University of Salford
MANCHESTER

Towards Intelligent Diabetes Knowledge Management and Knowledge Discovery: A Data Mining Approach

FAHAD DARWISH

Submitted in Partial Fulfilment of the Requirements of the Degree of Master of Philosophy

THE UNIVERSITY OF SALFORD
SCHOOL OF COMPUTING, SCIENCE AND ENGINEERING
MANCHESTER, UK

2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

ACKNOWLEDGMENTS

"God Thank you for giving me the strength and encouragement throughout all the challenging moments in completing this MPhil thesis."

First, and most of all, I would like to express my sincerest gratitude to my supervisor, Dr. Mohamad Saraee, for his patience, motivation and support throughout my study and related research.

I would also like to express my profound gratitude to the rest of the staff and team at the School of Computing, Science and Engineering at The University of Salford for their insightful comments and encouragement throughout this research.

I extend my thanks to the medical health care professionals and kind patients who were involved in the validation and who participate in the survey for this research. Without their kind participation and input, the project would not have been successfully conducted.

Finally, I would like to extend my sincere appreciation to my family: my parents and two brothers for their support and continuous encouragement throughout the process of researching and writing this thesis. This achievement would not have been possible without them. Thank you all.

PUBLICATIONS

- ❖ Darwish, Fahad., Silva, Charith., Saraee, Mo. (2019). Diabetes Self-Management System: Review of existing systems and potential enhancements. *2nd International Conference on Geoinformatics and Data Analysis ICGDA 2019.* Prague - Czech Republic, Paper accepted and was presented on 15th March 2019. Paper was published on International Conference Proceedings Series by ACM, indexed by Ei Compendex and Scopus and reviewed by Thomson Reuters Conference Proceedings Citation Index (ISI Web of Science). See http://www.icgda.org/# & Appendix C for more details

- ❖ Darwish, Fahad., Saraee, Mo. (2018). Diabetes Self-Management System: Review of existing systems and potential enhancements. *11th IEEE International Conference on Developments in eSystems Engineering DeSE 2018.* Cambridge - United Kingdom, Abstract accepted and was presented on 3rd September 2018. See http://dese.org.uk/wp-content/uploads/ProgrammeFull_DeSE2018_V7-1-1.pdf or Appendix B for more details.

- ❖ Darwish, Fahad., Saraee, Mo. (2018). Intelligent Diabetes Management System: For Better Management and Reduce cost to health providers. *In SPARC (Salford Postgraduate Annual Research Conference).* Media City - Manchester - United Kingdom, Accepted and was presented on 5th July 2018. See http://usir.salford.ac.uk/47964/ & Appendix A for more details.

LIST OF ABBREVIATIONS

Abbreviation	Extended Name
ANN	Artificial Neural network
CRISP-DM	Cross-Industry Standard Process for DM
DM	Data Mining
DT	Decision Tree
ICT	Information and Communication Technology
IDF	International Diabetes Federation
KD	Knowledge Discovery
KDD	Knowledge Discovery in Database
KM	Knowledge Management
KS	Knowledge Sharing
LR	Logistic Regression
NB	Naïve Bayes
NHS	National Health Service
NLP	Natural Language processing
NSA	National Security Agency
RF	Random Forest
SA	Sentiment Analysis
SHA	Secure Hash Algorithm
SVM	Support Vector Machine
SECI	Socialisation, Extenuation, Combination and Internalisation
TF-IDF	Term Frequency-Inverse Document Frequency
UMLS	Unified Medical Language System
WHO	World Health Organisation

ABSTRACT

Self-monitoring and self-management play an increasingly vital role in the management of prevalent diseases afflicting millions of people worldwide. Conditions such as cardiovascular disease or diabetes mellitus can be managed with pharmaceuticals. However, lifestyle factors and behaviour modifications also play a crucial part in controlling outcomes. If carried out effectively, self-monitoring and management techniques can benefit the patient from a health point of view, empowering them to take control of their disease, and also support the health sector economically. A recent report showed that almost four fifths of the NHS diabetes budget was spent on managing preventable complications.

Although applicable to any medical conditions with lifestyle control elements, this thesis will be using diabetes mellitus as a recurring theme to highlight research conducted. Diabetes is a rapidly growing epidemic disease with global implications impacting humans socially and economically. By 2025 it is estimated that five million people in the UK will be diagnosed with diabetes., This research will therefore, be relevant for a large proportion of the population.

Knowledge Management (KM) has demonstrated to be a valuable approach to sharing knowledge and providing users with the information necessary to help self-manage their symptoms. Although, KM has not yet been applied sufficiently to support the growing number of diabetics in the UK. In this thesis, KM is merged with Knowledge Discovery (KD) to combat that and address the specific needs of the diabetic population. The integrated framework is implemented using data mining techniques within the proposed e-Toolkit to elicit useful knowledge encountered by patients regardless of their disease, such as adverse drug reactions. The knowledge is then disseminated through the proposed modified SECI Model for knowledge creation via the e-Toolkit.

The second part of this research investigates which patient data is necessary to disseminate to healthcare professionals. This help to bridge any communications gap that may exist between patients and health care professionals. In theory, the e-toolkit will provide patients with one place to record every important health factor, whilst simultaneously allowing the medical team with real-time monitoring. Thus, enabling doctors and patients to work together to find effective ways to reduce the damaging effects of this disease, including determining the common side effects through medication reviews.

Keywords: Diabetes, Knowledge Management, Knowledge Discovery, British National Health Service Web System, Doctor, Patient, Data Mining.

CHAPTER ONE: INTRODUCTION

Diabetes is a devastating disease which can affect people of all genders, ages and races, regardless of economic status. If not treated the effects can cause long-term harm to vital organs and can even prove fatal. Medical specialists concur that sufferers of the disease require close supervision and observation, and intense self-organisation to effectively manage symptoms (American Diabetes Association, 2018).

Self-management plays a huge role in the control of diabetes amongst patients. This is usually carried out by taking regular measurements of their blood glucose levels, as well as other testing. Blood glucose readings are taken upon awakening, pre mealtime and after a meal, pre an exercise or after an exercise, before going to sleep, and at other occasions, if and when needed (S C.M. Croxson, 2009; Fahad Darwish, 2019).

Part of this self-management is awareness of diet, particularly carbohydrate intake, including type and amount. It is also important for diabetics to understand how factors like this have an impact onto their personal condition, and that these factors can differ between individuals (Richard S. Surwit, 2013).

Diabetics themselves have a vital role in managing certain risk factors which can affect their condition. This includes day to day factors such as exercise, diet and medication control. Self-management and education in the health sector can be supported by Knowledge Management (KM). KM is bringing the right knowledge to the right individuals so they can make the best decisions at the right time (Petrash, 1996).

KM can facilitate knowledge sharing amongst different organisations and individuals through communications (Almuayqil, 2017). However, in the UK there are a number of limitations with KM in relation to the practices in the NHS, If KM is used correctly then it could potentially save the healthcare service a large amount of money (Independent, 2018; Currie, 2008).

The risk of developing serious, life changing complications can be prevented by in-depth observation and regulation of health factors affecting diabetics such as blood glucose levels, diet, exercise, insulin dosage, medicines interactions, blood pressure readings and healing time of the individual (Darwish, 2012; Diabetes UK, 2018).

Such factors can have a severe effect on the patient's quality of life and in extreme cases, may even lead to death. Examples of risks associated with these factors include reduced bladder control, nerve damage, loss of vision, kidney failure, coronary artery disease and parts or whole foot amputations (Blumenthal, 2018; Darwish, 2012). Therefore, it is vital that the doctor is up-to-date with the patient's readings to ensure correct diagnosis and to enable appropriate treatment plans to be advised.

