendix B Data Extraction Form

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
A framework of intelligent decision support system for Indian police	Gupta M., Chandra B., Gupta M.P.	Journal of Enterprise Information Management	2014	Quantitative	Knowledge Acquisition	Clustering, Classification and Association Rules
A knowledge management approach to data mining process for business intelligence	Wang, H. and Wang, S.	Industrial Management & Data Systems	2008	Qualitative	Human Knowledge	DM Centred Cycle
A knowledge management framework using business intelligence solutions	Gadu, M. and El-Khameesy, N.	International Journal of Computer Science Issues (IJCSI)	2014	Qualitative	Knowledge sharing	Both predictive and descriptive DM
A model of information technology opportunities for facilitating the practice of knowledge management	Wild, R. and Griggs, K.	VINE	2008	Qualitative	KM Life Cycle and knowledge generation	DM DSS
An enterprise-wide knowledge management system infrastructure	Lee, S.M. and Hong, S.	Industrial Management & Data Systems	2002	Qualitative	Knowledge development	DM life Cycle
An integrated electronic medical record system (iEMRS) with decision	Ting, S.L., Ip, W.H., Tsang, A.H.C. and Ho, G.T.S.	Journal of Systems and Information Technology	2012	Quantitative	Knowledge generation	Decision Rules

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
support capability in medical prescription						
An Intelligent Decision-Making Architecture for Banks: Business Intelligence And Knowledge Management Systems Integration	Rao, G. K. and Dey, S.	Journal of Economic Development, Management, IT, Finance and Marketing	2012	Qualitative	Knowledge acquisition	DM BI
An ontology-based business intelligence application in a financial knowledge management system	Cheng, H., Lu, Yi-Chuan and Sheu, C.	Expert Systems with Applications	2009	Quantitative	Knowledge creation	Classification
Application of knowledge management technology in customer relationship management	Bose, R. and Sugumaran, V.	Knowledge and Process Management	2003	Quantitative	Knowledge life cycle	DM XML and predictive modelling
Artificial Intelligence in Knowledge Management: Overview and Trends	Birzniece, Ilze	Computer Science	2011	Qualitative	KM Framework	Artificial neural networks
Automated knowledge discovery in advanced knowledge management	Grobelnik, M. and Mladenić, D.	Journal of Knowledge Management	2005	Qualitative	KM Life Cycle	Ripple-down-rules learning algorithm
Automating knowledge acquisition for constraint-based product configuration	Huang, Y., Liu, H., Keong Ng, W., Lu, W., Song, B. and Li, X.	Journal of Manufacturing Technology Management	2008	Quantitative	Knowledge generation	Association rule mining

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
Big data systems: knowledge transfer or intelligence insights?	Rothberg, H.N. and Erickson, G.S.	Journal of Knowledge Management	2017	Qualitative	KM Life Cycle	Big Data Analytics
Business intelligence for cross-process knowledge extraction at tourism destinations	Höpken, W., Fuchs, M., Keil, D. and Lexhagen, M.	Information Technology & Tourism	2015	Quantitative	Knowledge extraction	Data stemming
Converting computer-integrated manufacturing into an intelligent information system by combining CIM with concurrent engineering and knowledge management	Prasad, B.	Industrial Management & Data Systems	2000	Qualitative	KM Life Cycle	AI
Data Mining as Support to Knowledge Management in Marketing	Zekić-Sušac, M. and Has, A.	Business Systems Research	2015	Qualitative	Knowledge creation	Association rules, Neural networks
Data mining for building knowledge bases: Techniques, architectures and applications	Krzywicki, A., Wobcke, W., Bain, M., Calvo Martinez, J., & Compton, P.	The Knowledge Engineering Review	2016	Quantitative	Knowledge extraction	Text mining
Data mining in anti-money laundering field	Yasaka, N.	Journal of Money Laundering Control	2017	Qualitative	Knowledge sharing	DM life Cycle

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
Data mining's capabilities for knowledge creation in the GCC counties	Wahab, A. and Rasha, S.	International Journal of Innovation and Knowledge Management in the Middle East and North Africa	2012	Qualitative	Knowledge creation	DM life Cycle
Data-driven decision-making for the enterprise: an overview of business intelligence applications	Hedgebeth, D.	VINE	2007	Quantitative	KM Life Cycle	BI DSS
Decision Support System and Knowledge-based Strategic Management	Alyoubi, Bader A.	Procedia Computer Science	2015	Qualitative	KM Life Cycle	DSS and GSS
Development of an intelligent e-healthcare system for the domestic care industry	Wong, B., Ho, G.T.S. and Tsui, E.	Industrial Management & Data Systems	2017	Quantitative	Knowledge creation	Fuzzy Association Rule Mining (FARM) and Rule mining
Eight questions for customer knowledge management in e-business	Rowley, J.	Journal of Knowledge Management	2002	Qualitative	Generating customer knowledge	DM DSS
Enhancing business networks using social network based virtual communities	Lea, B., Yu, W., Maguluru, N. and Nichols, M.	Industrial Management & Data Systems	2006	Qualitative	KM Life cycle	DM and text mining

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
Extensible markup language and knowledge management	Otto, J.R., Cook, J.H. and Chung, Q.B.	Journal of Knowledge Management	2001	Qualitative	KM Life Cycles	XML Markup
Facilitating knowledge management through information technology in hospitality organisations	Okumus, F.	Journal of Hospitality and Tourism Technology	2013	Qualitative	Knowledge sharing and dissemination	DM life Cycle
Gaining customer knowledge through analytical CRM	Xu, M. and Walton, J.	Industrial Management & Data Systems	2005	Qualitative	Customer knowledge acquisition	Pattern discovery association rules, sequential patterns; clustering, classification and prediction
How the Internet of Things can help knowledge management: a case study from the automotive domain	Uden, L. and He, W.	Journal of Knowledge Management	2017	Qualitative	Knowledge creation	IoT DM
Improving Knowledge Management by Integrating Hei Process and Data Models	Natek, S. and Lesjak, D.	Journal of Computer Information Systems	2013	Quantitative	KM life cycle, Knowledge process	DM data modelling