Knowledge Discovery (KD) through Data Mining (DM) techniques will identify knowledge which if applied, could improve treatment procedures. KM technology can be used to spread knowledge between patients and medical professionals. The ultimate aim of this thesis is to develop an e-Toolkit to help patients and doctors communicate directly, thereby educating each other on the individual's case, allowing assessment and review to be carried out with ease. (Rhoda K. Moise, 2017).

Nonaka and Takeuchi (1995) has proposed a knowledge creation model which facilitates the conversation of explicit and tacit knowledge using the four modes of socialisation, externalisation, combination and internalisation as termed (SECI). The SECI model is an established method in the knowledge management domain and has been successful in managing knowledge in a range of domains, such as healthcare (Almuayqil, 2017).

The integration of KM and KD can be achieved by the use of SECI model to help the healthcare professionals and patients in the United Kingdom in managing several long-term diseases, like diabetes mellitus and hypothyroidism. Such an approach could also help distinguish factors such as the side effects of medications, as is discussed in chapter 4 below. The discovered knowledge will then be disseminated to patients through the proposed SECI model on the e-Toolkit. Such an integration can be a powerful tool for making effective decisions and will help in self-management (Li, Zhu, & Pan, 2017).

1.1   Research Motivation and Problem Statement

The number of people that have been diagnosed with diabetes is set to double within the space of 30 years, growing from 135 million in 1995 to 300 million in the year 2025 (Abolghasemi, 2015). The International Diabetes Federation (IDF) predicts that this number will increase even further, to over 552 million by the year 2030. This number equates to around 6.4% of the world's population (The Lancet , 2011; United Nation, 2019).

The global medical community are struggling to keep up with the demand for diabetic care as the numbers continue to rise and potentially spiral out of control. According to the International Diabetes Federation (IDF), on average, around 3.2 – 4.0 million people, between the ages of 20 and 79 years, are predicted to pass away due to diabetes in 2019 (Muddasarul Hoda, 2019; idf.org, 2017). So far, diabetes have been a cause of mortality for around 10.7% of the global population from this age group (diabetes.org, 2010). This percentage is much higher than the combined deaths from the HIV/AIDS infectious diseases, which is 1.1 million deaths (Nam Han Cho, 2017).

Astonishing reports by the journal "Diabetic Medicine" indicate that the treatment cost of diabetes in the UK is set to rocket. The report forecasts that "…Diabetes will cost the NHS £16.8 Billion (17% of the NHS budget) by the year 2035 .." (NHS, 2012).

Baroness Young, who was the Chief Executive of the health charity Diabetes UK, mentioned that "… already huge sums of money spent on treating diabetes will rise to unsustainable levels that threaten to bankrupt the NHS…". As it stands almost 80% of the NHS diabetes budget is used to treat complications to the disease which actually may have been prevented with proper self-care and education (BBC, 2017). This shows the importance of quickly finding newer, alternative methods of managing the disease.

As already highlighted, self-monitoring and self-regulation is of extreme importance to successful management of the disease. A report by Yi Weinger illustrates that a large proportion of diabetics do not or cannot make such commitments. This can be due to the lack of self-efficacy, the absence of care education or the absence of diabetes management resources and organisation (Weinger, 2003; Darwish, 2012).

The patients that do attempt self-monitoring may struggle to constantly track the various factors that need to be recorded, and so they are usually not consistent with their choice of system, and are repeatedly switching from one system to another (Weinger, 2003). A quick search using Google shows many websites that are devoted to helping those who are having trouble with the maintenance of their diabetes. These include the National Institute of Diabetes and Digestive and Kidney Diseases, Diabetes Education Scotland, the Palo Alto Medical Foundation and the National Health Service.

Due to a knowledge deficiency on the disease process and lifestyle factors amongst patients, healthcare professionals can find it challenging to influence patients to modify their risk factors. This can prolong a patients' exposure to behaviours which can increase their risk of further problems. This is preventable with appropriate education and training yet is "becoming increasingly unaffordable for health care systems around the world" (BBC, 2017).

Learning about diabetes is vital towards the treatment plan of diabetics. However, there still remains a barrier between diabetics and the professional medical community. A renowned diabetic expert, Doctor Anderson, stated that Noncompliance happens when patients and the health medical team are chasing dissimilar goals (Resnicow, 2007).

Ineffective communication between the patient and their doctor only worsens this issue and negatively impacts the patient-doctor relationship. Implementing a standardised digital system to manage diabetes could improve communication between the doctor and the patient as the same data is shared between both parties via a direct platform.

Research was directed towards utilising computers and smart phones that can be used by diabetics for creating or further developing more suitable ways to record health risk factors. However, due to the applications not having the patient's or medical professional's best interests in mind due to their commercial purpose, they usually are not as effective. Our findings, from the Survey Analysis in Chapter 3, show that most websites or portals are created by the manufacturers of blood glucose and insulin pumps therefore with only buying customers benefitting.

Another drawback of those websites is that they are focused on product features and only record blood glucose readings and insulin dosage and do not take into account other factors which are necessary for the correct management of the disease (Larme & Pugh, 2001; Paul Kurdyak, 2017).

The use of such websites is restricted and can only be used by patients who have used or purchased the device, making accessibility limited. Patients who can use these websites would still have to employ another method to record their own figures (Peggy, 2007; Chin et al., 2001). However, as mentioned later in chapter 4 the proposed e-Toolkit can share figures directly with medical professionals without the need for face to face meetings.

The data collected from the survey and the literature in chapter 3 of our research concludes that existing electronic systems neither offer a way for NHS professionals to have direct and complete access to the patient readings and updates, or a base for the NHS professionals to communicate with diabetic patients.

The professional's ability to influence or to intervene is mainly reliant on upon the routine readings of patients, which is challenging as the data must either be emailed or printed and taken to scheduled appointments. (diabetes.org.uk, 2019)

One of the first issues that arise from using the current systems is that the only factors considered are insulin and blood glucose readings. Medical professionals are not able to access a patient's complete day-to-day health status and thus approach their treatment holistically (Peggy, 2007; Chin et al., 2001).

Another issue with current systems is that the equipment used by patients to record data does not allow information to be emailed to the clinic, preventing communication between the patient and the healthcare professionals.

Furthermore, there are no KM techniques or tools and the current system does not allow the patient to voice their opinion/ experiences of specific medications. Finally, an overwhelming number of patients and records can make it difficult for healthcare professionals to keep up to date with information.

1.2   Aim and Objectives

This research attempts to answer the following questions:

Q1. Can a universally accessible e-Toolkit improve diabetes management for both NHS staff and diabetic patients?

Q2. Can the integrated framework of KM and KD address the issues that face the medical and diabetic community?

Q3. Can the proposed improved SECI model framework address Knowledge Management and potentially help reduce the NHS budget?

The objectives of the research are to develop a framework to integrate KM and KD via the developed SECI Model to support the e-healthcare of diabetes mellitus patients and assist the NHS professionals in the UK.

The following objectives are to be carried out:

1. To conduct a literature review for the following:

    a. Knowledge Management and its models.

    b. Knowledge Discovery applications in the healthcare sector.

    c. SECI Models and their applications

    d. Diabetes mellitus and its complications and prevalence globally and in the United Kingdom.

2. To carry out a survey in order to establish limitation of the existing diabetic monitoring systems.

3. To develop a framework to integrate Knowledge Management and Knowledge Discovery within a developed SECI model

4. To carry out a Data Mining analysis of medication reviews to find the most common issues with certain medicines in order validate our framework idea, as well as applying non-medical datasets such as Amazon reviews. This is to prove that the framework can be applied in different fields and not only medical.

5. To build an e-Toolkit to support diabetes education and self-management in the UK.

6. To validate the developed framework and evaluation of the designed e-tool kit.

The scope of the research is intended to support and help diabetic patients in the United Kingdom and, potentially, around the globe.

The e-Toolkit Aims to achieve the:

- Creation of a universally accessible e-Toolkit that will help the diabetic patients in the UK and globally. The e-Toolkit will have many purposes offering an area which is simple to use such that diabetic patients are able to document risk factor data. Diabetics patients will have access to this e-Toolkit from anywhere in the world using their smart phones, tablets, PC's or any other device which has internet access.
- Sentiment Analysis techniques and other Data Mining tools will be used within the e-Toolkit to potentially find side effects of medications and give suggestions of treatment options for healthcare professionals according to the data which patients have input.
- The ability to communicate with the patient's appointed medical professional via a direct messaging system, with measures in place for regular monitoring. Therefore, the e-Toolkit can be incorporated into patient's day-to-day routine in real time.
- The ability to display data graphs and charts and high risk patient alerts to medical professionals, and the ability to share and store parts of patient medical data (e.g. test results)

The e-Toolkit will seek to overcome the limitations of current available systems which were discussed in the survey and literature review. This e-Toolkit will address a current problem with which the NHS is struggling and will aim to reduce the portion of the NHS budget which is currently spent on diabetes.

1.3   Research Methodologies

Research strategy can be explained as "the general plan of how the researcher will go about answering the research questions" (Saunders, Thornhill and Lewis, 2015). This thesis will employ the research onion methodology presented by (Saunders et., 2009). It shows the steps that must be followed when creating a research strategy.

The first onion peel is the research philosophy, the primary stage for the appropriate research method, which is continued in the second stage. Research strategy is then chosen in stage three, followed by the time horizon in stage four. The fifth stage is where the data collection methodology is chosen.

Onion research has many advantages such as the creation of a series of stages which allow for the understanding and application of different methods of data collection (Saunders et., 2009). an illustration of the onion method is shown on figure 1.
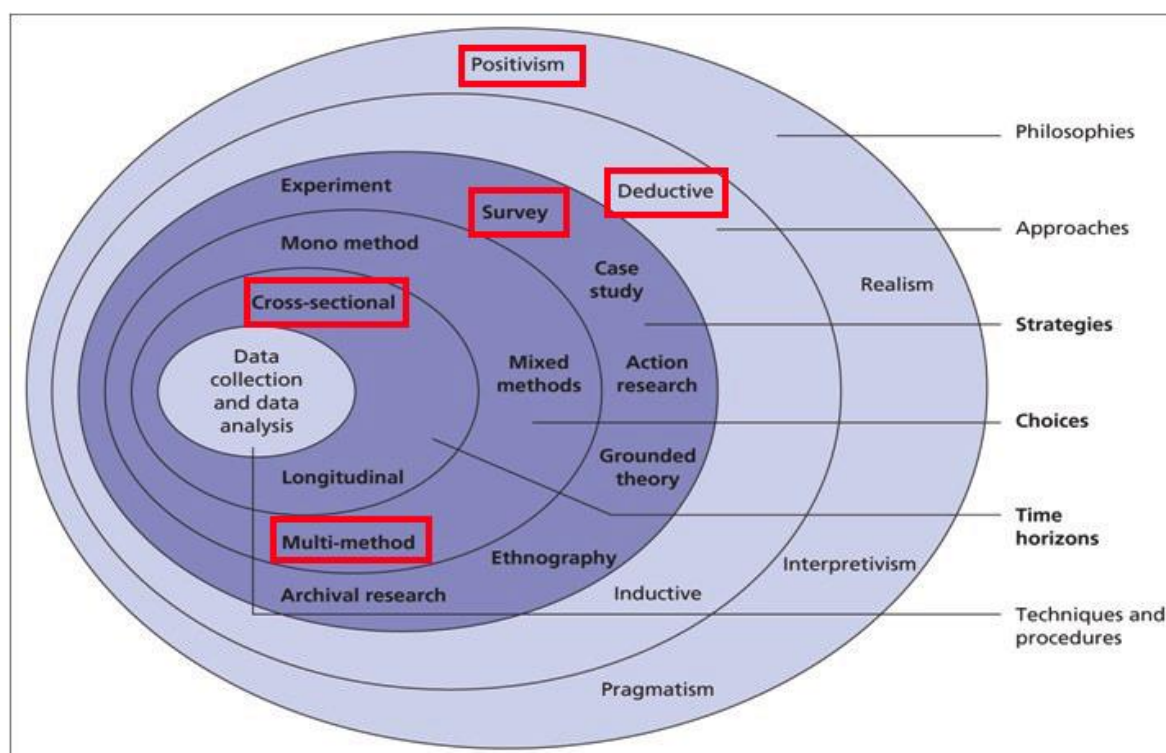
Figure 1 - Research Onion with our selection of research methodologies for the current thesis Source: (Saunders et., 2009)

The onion layers allow the researcher to examine the above methodologies and easily identify how and where to fit particular features, as will be described in the following sub-sections.

1.3.1 Research philosophy

Research philosophy can be sectioned into four subcategories: realism, pragmatism, positivism and interpretivism. The selection of a category is based on the queries of the study. Research philosophy brings about assumptions linked with how a person views their surroundings (Saunders et., 2009). The researcher is impartial in data collection and views a society with a philosophy of positivism (Snieder R. & Larner, K., 2009).

The researchers do not have an influence on the research being carried out and have no desire to change any components of the data. Being unbiased during a study is vital and this point is stated in the positivist philosophy. A logical outlook to the growth of knowledge is specified in positivism. This philosophy presumes that objects are present regardless of our awareness of their presence and can be split into direct and critical realism (Saunders, Thornhill and Lewis, 2015).

On the other hand, interpretivism aims to generate a greater analysis of the social world according to (Saunders, Thornhill and Lewis, 2015). However, pragmatism states that ideas are only applicable when they back action. When the study being carried out does not indicate positivism or interpretivism, then pragmatism can be used. In the case of positivism, the researchers are neutral, and they have no effect on the study. Positivism postulate that reality exists regardless of the study matter. In theory this means phenomena is consistent between subjects (Isadore Newman, 1998).

Therefore, positivism is most fitting philosophy to use to analyse an observable reality, for instance in this thesis. The researcher uses the survey data from the participants to determine aims and objectives of the research study and discover new ideas. They can also eradicate existing ideas on which the diabetic community have already given their opinions and views via the proposed e-Toolkit. Therefore, a positivistic approach is fitting for this research study as the researcher aims to use unbiased approaches to support the diabetic society.

1.3.2   Research Approach

Approaches to research can be separated into two groups: inductive and deductive (Saunders et., 2009). The deductive method examines a current theory to classify ideas that will be examined using information and data collected by the researcher whilst the inductive process collects data and constructs a theory based on the data analysis results (Kothari, 2011). Selection of an approach is built on factors such as the research problem nature, the research philosophy, and the research study area.

This thesis will focus on the deductive approach as it emphasises recognising diabetes mellitus complications as well as e-health obstructions in the UK. The researcher has applied this approach by creating an e-Toolkit which will apply to the suggested framework alongside numerous literature reviews.