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
Information and reformation in KM systems: big data and strategic decision-making	Intezari, A. and Gressel, S.	Journal of Knowledge Management	2017	Qualitative	KM Life Cycle and systems	DM Life Cycle and big data
Integrating web-based data mining tools with business models for knowledge management	Heinrichs, John H. and Lim, Jeen-Su	Decision Support Systems	2003	Quantitative	Knowledge creation	Web-based DM
Intelligent Knowledge Beyond Data Mining: Influences of Habitual Domains	Yu, Xiaodan; Shi, Yong; Zhang, Lingling; Nie, Guangli; and Huang, Anqiang	Communications of the Association for Information Systems	2014	Quantitative	KM and human knowledge	DM life cycle
Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector	Sumbal, M.S., Tsui, E. and See-to, E.W.K.	Journal of Knowledge Management	2017	Qualitative	KM Life Cycle	Prediction
Knowledge elicitation and mapping in the design of a decision support system for the evaluation of suppliers' competencies	Cannavacciuolo, L., Iandoli, L., Ponsiglione, C. and Zollo, G.	VINE	2015	Quantitative	KM Life Cycle	Fuzzy Logic
Knowledge elicitation approach in enhancing tacit knowledge sharing	Ting, S.L., Wang, W.M., Tse, Y.K. and Ip, W.H.	Industrial Management & Data Systems	2011	Quantitative	Knowledge sharing	Association rule mining technique

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
Knowledge integration in organisations: an empirical assessment	Kenney, J.L. and Gudergan , S.P.	Journal of Knowledge Management	2006	Qualitat ive	Knowledge integration	Data mining software
Knowledge management and data mining for marketing	Shaw, M. J., Subramaniam, C., Tan, Gek Woo and Welge, M. E.	Decision Support Systems	2001	Qualitat ive	KM Life Cycle	DM Life Cycle
Knowledge management and its link to artificial intelligence	Liebowitz, J.	Expert Systems with Applications	2001	Qualitat ive	KM Life Cycle	AI
Knowledge management for the analysis of complex experimentation	Maule, R., Schacher, G. and Gallup, S.	Internet Research	2002	Qualitat ive	Knowledge generation	AI
Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing	Nemati, H. R., Steiger, D. M., Iyer, L. S. and Herschel, R. T.	Decision Support Systems	2002	Qualitat ive	Knowledge generation	DM, AI and DSS
Knowledge4Scrum, a novel knowledge management tool for agile distributed teams	Sungkur, K. R. and Ramasaw my, M.	VINE	2014	Quantit ative	KM Life Cycle	Clustering, Classification tree
Knowledge-collector agents: Applying intelligent agents in marketing	Moradi, M., Aghaie, A. and Hosseini, M.	Knowledge-Based Systems	2013	Quantit ative	KM Life Cycle	Fuzzy AHP logic, Decision

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
decisions with knowledge management approach						support system
Linking key figures and internet business news for personalised management information	Meier, M. and Mertens, P.	Journal of Systems and Information Technology	2000	Quantitative	Knowledge creation	Text mining
Managing extracted knowledge from big social media data for business decision-making	He, W., Wang, F.-K. and Akula, V.	Journal of Knowledge Management	2017	Qualitative	KM Life Cycle	Text mining
Metadata as a knowledge management tool: supporting intelligent agent and end user access to spatial data	West, L. A. and Hess, T. J.	Decision Support Systems	2002	Quantitative	KM Life Cycle	DM DSS
Project-based knowledge maps: combining project mining and XML-enabled topic maps	Liu, D. and Hsu, C.	Internet Research	2004	Quantitative	Knowledge acquisition	Clustering and Association rule Mining
Relevance of data mining for accounting: social implications	Mraović, B.	Social Responsibility Journal	2008	Qualitative	Knowledge creation	DM Life Cycle
Role of knowledge management and analytical CRM in business: data mining based framework	Ranjan, J. and Bhatnagar, V.	The Learning Organisation	2011	Qualitative	KM Life Cycle	DM life cycle

Article Title	Author	Journal Title	Year of Publication	Design	Knowledge Types/Task	DM Task
Student data mining solution–knowledge management system related to higher education institutions	Natek, S. and Zwilling, M.	Expert Systems with Applications	2014	Quantitative	Knowledge creation	DM Life Cycle
The concepts of big data applied in personal knowledge management	Liu, C.-H., Wang, J.S. and Lin, C.-W.	Journal of Knowledge Management	2017	Qualitative	KM Life Cycle	Big Data
The integration of business intelligence and knowledge management	W. F. Cody, J. T. Kreulen, V. Krishna and W. S. Spangler	IBM Systems Journal	2002	Quantitative	KM Life Cycle	Classification and OLAP
The Interaction of Business Intelligence and Knowledge Management in Organisational Decision-Making	Vinekar, V., Teng, J.T.C. and Chennamaneni, A.	Journal of International Technology and Information Management	2009	Qualitative	KM Life Cycle	DM Life Cycle
The role of AI-based technology in support of the knowledge management value activity cycle	Fowler, A.	The Journal of Strategic Information Systems	2000	Qualitative	Knowledge acquisition and organisation	AI and Neural Network
Web Service for Knowledge Management in E-Marketplaces	Singh, R., Iyer, L., & Salam, A. F.	E-Service Journal	2003	Qualitative	Knowledge sharing	XML

Appendix C List of Codes and Code Groups

Code	Code Groups
A data warehouse should always provide its users with accurate, consistent, and real-time data. It should be ffexible to support all corporate operations and changes.	Data warehouse
A DM result might not trigger an action, but can be learned by business insiders to develop tacit knowledge about the interesting patterns of data. Internalisation is the process of transformation from a DM result to tacit knowledge.	DM result to knowledge
A knowledge repository of previous data mining efforts could be used to store knowledge of 'normal' data so that outliers are more effectively identified in the data preparation stage. Knowledge repositories could also be used to store knowledge of application goals, which could greatly aid in directing the data mining effort.	Knowledge repository
Aggregation of human knowledge and feedback into KB construction seems inevitable despite the trend towards automation.	Automation
AI-based technology, on its own, does not provide a unique solution to organisational KM needs. It still does not substitute for human intelligence and possesses only a limited capability to address the issue of tacit knowledge.	Knowledge acquisition for Expertise
Amarket basket analysisB w x 3 and gives us the relationship between different products purchased by a customer. This type of knowledge can be useful in developing marketing strategies for promoting products that have dependency relationships in the minds of the customers.	Customer relationship management
An action must have its outcome. An action outcome is the assessment of the business decision and its execution in terms of tangible and intangible costs and beneffts. Metrics and measures of costs and beneffts are shared by the organisation.	Measurement
An enterprise-wide KM system infrastructure.	Enterprise wide knowledge
The majority of models used in the KM ffeld, such as the tacit and explicit knowledge framework for a dynamic human process of justifying personal belief toward the truth, are typically non-technology oriented.	Non-technology driven
As the other drawbacks of ES towards KM the following issues are named: ES are unable to respond to vague questions or give vague answers; There are difficulties in the maintenance and updating of the knowledge base and learning from experience.	Expert system
As thousands of knowledge sets gradually piled up in the knowledge database, intelligent screening of ontology, spontaneous push-and-pull knowledge dissemination, and performance ranking will essentially and inevitably lead the way to more powerful knowledge generation.	Knowledge base
As with any KM-related activity, support from top management is neces	Management support

Code	Code Groups
AR scenario: There are condition-outcome relationships among attribute/value pairs, if a consumer purchases A, then she/he is likely to purchase B.	DM tasks
Besides the signiffcant of the approach in product conffguration, the rule acquired can be used to assist the new product design. It can be enhanced to generate the design knowledge for a portion of the new design that has been existed or similar to the existing modules.	Design knowledge
BI and KM integration assists today's managers in improving/optimising decision-making process by sharing data and information across the organisation; getting the details from internal and external sources; and forecasting the future trend and taking better decisions.	BI and KM integration
BI is a part of KM	BI and KM integration
Big data applications could broaden the scope of PKM that transforms “data” into “information”, and ffnally into “knowledge”, to contribute the development path of PKM utilisation.	Big data
Big data can be understood primarily as the huge amounts of data generated by the Internet, mobile phones, tablets, etc. They are unable to be stored and retrieved by a traditional database management system. From a management perspective, it also can be characterised by volume (about 2.5 exabyte of data were created every day in 2012 and it seems to be much more nowadays), velocity (real-time information creates competitive advantages) and variety (data from different sources)	Big data
Big Data provides new opportunities for extending the utilisation KM.	Big data
Big Data technology is used as a solution to analyse the big social media data related to the organisations and their competitors, and to visualise and benchmark comparisons among competitors across events, products, issues and any other areas that may affect business performance.	Big data
Big social media data can play a critical role in the development of new business knowledge and increase the opportunity for organisations to discover BI for decision-making.	Big data
BI helps with the Intelligence phase – identifying internal strengths and weaknesses, or identifying external threats and opportunities.	BI and KM integration
The multiple data formats and distributed nature of the knowledge on the Web make it a challenge to collect, discover, organise and manage the knowledge in a manner that is useful for marketing decision support.	Web knowledge
By using a common tool that is accessible through the corporate intranet, where the combined knowledge of the system is incorporated and used by analysts across the enterprise, planners can generate knowledge about the impact of a decision (a scenario deffned by the combination of simulation input factors).	Knowledge repository
Classiffcation: An observation with a certain set of attributes can be assigned to a class A ffrm with certain attributes is likely to bankrupt.	DM tasks

Code	Code Groups
Classification algorithms, namely decision tree, Artificial Neural Network (ANN) approach, Bayesian approach, etc. have also been stored in the model base for catering to various applications in crime DM. Crime zones can be treated as classes for supervised learning. The various classification models have been applied as per the application areas.	DM tasks
Cluster analysis There are distinct segments in a given set of observations Certain consumers form a market segment.	DM tasks
Clustering algorithms are used for identifying crime zones and crime hot spots. Parametric Minkowski distance measure have also been applied to introduce the weightage scheme in the clustering process since all types of crime do not have equal weights: murder, for example, will have more weights over kidnapping and hurt. These weights can either be obtained by the user's input or by applying standard algorithms of weight generation and selection.	DM tasks
Companies which use the IoT can gather data about how their products behave and interact, and can then use it to understand and predict future behaviours.	Big data
Competitive Intelligence	IT enablement Competitive intelligence
Conversely, tacit knowledge is more readily apparent in the CI metric. CI operations gather inputs from all kinds of intangibles, but they are most interested in new developments and need their own analytic skills to spot patterns in their inputs – suggesting the presence of important tacit knowledge in what they are studying and in their own activities.	Competitive intelligence
Core to CI is the ability to deductively address speciffc questions and inductively piece together competitive landscapes.	Competitive intelligence
CRM	Customer relationship management
CRM provides real-time information about customer's buying patterns, pre-and post-sales behaviour and factors for customer retention.	Customer relationship management
Customer data	Customer relationship management
Unlike the cvonventional database, that is mainly designed for operations in business, a data warehouse- a centralised read only database often remote, containing snapshots of corporate data, is one of the essential IT application supporting knowledge capture.	Data warehouse for knowledge capture
Data and corresponding semantic structures change in time. As the consequence, we need to be able to adapt ontologies that are modeling the data accordingly – we call this kind of structures ‘‘dynamic ontologies’’. For most of such scenarios an extensive human involvement in building models from the data is not economical anymore, since it gets too costly, too inaccurate and too slow.	Automation