1.3.3   Research strategy

The third layer of the research onion methodology is strategy. The researcher has chosen the 'survey' strategy to determine that the concept of an e-Toolkit will be accepted by the medical and patient community. The survey strategy also allows for the concerns about the portal to be assessed as well as establishing key features that are required for the creation and longevity of the e-Toolkit. Survey research is explained as the collection of data from a group of people through their answers to questions (Check J., Schutt R. K., 2012).

This research strategy has a number of aspects including data collection instruments, recruiting participants, ethics and more. Survey research can use either qualitative research approaches, involving open-ended questions, or quantitative research approaches, As explained by Bryman book "Surveys tend to be used in quantitative research projects and involve sampling a representative proportion of the population" using mathematical rated items, questionnaires or a combination of both approaches can be used (Bryman, 2012).

1.3.4   Choice of methods

The selections indicated in the onion research touches on methods of mono, mixed and multi. A large variety of techniques are used in the multi method (Bryman, 2012). However, only one research practise is used in the mono method and two or more practices are used in mixed method, which both mention the practise of qualitative and quantitative approaches. The only thing that differs is that the mixed method uses a shared strategy that generates a solo dataset (Flick, U. , 2011). Whereas the multi-method strategy focuses on divisions of separate segments in the research with specific datasets individually created, and then each is examined using the techniques via qualitative and quantitative (Feilzer, M. Y., 2010).

Quantitative, qualitative, and mixed approach are the three main research methodologies (Williams, C., 2007). Researchers who need to produce further knowledge and quantify numeric data tend to use quantitative methods (Williams, C., 2007) . Finding out knowledge from real situations is known as a qualitive approach (Creswell, J. W. , 2013). This type is not as organised and includes explaining and describing gathered data (Williams, C., 2007; Leedy, 2001).

Imperative trends can be studied and taken from data which can be assembled from questionnaires and used as quantitative research methodology, this feature is based on the positivist research philosophy (Williams, C., 2007). When questionnaires are used in survey studies it tends to be linked with the deductive method. We will use two questionnaires to group applicable data quantitatively, thus the research method decided for the study is a quantitative multi-method approach.

1.3.5   Time horizons

This layer is the time scale considered for the completion of the project (Saunders, Thornhill and Lewis, 2015). As stated by Bryman – "within the research onion, there are two kinds of time horizons; the cross sectional and the longitudinal" (Bryman, 2012). The cross-sectional time horizon is already recognised, where the data collection is essential and at a particular timeframe (Flick, U. , 2011). In the longitudinal time horizon, data collection is repeated over a lengthy period and is key in research for investigating change over time (Wayne Goddard, 2004).

The researcher of this thesis, over the course of four months, conducted the primary research in the UK. Within this time frame, the researcher distributed the surveys and consent forms to volunteers in the medical field particularly the survey for medical professions and the survey for the diabetic patients. When the data was collected it was not needed for any further collection of further data. Therefore, this project can be characterised as having a cross-sectional time horizon.

CHAPTER TWO: BACKGROUND AND LITERATURE REVIEW

2.1   Introduction

This chapter will carry out a literature review in order to investigate the previous related studies and examine the approaches that were adopted in these studies. For simplicity this will consist of three main parts. The first part covers a background on diabetes. This is followed by a review of existing diabetic monitoring systems, and finally literature on Knowledge Discovery, Knowledge Management and SECI Models and Related Research.

2.2   Diabetes Mellitus

Diabetes is a metabolic, chronic disease characterised by high levels of blood glucose (or blood sugar), WHO has stated that diabetes could "Leads, over time, to serious damage to the heart, blood vessels, eyes, kidneys, and nerves" (World Health Organization, 2018). It typically occurs when the body cannot store glucose (a kind of sugar) correctly due to the production of low levels of a hormone called insulin which is created in the pancreas that help the cells to take in glucose from the carbohydrates in the foods consumed.

Eventually the cells turn the glucose into energy. Glucose is a key supplement for cellular energy. However significant detrimental health effects can occur if insufficient amounts of insulin are produced. This can include impairment of brain and liver cell function due to inadequate glucose uptake into the cells (Diabetes UK, 2018).

Moreover, if there is excess glucose in the blood stream (not in cells), it can build up in some of the organs, for example the eyes, leading to diabetic retinopathy (high blood glucose levels can destroy and damage the eye retina) which can lead to blindness. Any excess glucose in the blood is filtered and removed by the kidneys and bladder. This surplus of glucose creates a high blood pressure within the body, so the kidneys become damaged by the increasing pressures within the filtration system. This pathology is known as diabetic nephropathy and can lead to harmful toxin build-up within the body (nhs.uk, 2017; Diabetes UK, 2018; William D. McArdle F. I., 2010).

2.2.1   Types of Diabetes

In order to appropriately research and implement the aim of this thesis, it is vital to have a sufficient understanding of the disease. This section discusses in detail the most common types of diabetes.

2.2.1.1   TYPE-1

Type 1 diabetes is an auto-immune condition which tends to present itself in patients under the age of 20, but can affect patients at any age. The auto-immune condition means the immune system attacks healthy body tissue, specifically the cells in the pancreas (Zheng, 2017; Unwin, 2011). This means that the damaged pancreas is then unable to produce insulin, meaning glucose cannot be moved out of the bloodstream and into the cells. It is considered to be a genetic disorder meaning you are more likely to be affected if a close relative also suffers from the condition (nhs.uk, 2018). It is reported that you are 6% more likely to suffer from diabetes if you have a close relative who also has the disease compared to 0.5% if you do not (nhs.uk, 2018; Blumenthal, 2018; Cihakova, 2001).

Juvenile diabetes is the term given to diabetes effecting younger age group, particularly teens and children (diabetes.co.uk, 2018).

2.2.1.2   TYPE-2

Type-2 diabetes was previously known as non-insulin-dependent or adult-onset diabetes due to mainly affecting people over the age of 40. However, as stated on Diabetes UK "Type 2 diabetes is becoming increasingly common in young adults, teens and children and now accounts for approximately 90% of all diabetes cases around the world". More often it can occur in the body for years before a diagnosis is given (Diabetes UK, 2018).

Type-2 diabetes is different from Type-1 as the condition can be reversed if the correct measures are implemented early enough, such as changes to diet and exercise (Hollingsworth, 2013). The IDF (International Diabetes Federation) has reported that "…Up to 80% of type 2 diabetes could be prevented through healthy diet and regular physical activity…" (IDF, 2019). This shows that Type-2 can be managed without the use of insulin or medications with good diet and the right monitoring management systems.

Type-2 diabetes is often defined by the lack of the ability to produce enough insulin to meet the body's requirements or defective insulin which is rejected by the cells in the body. This second effect is called 'insulin resistance' (National Institutes of Health, 2019).

A chain of events transpires when the cells are incapable of absorbing glucose. Firstly, not used glucose can builds up over time in the bloodstream. The medical definition for this is hyperglycaemia.

The increased levels of glucose prompt the pancreas to boost insulin production in an attempt to utilise that glucose, however the insulin is unusable and the pancreas exhausts itself (Diabetes UK, 2018; Merck Manual, 2007; The New Encyclopaedia Brittanica, 1998)

Secondly, the cells send signals to the body that they are starving for glucose, which puts pressure on the pancreas due to the liver releasing the stored glucose and leads to the hyperglycaemia getting worse. In the meantime, the extra glucose is filtered by the overworked kidney and bladder and the excess glucose damages various organs around the human body (A.D.A.M. 3, 2018).

2.2.1.3   Gestational

Gestational diabetes happens when the developing foetus is not provided with the right amount of insulin especially as during pregnancy the body needs more than usual. Sometimes the activity of insulin is being blocked by an increase of pregnancy hormones (NHS.UK, 2018) .