Code	Code Groups
Data driven extraction	Data driven activities
Data integration	System Integration
DM for small student data sets from research is relevant example of effective use of DM technology to develop the HEI's KM systems in the education domain.	DM tasks
DM poses numerous questions relating to privacy, legality and ethics. The trouble is that there is a permanent threat of using DM applications beyond the limits of its original purposes. ''Knowing about ... '' represents information.	DM challenges
DM poses numerous questions relating to privacy, legality and ethics (Clark, 2003). Each time we participate in a telephone survey, ffll in a credit form, visit a doctor, every web site we visit. . . it all goes into the databases. Pessimists would say: Is that not that scary? It may be an exaggerated attitude, but knowing a man gives one power over the man, and the possibilities of misuse are huge. This concerns primarily the privacy implications. The trouble is that there is a permanent threat of using DM applications beyond the limits of its original purposes. Therefore it is particularly important, believes Seifert (2004), to deffne the extent and set boundaries in which government agencies can mix commercial data with government data.	DM challenges
DM techniques can help us reveal patterns and relationships, but they cannot tell us the true value or signiffcance of these ffndings. These assessments still remain a prerogative of man and his willed action	DM tasks
DM tools contribute to efforts for wide integrations of data by being supported by a strong and mature technological infrastructure made up of three core components: 1. massive data collection; 2. powerful multiprocessor computers; and 3. DM algorithms.	System Integration
Data regarding the internal and external factors, which are related to the organisation and the environment, and the processing of such data, are therefore vital for making strategic decisions.	Data driven activities
Data warehousing technology can help to make by integrating various sub-systems into a BI framework.	Data warehouse
Developing such a system builds on an enterprise wide integration of technologies working together such as data warehouse, web site, intranet/extranet, phone support systems, accounting, sales, marketing and production.	System Integration
Deviation analysis There is an abnormal observation in a set of observations A competitor is taking an abnormal action.	DM tasks
DM can be beneffcial for KM in the following two major aspects: (1) To share common understanding of the context of BI among data miners. For example, given a marketing survey database, the data miners share the scope of the database, the deffnitions of the data items, the metadata of the database, and the a priori knowledge of DM techniques to be applied to the database.	DM result to knowledge

Code	Code Groups
DM constitutes one step in the knowledge discovery process. It is in data mining step that the actual search for patterns of interest is performed. It is important at this stage to choose the appropriate DM algorithm (NNs, linear/logistic regression, association rules, etc.) for the DM task.	DM tasks
DM is to reveal interesting patterns in the data to verify a hypothesis or hypotheses for the data miners. A hypothesis mirrors a priori knowledge (or seed knowledge) for DM. A DM algorithm is designed to verify a speciffc type of hypothesis. Typical categories of DM algorithms, their corresponding general types of hypotheses for DM, and examples of seed knowledge are summarised.	DM tasks
DM techniques can be more hazardous than helpful if the frontline users do not fully understand how to apply those techniques in pertinent context.	DM tasks
Domain experts need a ffexible system environment where they can freely select data and models and run different settings of parameters for decision support purposes.	Knowledge acquisition for Expertise
DSS can also enhance the tacit to explicit knowledge conversion by eliciting one or more what-if cases (i.e., model instances) representing situations that the knowledge worker wants to explore. As the knowledge worker changes one or more model coefficients or right hand side values (e.g., in a linear programming model) to explore its effect on the modelled solution, s/he is estimating ranges of those parameters/values that reflect the actual and/or potential decision-making environment represented by the model.	Decision Support
DSS has made significant research contributions in knowledge extraction/acquisition with knowledge engineers of expert systems and the mathematical models of management scientists.	Decision Support
Dynamic capabilities should allow hospitality businesses to make timely decisions, reduce cost, improve quality and launch new and better products and services before their competitors.	Decision Support
Employees may analyse potential growth and profitability of customers, and reduce the risk exposure through more accurate financial credit scoring of their customers.	Customer relationship management
Exchange of data	Data driven activities
Existing Big Data platforms from vendors like IBM, SAP, Oracle and Microsoft can be integrated to store, manage, analyse and compare data across numerous social media sources. Such comparisons are supported by Big Data analytics and contextual data found from big social media data. Extracted knowledge will be stored into a knowledge repository on a daily basis and shared with managers and other employees in the organisation through the organisation’s existing KMS.	Big data
Expertise development	Knowledge acquisition for Expertise
Explicit knowledge is generated with the help of OLAP, DM and other reporting tools. Generated knowledge must be filtered, organised, and stored in a central knowledge repository to make them available efficiently and	Knowledge repository