Usually after the birth of the baby the condition disappears, however it is likely to recur in future pregnancies of woman who have already experienced it once (NHS.UK, 2018).

A recent study found that women who have gestational diabetes mellitus (GDM) had a 10-fold bigger chance of developing Type 2 diabetes compared to women with no GDM (Herath H, 2017).

2.2.2   Consequences of Uncontrolled / Mismanaged Diabetes

Inefficient control of diabetes can lead to complications that affect the entire body, as discussed by Bramble, "Complications could be ranging from heart disease and a higher risk of heart attack or stroke, visual problems including blindness, kidney disease that can lead to kidney failure, neuropathy (nerve damage)", an increased risk of infection and problems effecting the oral dentition. These complications if not managed can ultimately lead to death (Bramble., 2003).

Fortunately, managing blood-glucose levels aid in the prevention of the aforementioned problems or if they have already been developed, it will help in at least managing those problems/complications (Margaret O'Malley, Kelly Kennedy, RD, 2018).

Although there are medicines available to help treat and control the above complications. However, those medicines can be avoided with the right education and self-management i.e. maintaining exercise, diet, blood glucose readings etc from the beginning. Prevention is better than cure, and even though there are medicines to help treatment of complications, if these complications are avoidable in the first instance then patients must take some responsibility where they are in a position to help themselves (Mayoclinic.org, 2018). This essentially details how diabetes is a disease in which prevention through lifestyle means is better than treatment with medication once the disease has progressed.

2.3   Responsibilities of the Healthcare Team

In the case of a diagnosis of diabetes, in addition to the GP, an entire team of medical professionals can become involved. This may include endocrinologists, dieticians, dentists, chiropodists, diabetic nurses and podiatrists who look after the patient's feet to prevent problems such as gangrene and ulceration in later years (diabetes.org, 2018).

The entire multidisciplinary team will try to educate the patient about their condition, providing the help they need. However, the healthcare professionals can only do so much when diabetic do not take the condition seriously and neglect to self-monitor their selves.

As stated on Diabetes UK - Doctors will usually arrange for patients to have various tests throughout the year such as the following:

- "C-peptide test – Measures the amount of insulin the body produces
- Fasting plasma glucose test – Identifies the impaired fasting glucose (resistance to insulin levels).
- GAD antibodies - Helps identify type 1 diabetes.
- HbA1c test – Calculates the average blood glucose levels over the last 2-4 months
- Oral glucose tolerance test (GGT) - Helps identify diabetes or insulin resistance.
- Type 2 diabetes test – Detects risks of developing type 2 diabetes" (diabetes.org.uk, 2018).

Doctors will also monitor blood pressure levels, examine kidney function, test insulin levels and adjust dosages if necessary, prescribe medications and monitor life style activities; especially physical activities such as swimming, running and walking.

The medical team can also monitor sleeping patterns and perform weight checks. The dentist should be seen every three months to check for signs of gum disease or inflammation and a routine visual examination should be done every six months (Dr Sandra Cabot MD, Margaret Jasinska ND, 2007).

2.4   Responsibilities of the Diabetic Patient

Patients who have recently been diagnosed with diabetes will learn to first keep up with their condition. They will have a daily routine and will need to keep thorough information of how numerous aspects affect their blood sugar levels such as energy, mood, blood pressure, bladder control etc. Patients should also learn how to measure and record their blood glucose and insulin levels and administer their own medications. In addition, they will learn of adjustments to diet which types of food they should east and which they should avoid and the timing of meals. They will also need to be educated on the major side effects of the medications they take for diabetes. It is also important for them to be able to detect and treat hypoglycaemia, when to call the healthcare team to report high glucose levels or with any questions or concerns. (A.D.A.M. 3, 2018; Elizabeth Blair, 2013; Fahad Darwish, 2019)

2.5   Knowledge Management and Knowledge Discovery and how they are used to manage diabetes and other diseases

Knowledge Discovery (KD) and Knowledge Management (KM) are two fields with the ability to empower people. KM is a dynamic process of capturing, sharing, storing and generating new knowledge (Ni, 2010). The main aim of KM is to increase knowledge within an organisation through communication, offering opportunities to learn and disseminate knowledge (McInerney and Koenig, 2011). The utilisation of KM in healthcare is associated with a number of advantages. For example, it has been stated that medical staff are benefitted when KM is implemented by reducing the amount of medical prescriptions required (Morr, C. El. and Subercaze, J., 2010).

In addition, according to Computerworld (2002), and other numerous studies it has be stated there has been an error decrease of 55% in the healthcare system due to KM. This type of management can be proven to be very successful in the medical field with an enhancement

of quality of care and improvements in communication which will reduce medical mistakes. KM can contribute to this objective by the development and sharing of medical knowledge within the hospital (Morr, C. El. and Subercaze, J., 2010). In fact, the reduction of medical errors and the increased cooperation facilitated by KM in healthcare can result in improved cost efficiency. The practice of cooperation aids to the sharing of information among different healthcare providers, whereas KM can reduce adverse drug effects, therefore reducing medical expenses caused by medical errors.

Applying KM in the healthcare sector has many benefits, some of which include creation, identification, acquisition and dissemination of healthcare knowledge. These are extremely fragmented, proliferative and crucial for decision-making (Nicolini et al., 2008). The practice of cooperation can encourage the sharing of information among different healthcare providers, whereas KM can reduce adverse drug effects and reduce medical expenses caused by medical errors.

Knowledge Discovery in Databases (KDD) is a field concerned with the progression of methodology and techniques for making sense of data. Generally, a KDD contains multiple stages, beginning with a database and involving selection, pre-proccession, sub-sampling and transforming of a selected dataset (Fayyad, 1996). The following stages after include applying a data-mining (DM) algorithm to extract a knowledge pattern and then evaluating the associated knowledge and identification. These steps are illustrated in Figure 2:
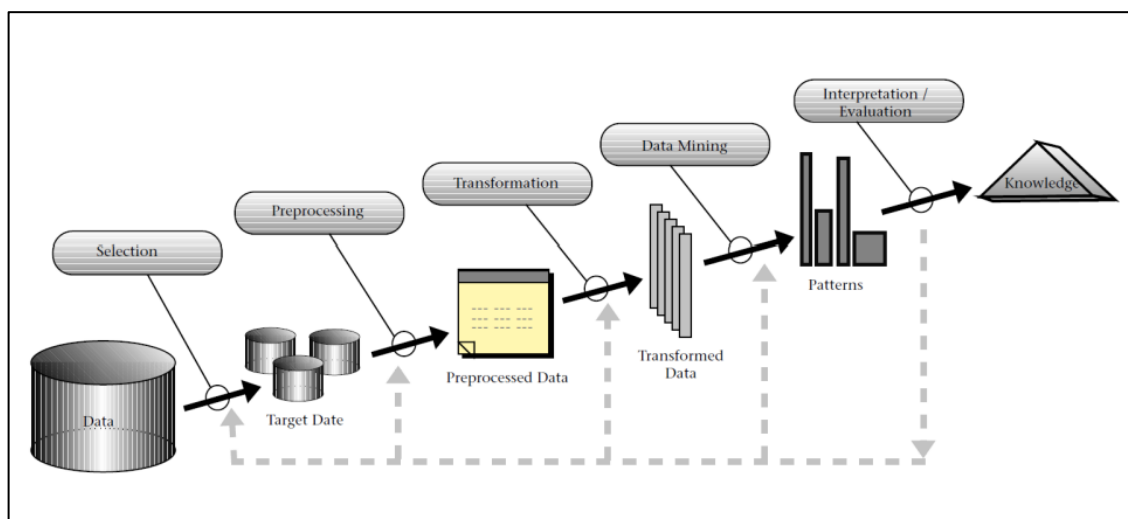
Figure 2 – KDD General Steps - Source: (Fayyad, 1996)

Knowledge Discovery (KD) can be implemented through data mining to produce useful knowledge. KD and DM have been very popular and have been applied successfully in various domains including education, banking and customer relations management (Jiawei Han, 2012). Foremost, DM has been successfully implemented and shown itself as a sufficient tool in assisting disease management and advancing the quality of services in healthcare.