Code	Code Groups
effectively. Tacit knowledge can be directly captured by KM system from the human experts.	
Externalisation or Articulation is the conversion of the tacit knowledge to explicit knowledge. It occurs through and is largely facilitated by the DSS systems that an organisation utilises.	Decision Support
First is the knowledge acquisition process that elicits knowledge automatically from the EMR, capturing the physician's decision logic for the medical treatment. This method of automatic knowledge acquisition is an indirect approach that does not require any human intervention.	Automation
First, people often ffnd that "knowledge" gained from DM does not always lead to an action in all situations, particularly when the piece of "knowledge" is hard to apply.	DM result to knowledge
for the KB to be used effectively in knowledge mining, the extracted knowledge must be sufffciently reliable. Thus, some form of human input is typically incorporated into the process. Human input typically provides background knowledge in the form of curation rules to improve the quality of the KB, and human-labelled examples to improve supervised machine learning processes.	Knowledge base
From the KM perspective, expanding the KM scope to include massive amounts of social media data is inevitable for companies to generate organisational knowledge for business purposes. To better manage extracted knowledge from the big social media data and derive more business value, there is also a strong need for forward-thinking companies to build Big Data analytics capabilities.	Big data
Fuzzy Association Rule Mining (FARM) was deployed in the e-healthcare system for elderly care services developed in this research in order to have a more realistic and practical classification of the data attributes in their relationships.	DM tasks
There still is a lack of attention on theories and models of DM for knowledge development in business. Conventional theories and models in this area ought to be re-examined and developed in such a way that a distinction is made between two important variables: DM centred information and business centred knowledge.	DM result to knowledge
Hypotheses pertinent to business actions are always depending upon the knowledge sharing among data miners and business insiders.	Knowledge acquisition for Expertise
ICT components	IT enablement
If we can put together some interesting metrics, understand what they mean in terms of the contribution of different intangibles (Big Data/information, knowledge, intelligence) to competitiveness and then allow speciffc ffrms in speciffc industries to evaluate their own circumstances, we can better advise practitioners on what intangibles to apply in what manner to what situation. Indeed, another core contribution from KM/IC is matching the right tools (e.g. communities of practice for tacit-to-tacit exchange) to the right circumstance, and this approach expands that capability.	Competitive intelligence

Code	Code Groups
Improved quality of elderly care service: apart from the fast intervention of the caregivers to support the services in responding to the instant alerts triggered by the e-healthcare system in this study, the development of the data collection through IoTs can replace part of the regular work of the caregivers in vital sign measurements.	Big data
In a time of globality, companies operate in a business environment in which high speed dictates mutual interactions and, at the same time, requires a radically different approach to data collection, storage, and processing based on an integrated view of the data.	Big data
In association rule mining, support and conffdence are commonly used as indicators to assess the probabilities of rules. Support determines how often a rule is applicable to a given data set, while conffdence determines how frequently one item appear in transactions that contain another item. Rules with sufffcient level of support and conffdence that are higher than deffned thresholds are considered interested rules.	DM tasks
In order to improve the efficiency of decision-making and to adapt to changing environment and markets, KM and BI need to be integrated.	BI and KM integration
In the case of medical prescription, physicians rely heavily on their knowledge and experiences to select appropriate medicine. It is almost impossible for a physician to utilise only their individual knowledge to consider all the important differences between current and former similar cases.	Knowledge acquisition for Expertise
In the future, Big Data will be able to develop a system that could mine knowledge-based on automatically learning and classifying artiffcial intelligence or other advance technologies.	Big data
In the long run, however, for KM to be truly successful, it needs to be integrated within the overall strategic vision of the organisation.	System Integration
In this study, an integrated EMRS with prescription decision support capability, abbreviated as iEMRS, is proposed to combine the concept of clinical DSS (CDSS) and traditional electronic medical record system in providing knowledge that can serve as references for quality decisions in prescription.	Decision Support
Information and communications technology approach to EMRS is accepted in hospitals as well as clinical functions and logistics activities to enhance the transmission of patient medical records and communication among physicians.	IT enablement
Integrated BI and KM provide the robust system with the capability of process-driven decision-making.	BI and KM integration
Integrative flow of knowledge to inform e-business processes is a challenging task with significant potential competitive benefits for organisations that exhibit technical and managerial leadership.	Competitive intelligence
Integrative technologies allow firms to create coordination structures that mitigate the risk inherent in dynamic market conditions through the availability of accurate information and specific knowledge for the planning process to e-marketplace participants. Additionally, these technologies allow organisations to coordinate activities and exchange information, including demand	System Integration

Code	Code Groups
conditions and supply capabilities, across collaborating firms in upstream and downstream marketplaces.	
intelligence amplification by machine and mind can outperform a mind-imitating AI system working by itself. The combination of formal, explicit knowledge in the machine, and the non-formal, tacit knowledge of the users, can thereby result in problem-solving capabilities which surpass either one of these components acting alone.	Knowledge acquisition for Expertise
Intelligent support systems can play a vital role in improving outcome in the crime investigation, criminal detection and other major areas of functioning of police organisation by facilitating recording, retrieval analysis and sharing of the information.	Decision Support
Intensive competitive intelligence activities could empower businesses to develop new products and services, optimise business processes, enhance value creation and foster innovation. We believe that the increase in business knowledge and innovation will likely lead to an increase in economic activities which will have a signiffcant impact on today's knowledge economy.	Competitive intelligence
Interactions between people that traditionally take place face to face are now conducted online using convergent synergy of web conferencing, real-time collaboration technologies, instant messaging, shared online work spaces, and interactive white boards through the Internet, such as HTTP protocols.	Big data
Internet-based technologies have great ability to support the essential tasks of KM, including knowledge acquisition, organisation and distribution that foster the demand of virtual organisations. Accordingly, Internet-based technologies provide an accessible, scalable, and effective platform to support KM functionality, and to consolidate project management activities.	IT enablement
IT applications can be used in coding and formalising this knowledge as well as in disseminating it within the organisation.	IT enablement
IT enablement	IT enablement Customer relationship management
It is important to note that DSS, and more specifically GSS facilitate more efficient KM in each of the four steps in the KM cycle.	Decision Support
It proposes a knowledge sharing model for business knowledge workers to make DM more relevant to BI.	Knowledge acquisition for Expertise
KM application development	Knowledge acquisition for Expertise
KM system	KM system flexibility
KM technologies coupled with DM principles are capable of managing the ffrm's explicit and implicit knowledge.	DM result to knowledge

Code	Code Groups
Knowledge base is divided into different categories, like product's knowledge, competitor's knowledge, and customer's knowledge. This layer also contains a knowledge map. The knowledge map shows which knowledge is used, this knowledge is useful for what specialist, what is its relationship with other knowledge and etc. With use of knowledge map, agents can retrieve relevant knowledge for decision-makers more effectively, because the knowledge map shows the relationship between knowledge and their usage.	Knowledge base
KBs are becoming indispensable tools for automatic understanding of human-generated data, summarisation and question answering.	Knowledge base
Knowledge creation is driven by providing incentive and greater investment in research and development as well as by building technological capacity to ensure access to knowledge.	IT enablement
Knowledge development	Knowledge acquisition for Expertise
Knowledge engineering is concerned with eliciting rules and facts from experts.	Knowledge base
Knowledge extracted from BI can enable the implementation process of innovative products/services.	Knowledge base
Knowledge sharing in the workplace is often implemented with the ass	IT enablement
Knowledge sharing in the workplace is often implemented with the assistance of IT, through digitised filming of the physical demonstration of a process.	IT enablement
Knowledge workers	Knowledge acquisition for Expertise
Many companies today realise that they must become knowledge-creating companies or learning organisations in order to survive and flourish in a rapidly changing business environment. That means constantly creating new business knowledge, disseminating it within the organisation, and quickly building it into new products and services.	Knowledge base
Many of these functions require extensive knowledge and analytical ability gained only from years of experience. Without this core knowledge and the benefft of past trial and error, process managers make less informed and potentially costly decisions.	Knowledge acquisition for Expertise
Market research databases. These databases can contain customer satisfaction information and quality performance information on products and services that the firm provides.	Knowledge base
Marketing concepts	Customer relationship management
New cases may subsequently be added to the repository of retained knowledge to create a form of database constituting problem descriptions, pertinent questions, proposed solutions, consequences and a convenient	Knowledge repository

Code	Code Groups
means of indexing and retrieving cases that suggest a solution to current problems.	
NNs are selected due to its suitability for both classification and prediction type of problems such as a customer behaviour which can be observed as a binary or a multiple response classification problem, as well as a regression problem of predicting a continuous output.	DM tasks
Nowadays, social media have become ubiquitous and are playing an increasingly critical role in today's business environments. Numerous companies are using social media networks such as Facebook and Twitter to provide various services and also interact with their customers. social media data have several unique characteristics. First, social media cover general users' opinions about almost every aspect of our lives. Second, there is always fresh content on social media, and the content is generated consistently and timely by numerous online users. Third, social media contents are associated with metadata in various attributes such as user, location, likes, time and dislikes. Fourth, social media data have quality issues and also contain a lot of noise and spam, which need to be sifted through to ffgure out trustable data.	Big data
One of the problems in which ARs can be applied is the market basket problem which assumes that there are a large number of products that can be purchased by a customer, either in a single transaction or over time in a sequence of transactions.	DM tasks
Our experience was that the structured blogs system was very useful for knowledge sharing to make DM meaningful for BI.	Data driven activities
Performance analysis is another crucial task for effective policing. DEA as a performance measurement technique has been applied in combination with clustering technique to rank PAUs on the basis of their effective enforcement of crime prevention measures.	Measurement
Physicians can choose most of the drugs in the result of association rule mining when they determine the treatment. In other words, the proposed system can generate more relevant prescription decision(s) for physicians to encounter different patients' complaints whereas the suggested medicine in the decision support is nearly the same as the decision of physicians.	DM tasks
Planning involves the evaluation of a variety of potential policies that impact the success of an organisation in achieving its goals and making effective decisions related to the resources, manpower, and processes needed to implement selected policies. Whether the objective is to develop a strategic, tactical, or operational plan, planning sets the direction an organisation should take to achieve its goals and guides it to follow that direction.	Decision Support
Practically, it is hard to ffnd an expert of DM who is also an excellent business insider, and vice versa. In other words, knowledge workers involved in DM and its applications are usually divided into two groups: business insiders and data miners. A business insider is a CEO or middle level manager who possesses best knowledge in business problem-solving and decision-making. She or he must understand the concepts of DM, BI, and KM in the organisation, although might not be familiar with detail DM techniques and procedures. A business insider's objective of taking part in conducting DM and the development of KM is to improve the business performance of her or his organisation. A data miner, on the other hand, is an expert of DM, and best understands DM	Knowledge acquisition for Expertise

Code	Code Groups
techniques in the organisation. She or he must understand the nature of the business and be able to interpret DM results in the business context, but is not directly responsible for business actions. The collaboration of these two groups of people makes DM relevant to genuine BI.	
Previously unknown information from the operational customer database.	Knowledge base
Product variation grows exponentially due to the increasing demand of customised products. Manually keying in the conffguration knowledge into a computer-based conffgurator is time consuming and error prone. Especially for complex product, the design constraints among components and attributes are complicated and large in amount.	Knowledge base
Recent advances in ICT have facilitated access to knowledge by providing new opportunities for growth and development	IT enablement
Regression: There is a certain function that can describe the relationship between the attributes of observations.	DM tasks
Representations can be in the form of theoretical advice to the decision-makers.	Decision Support
Essentially provide great value to BI generation. The integration of decision support and KMPs is crucial for enterprises to create their niche BI and to maintain global competitive advantages.	Competitive intelligence
Sequence analysis There is a certain pattern of time-dependent events An online shopper often search the correlated web sites.	DM tasks
Since knowledge plays such an important role in DM, one important implication for practice is that business value from BI can be enhanced by providing KM systems for DM.	Competitive intelligence
So with current technologies it is not possible to fully automise the KMP. There is still a need of human intelligence, but there are starting points for enhancement and improvement for more computer support.	Automation
Sometimes it is more important to explore the hidden knowledge in an organisation's data and transform it into an explicit knowledge to improve the decision-making process rather than relying on experience or rule of thumbs.	Decision Support
SSC algorithm is used to cluster crime locations and to find crime hot spots in India, as the density of crime incidents will be continuous over an area, higher in some parts and lower in others. For this application, NCRB experts have provided the information about crime zones of a few locations based on their domain knowledge.	DM tasks
Structuring data	Data driven activities
System integration	System Integration
Tacit to explicit knowledge	IT enablement