Studies have stated that the process of managing certain diseases including diagnosis and treatment has been improved by DM (Yin Kia Chiam, 2019). These diseases vary from cancer and heart conditions to diabetes mellitus (Ioannis Kavakiotis O. T., 2017). A study carried out by (Resul Das, 2009) used a neural network classification algorithm from the DM software and performed different diagnosing experiments of heart disease, obtaining an accuracy rate of 89.01%. (See Figure 3). The importance of medicine and its impact on human lives made the implementation of DM in such a domain essential.
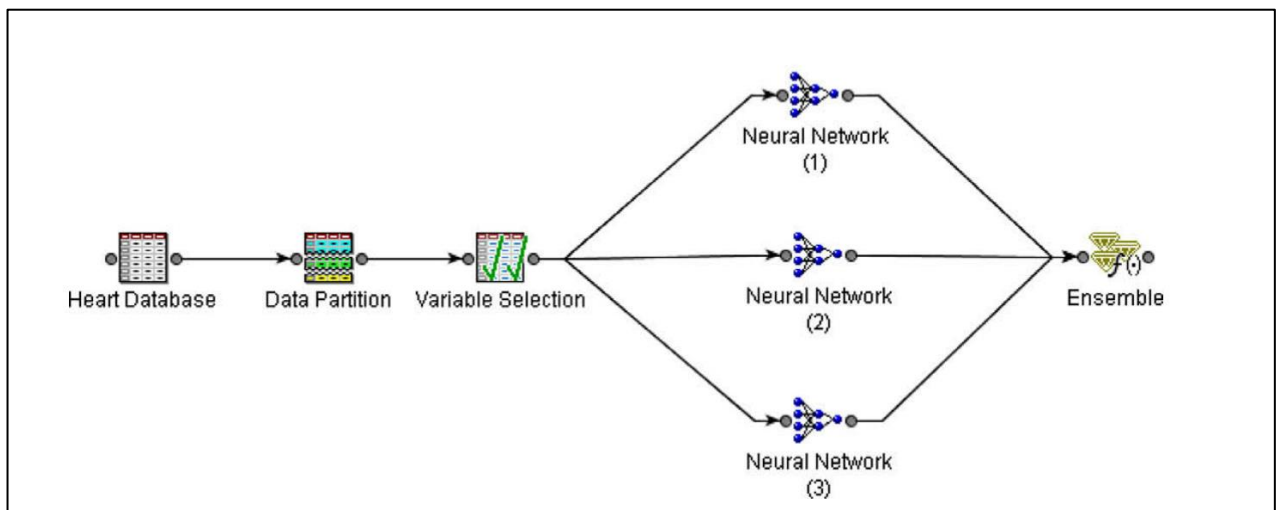
Figure 3 - Heart disease diagnosing system - Source: (Resul Das, 2009)

Mining medical data aims to promote healthcare services. We can define medical DM as the process of extracting hidden, potentially useful and novel information from medical data for the purpose of improving accuracy, time and cost efficiency and decision support systems (Niaksu, 2015). Mining data in the medical domain has multiple goals. The primary goal of this process is to extract the desired knowledge, which can be used to identify valuable medical information, such as risk factors of a disease.

The literature has evidence of achieving this goal successfully in different diseases (A. Sallaberry, 2011). A further benefit of DM in the domain of medicine is the efficiency of time and cost, which can lead to the prevention of unsuccessful treatment of diseases. For some specific illnesses, quick thinking and abrupt treatment are required to prevent any harmful effects.

A study from (Rusdah, 2013; Tamer Uçar, 2011) showed that the diagnosis of tuberculosis is time consuming. However, using an interface system the presence of the disease can be predicted in less than the original diagnosis time of 45 days. This interface system used DM which facilitated temporary treatment and also showed accurate and consistent results.

The establishment of the Medical Decision Support Systems (MDSS) is enabled by medical domain of DM, providing a use in the automation of several medical care processes. A study by (Nura Esfandiari, 2014) used MDSS to use six different prediction algorithms and compared it with six DM methods to estimate the mortality rate arising from head injuries. One of the DM methods, Boosted Tree Classifiers and Regression (BTCR), demonstrated a high-performance rate of 93%. Therefore, it was concluded that the BTCR technique is one of the most powerful methods of predicting mortality in head injuries (Sut N. S., 2011).

In the medical domain, the use of DM improves efficiency by reducing the likelihood of human error for chronic diseases, including diabetes mellitus where the process of mining patient data improved performance and outcomes for the prediction of diabetic individuals, particularly those with Type-2, through a classification technique (Patil, 2010). In the medical field, data management has been utilised by healthcare professionals and organisations with positive advantages by supporting various medical activities (Esfandiari et al., 2014; Jothi et all., 2015; Iru et al., 2016).

DM has been praised for its medical contribution in publications around the world, for medical activities including diagnosis, treatment, monitoring and others (Sarah Itani F. L., 2019). The support of DM is important to help patients in different countries to overcome health conditions such as diabetes mellitus, heart disease and Alzheimer's disease plus many other serious illnesses (Ioannis Kavakiotis O. T., 2017).

A DM methodology is required for the extraction of tacit knowledge taken from our e-tool kit (as shown in chapter 5), knowledge that could help better manage the diabetics. We should first specify the algorithms and software tools suitable to a given purpose (i.e. descriptive or predictive approach) (Kurgan, L.A. and Musilek, P., 2006).

As mentioned, KD and DM are two very promising fields. On one hand KD provides the ability to extract useful knowledge via a non-trivial process, whereas KM enables the dissemination of the discovered knowledge and provides useful ways for successful utilisation of that knowledge.

The integration and combination of KD and KM is needed to maximise knowledge extraction from patients, medical professionals and data and its dissemination of knowledge discovered as explained in chapter 4, which is the end goal of this thesis. DM, in particular KD field, has been successfully utilised to extract knowledge from a great amount of data in many healthcare and medical domains. (Sarah Itani F. L., 2019), In this thesis, we aim to integrate KD and KM to support diabetics in the NHS as this disease is affecting a large part of the population. One of the novelties of this thesis is the creation and dissemination of knowledge within the SECI model.

Chapter 4 demonstrates how DM can extract valuable tacit knowledge relating to diabetes mellitus in particular. This implicit knowledge can facilitate decision-making in relation to diabetic medicine side effect prediction, as well as diagnosis and treatment for both healthcare professionals and diabetic individuals.

The proposed methodology mentioned in chapter 4 on this thesis uses Sentiment Analysis. This is because of its relevance to the studies objective in extracting useful knowledge or extracting the most frequently mentioned words from reviews in a particular field to find beneficial knowledge. As an example, in this thesis the use of medical reviews on a particular drug and found that headache pain was the most common side effect that patients talked about.

KD can be used by healthcare organisations to analyse their datasets to discover valuable new knowledge about their patients in order to provide better healthcare services. KM can be employed to extract and disseminate the best practices of healthcare professionals in order to provide efficiency and effectiveness in their healthcare services. (Almuayqil, 2017).

This research proposes an integrated framework of KM and KD. The framework proposed will be applied to diabetes. The KD part of the framework is responsible for extracting useful trends related to diabetes/medicines complications encountered by diabetics in the UK, who are from different backgrounds and nationalities.

Knowledge Discovery Methods that could potentially be applied in this part of the framework are Data Mining, Text Mining, Web Mining, and Link Mining.

The current thesis, however, focuses on Text Mining (Sentiment Analysis) and Data Mining techniques to extract implicit knowledge relating to diabetes mellitus. The KM part of the framework applies the SECI model to disseminate useful knowledge relating to information that the patients have input themselves. This may take the form of reviews of a specific medication, as used in this thesis self-management and education in order to resolve the complications and diabetes problems discovered from the KD part.

As well as the potential of applying this framework to the public health sector in order to minimise the effect of diabetes the SECI model is used in this framework to disseminate knowledge discovered. This knowledge is extracted from the information gained from the dataset/database and other sources from within the e-Toolkit such as the blogs. This is explained in Chapter 4. "The SECI model is a useful theory that converts human knowledge and expertise between tacit and explicit." (Ikujiro Nonaka, 2008).

2.6   SECI - Socialization, Externalization, Combination, Internalization (knowledge creation model)

In this research, the four modes of the SECI model are applied to create and disseminate knowledge associated with diabetes mellitus within our e-Toolkit. To understand the SECI knowledge creation model, the section below is a literature review of the SECI model and its use as it is one of the main objectives of this thesis.

SECI is a knowledge creation model consisting of the three elements of process, knowledge and creation. This occurs through the tacit knowledge and explicit knowledge conversion. (Ikujiro Nonaka and Hirotaka Takeuchi, 1995) highlighted that although organisations produce and achieve knowledge effectively, the capturing and sharing elements lack enormously. However, Nonaka and Takeuchi proposed a knowledge creation model with the four fundamentals of Socialisation, Externalisation, Combination and Internalisation (SECI). This model has been employed in various fields for the intent of knowledge creation and dissemination.

Numerous studies have revealed effective use and implementation of the SECI model, such studies include a library application to build a knowledge sharing (KS) system (Jixia Cao, 2010). In an different study, the knowledge creation model was operated in software development to help the investors in the course of the software requirements in elicitation, specification and validation (Chikh, 2011). Furthermore, the SECI model was used in a study done by Yangyang Li to examine the performance in the manufacturing division and demonstrated an effective operation (Yangyang Li a, 2009).

Nonaka and Takeuchi identified two categories of knowledge which are explicit and tacit. It is significant to share and convert explicit and tacit to have the ability to produce knowledge. Explicit knowledge can be communicated in systematic language, while tacit knowledge is attained via skills and knowledge and cannot be effectively articulated.

Knowledge creation occurs when tacit is converted into explicit knowledge (Ikujiro Nonaka and Hirotaka Takeuchi, 1995). Organisations use current knowledge assets to form original and new knowledge through the use of the SECI model in the shared context area.

The four SECI conversion modes are shown in Figure 4.
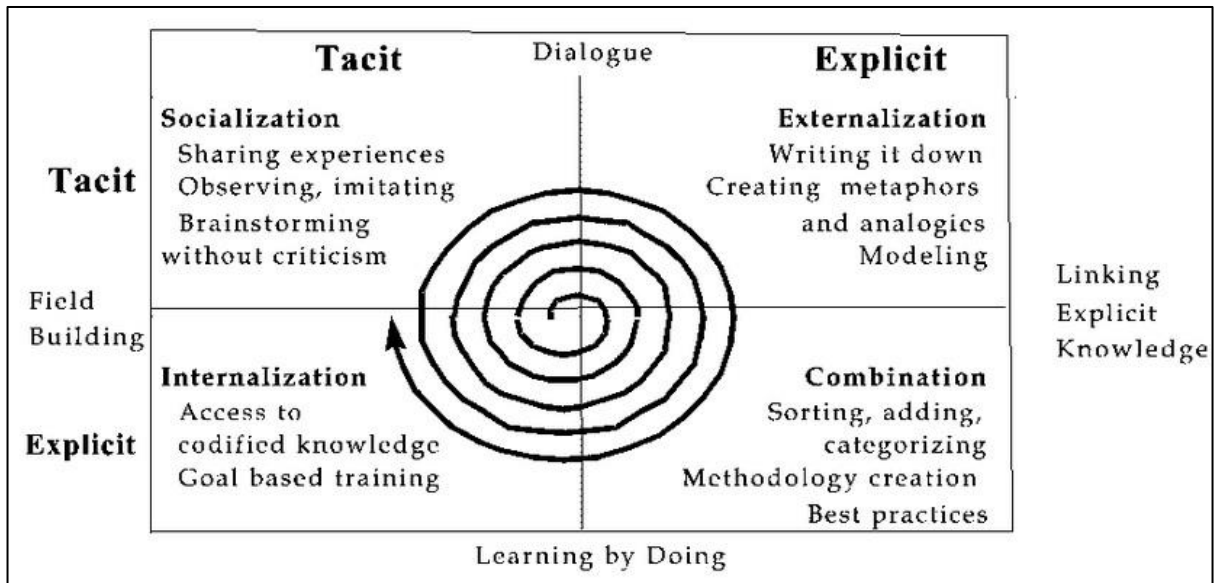
Figure 4 - The SECI Model Source: (Ikujiro Nonaka and Hirotaka Takeuchi, 1995)

To better understand the conversation from Tacit to Explicit (Takeuchi Hirotaka, Shibata Tsutomu, 2006) has been illustrated in Figure 5.
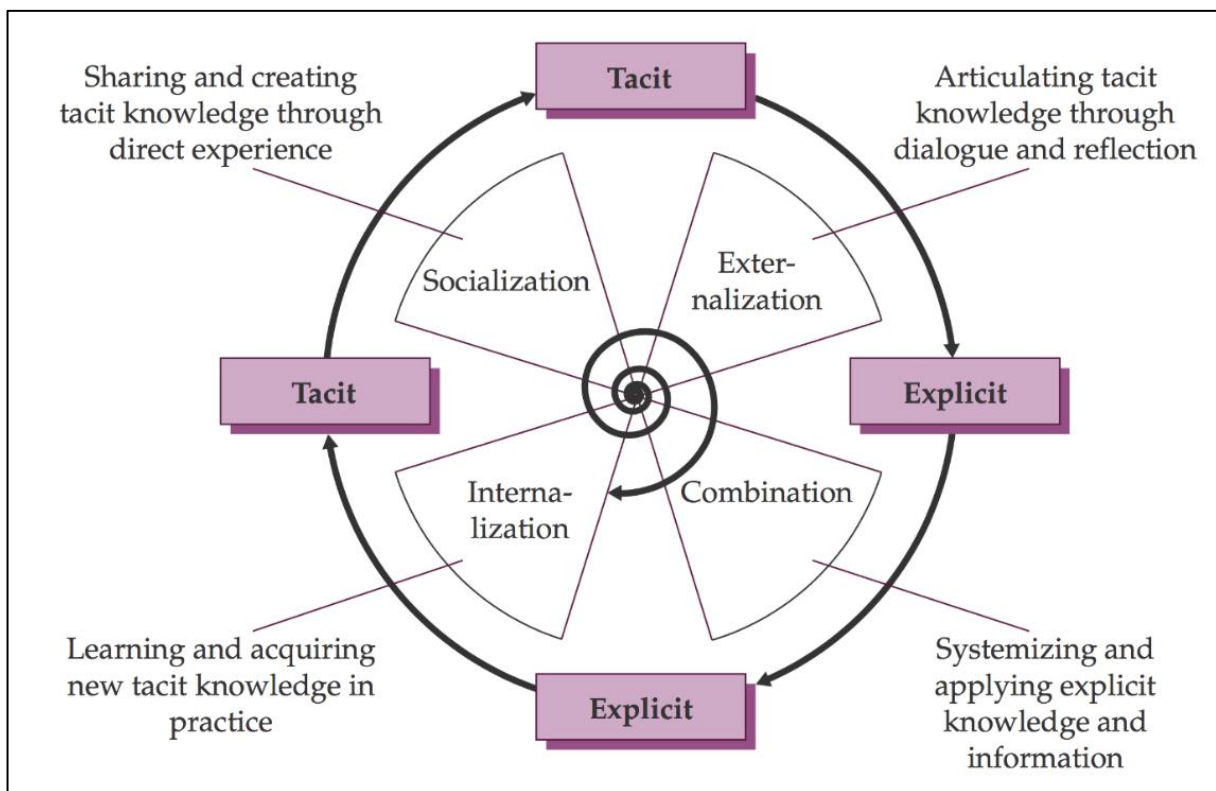
Figure 5 – SECI Model By Takeuchi & Tsutmu - Source: (Takeuchi Hirotaka, Shibata Tsutomu, 2006)