Code	Code Groups
Technologies such as expert systems, simulation, statistical tools, knowledge-based systems, and multiple AI technologies are integrated into the SKMS allowing the combination of different perspectives for acquiring the required explicit decisional experiential knowledge by the means of Decisional DNA.	System Integration
The rise of Big Data also provides opportunities to mine these resources with more powerful methods than could be justiffed before. This is because powerful learning algorithms are prone to overfft and require relatively large samples of data to avoid this problem.	Big data
The various internal data sources to be used are obtained from the firm's e-commerce systems, sales transaction systems, financial and accounting systems, human resource systems, and plant operation systems.	Data warehouse
The advantage of these techniques lies in the fact that they can be implemented on the existing software and hardware platforms, and can also be integrated with new products and systems. The size of technological infrastructure required by DM applications depends on the size of the database and query complexity. These advanced techniques are based on a full integration of a DM Server with the data warehouse and OLAP Server (Online analytic processing Server), which enables direct application of users' business models to the data warehouse.	System Integration DM tasks
The aim of Big Data and KM is almost same which is to use knowledge for better decision-making. However, the difference exists in the way it is performed. Big data are structured, unstructured or semi-structured data available through a variety of sources.	Big data
The assessment of organisational competencies is based mostly on objective indicators or documentable judgements related to the accumulation of critical resources, availability and use of management systems and use of formalized organisational knowledge applied in the implementation and control of critical processes.	Measurement
The benefits of integrating of BI with KM are 1) ensure a real support in deploying successful businesses across the organisation by smoothly managing multicultural teams of employees in providing highest quality products and global services to multicultural customers, 2) end-users preference and experience can be included in BI implementation, and 3) provide better understanding on business context, interpretation results and training to the end user.	BI and KM integration
The central theme of BI is to fully utilise massive data to help organisations gain competitive advantages.	BI and KM integration
The DM techniques like clustering for customer data segmentation, decision tree for customer preferences are applied on the customer data. This results in forming knowledge-based customer data warehouse. The knowledge extracted by the DM tools from customer data are applied for KM of customer data.	Data warehouse DM tasks
The effects of organisational form on the efffciency, scope and ffexibility of ffrm knowledge integration is inffuenced by the ffrm's combinative capability.	Measurement
The extracted information/knowledge obtained by applying BI tools must be stored for future use and sharing within the organisation. Knowledge repositories are widely recognised as key components of most KM systems	Knowledge repository

Code	Code Groups
used for storing such information/knowledge. They are collection of both internal and external knowledge and seek to capture both tacit and explicit knowledge.	BI and KM integration
The first trend is a greater reliance on semantic contexts and integrated approaches to knowledge extraction.	System Integration
The goal of a DM task is to verify hypotheses which have been kept in the data miner's mind.	DM tasks
The important components of IPS are Adaptive Query Interface (AQI), Knowledge Acquisition System (KAS), Database Management System (DBMS), Model Base Management System (MBMS) and visualisation. The objective of knowledge base is to enhance usability of the information and pattern retrieval process.	Knowledge base
The increased importance of BI infrastructure reflects three interacting trends, 1) More turbulent global environments, 2) Additional pressures to unveil valid risk and performance indicators to stakeholders, and aggravated challenges of effectively managing the densely interwoven processes.	Measurement BI and KM integration
The information and knowledge flows needed to support e-business require integrative systems that support e-marketplaces in providing the mechanisms for firms participating in the value chain.	System Integration
The insights gained by the decision support applications are used to refine the existing knowledge and feedback into knowledge organisation.	Decision Support
The integration of operational customer information with the analytical tools of DM drives analytic applications, to identify predictable trends in customer behaviours. The resulting information analysis can then be incorporated into organisations BI process for effective planning of marketing strategy.	Customer relationship management
The interaction process between the business insiders and data miners is actually a knowledge sharing process. In our view, the content of the entire interaction process (not just the DM results) is knowledge of the organisation. It includes: . linguistic standardisation of DM terms and concepts; . problem deffnitions; . DM documents; . DM resources; and . actions and outcomes.	Knowledge acquisition for Expertise
The Internet prompted the creation of a sixth type, the web-based DSS, which powers today's web-centric databases and knowledge portals.	IT enablement
The knowledge about customers, from customers and for customers help it gain maximum information pertaining to customers. This is a cycle as information derived out of operational CRM.	Customer relationship management
the knowledge about customers from these databases is critical for the marketing function.	Customer relationship management
The KM field is becoming amorphous as vendors are claiming that their products are ªknowledge managementº tools, whereas they simply may be document management or information management products.	Lack of understanding
The knowledge warehousing goal suggests three functional requirements for KW: (1) an ability to efficiently generate, store, retrieve and, in general, manage explicit knowledge in various forms, (2) an ability to store, execute	Knowledge base

Code	Code Groups
and manage the analysis tasks and their supporting technologies with minimal interaction and cognitive requirements from the decision-maker, and (3) an ability to update the KW via a feedback loop of validated analysis output.	
The main advantage of ANNs is their ability to operate with incomplete data. They are capable of profiling users to enable information to be targeted at specific individuals according to their preferences. KM can use this technology to advance knowledge distribution and sharing.	DM tasks
The medical knowledge in iEMRS assists physicians to make better prescription decision. With instant referencing to the comprehensive past medical histories of the patients and drug list recommendation (through discovering the pattern in the database), physicians agreed that the medicine selection process can greatly improved, especially in the area of drug-drug interaction checking.	Decision Support
The overall efficiency improvement of the e-healthcare system implementation by integrating the IoTs sensors and the FARM approach reduce the waiting time of the patients who are receiving care services.	System Integration
The purpose of including AI is to amplify the cognitive capabilities of the decision-maker in converting tacit knowledge into explicit knowledge, integrating this explicit knowledge by analysing it to detect new patterns and relations, and understanding the new knowledge by providing analogues and explanations.	Decision Support
The re-use of knowledge in repositories allows decision-making to be based on actual experience, large sample sizes and practical lessons learned, as well as real-time information obtained from sensors or from other devices.	Knowledge repository BI and KM integration
The selection of DM algorithms, hypotheses formation, model evaluation and refinement are key components of this discovery process. Because it takes cycles of trials and errors to progressively produce the most useful knowledge from the DM, a learning by experimentation approach 19 w x can be useful to ensure that the process can eventually AdiscoverB the useful knowledge.	DM tasks
The storage of information is the catalyst to converting explicit knowledge into new knowledge. This process is called integration or leveraging.	System Integration
The timely availability of problem -specific knowledge that is useful for the business problem under consideration is a useful service that extends the KMPes of organisations.	Enterprise wide knowledge
The valuable knowledge is more on the tacit side of the scale, so IT-based explicit KM systems are unlikely to be effective. Some tacit exchange systems might work, but ffrms should evaluate the nature of their own valuable intangibles (are they only present in “stars” incapable of easily sharing?) and whether person to person KM systems might be useful.	Enterprise wide knowledge
Then, explicit knowledge is transferred via documents, meetings, videos, e-mail messages, phone conversations, training manuals, rules, and policies. Finally, internalisation is the process of converting explicit knowledge back to tacit knowledge.	Knowledge base

Code	Code Groups
There are three major areas of application of DM for knowledge-based marketing — 1 customer profil- Ž . ing, 2 deviation analysis, and 3 trend analysis.	DM result to knowledge
There is a need for forward-thinking companies to expand their KM scope to include the massive amount of social media data. This expansion not only includes customer-generated opinions about their own businesses but also customer-generated opinions about their competitors.	Big data
Therefore, given unusual DM tasks or unfamiliar DM algorithms, it is important for the DM project teams to choose team members with diverse educational background and DM experience so that the team can make an optimal decision on choosing a DM method.	Knowledge acquisition for Expertise
These IT tools are perhaps developed for other purposes, rather than managing knowledge per se. It may, therefore, be difffcult to connect different IT applications or create interfaces and synergies among them in hospitality organisations.	IT enablement
These tools are continuously processing a variety of information and knowledge transactions at many different places during a product lifecycle evolution, which also needs to be accessed by other team members at many more places and applications.	Knowledge base
They also found that the Big Data analytics' adopters were ffve times faster in making good decisions than their competitors and twice as likely to be in the top quartile of ffnancial performance within their industries because of the BI obtained from the Big Data analytics.	Big data
To allow an effective and efficient processing of knowledge transactions during product realisation, C4 tools are being redesigned to reflect an organisation's collaborative and competitive posture.	Knowledge repository
To become a genuine BI tool for comprehensive knowledge discovery, DM must be integrated with KM for knowledge improvement in the organisation.	System Integration DM result to knowledge
To make the DM results actionable, the data miner must explain them to the business insider.	DM result to knowledge
To manage knowledge that crosses organisational boundaries and is distributed across supply chain partners. Customer knowledge is typically distributed across supply chain partners, and marketing is an important beneficiary of this knowledge.	Customer relationship management
To support all four data driven decision types, KM systems need to accommodate both data analytics and human insight.	Data driven activities Decision Support
Transparent exchange of knowledge among collaborating organisations requires a technical foundation for business process integration through standards-based software architectures.	System Integration
True CRM is possible only by integrating the knowledge discovery process with the management and use of the knowledge for marketing strategies.	Customer relationship

Code	Code Groups
	management System Integration
Unfortunately most DM techniques work best with very large samples.	DM tasks
Users can easily query and acquire knowledge in the knowledge base in order to share and exchange knowledge, and they can constantly enrich and update the knowledge base to achieve knowledge re-use and innovation.	Knowledge base
Using BI to identify internal strengths and weaknesses and external threats and opportunities only make the organisation more aware, but to achieve business benefit they need to find ways to exploit those opportunities and counter the threats.	Competitive intelligence
Using scalable distributed algorithms for harvesting knowledge.	DM tasks
We believe that one of the most important aspects of effective DM for BI is the knowledge sharing and planning phase that connects business insiders and data miners in the organisation.	Knowledge acquisition for Expertise
We suggest a model based on DM techniques, specifically ARs and neural networks (NNs). ARs and NNs are used in an integrative way to discover some interesting patterns of customer behaviour, specifically focused on items that are frequently bought together, and on the profiles of customers who buy those items.	DM tasks
While data and information have always been out there and have been helpful in steering the course of a business, new capabilities have expanded the range of what is possible.	Competitive intelligence
With better decision support in prescription process, the logic tier offers the treatment pattern algorithm to provide knowledge for physicians. Association Rule Mining is employed to discover the prescription pattern and present the patterns as rules to support the prescription.	DM tasks
With the IoT, products can be tracked anytime, making it possible to respond to customer behaviour. It is also now possible for products to connect with other products, leading to new analytics and new services for more effective forecasting, process optimisation and customer service experiences.	Customer relationship management Big data