The four types of SECI knowledge creation process is explained below:

- **Tacit to Tacit (Socialisation):** This mode includes "social interaction to enable the conversion of tacit knowledge from one person to another without the use of language" (Ikujiro Nonaka and Hirotaka Takeuchi, 1995) Valuable tools such as observation, imitation and brainstorming are required to accomplish such a conversion as illustrated in figure 6.
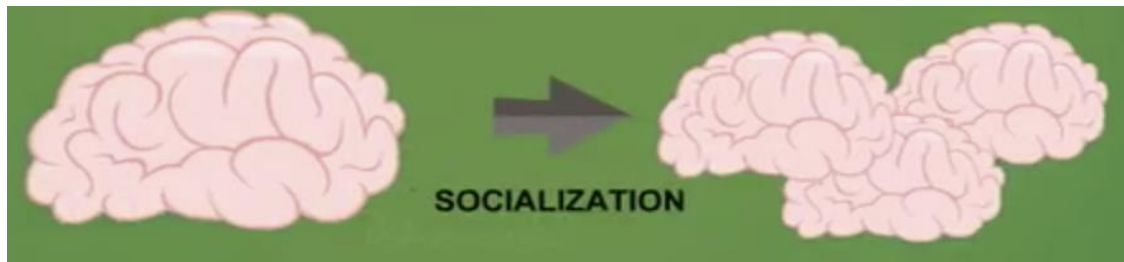
Figure 6 – Illustration of the Socialisation node – SECI Model – Source: (Chandra Wiranto, 2019)

- **Tacit to Explicit (Externalisation):** This mode allows the tacit to be converted to explicit knowledge (Nonaka, 1994). This means the knowledge is converted into readable and interpretable data by users through collaboration, interchanges and informal meetings. The newly assembled explicit knowledge can be coded and logged via various means, such as documentation and footages.
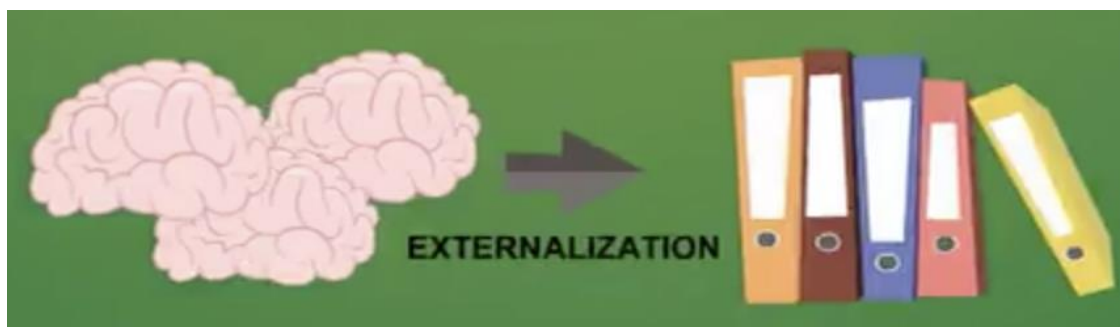
Figure 7 – Illustration of the Externalisation node – SECI Model – Source: (Chandra Wiranto, 2019)

- **Explicit to Explicit (Combination):** This is the third mode of the SECI model where knowledge is transformed from explicit knowledge to explicit knowledge (Nonaka, 1994).which is the process of integrating concepts into a knowledge system. Official organisations meetings are very assistive combination tools (Ikujiro Nonaka and Hirotaka Takeuchi, 1995)
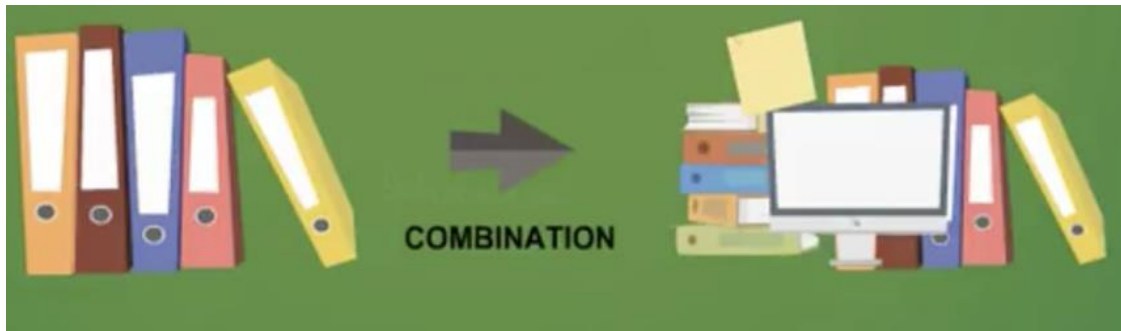
Figure 8 – Illustration of the Combination node – SECI Model – Source: (Chandra Wiranto, 2019)

- **Explicit to Tacit (Internalisation):** Tacit knowledge is converted and develops a section of an individual's brain (Ikujiro Nonaka and Hirotaka Takeuchi, 1995). To explain, the internalisation mode includes a "learning by doing" technique and the effective people who adopt the tacit knowledge will ultimately acquire the "know-how" skills.

Figure 9 – Illustration of the Internalisation node – SECI Model – Source: (Chandra Wiranto, 2019)

In summary, despite the significance of knowledge creation models, evidence have revealed that there are not enough support models in KM. The SECI model seems to be the only highly utilised model for KM and knowledge creation in different domains (Yangyang Li a, 2009).

According to Snowden (2002), the knowledge conversion model of Ikujiro Nonaka and Hirotaka Takeuchi contributes to the deficiency of other KM models and tools because the focus on this model has dominated the field of KM. Therefore, the model is adopted for the following four reasons:

1. The model examines the link between the two contrasting sorts of knowledge, explicit and tacit, which have been outlined above and which are highly utilised in healthcare (Dalkir, 2011).

2. It facilitates the transformation between explicit and tacit knowledge, a necessary process to utilise from the availability in knowledge in the healthcare domain (Dalkir, 2011).

3. The model only concentrates on both dissemination of knowledge and takes into consideration the process of knowledge creation, which is also a valuable process in the area of KM.

4. It is widely utilised in different domains, such as organisational learning, product development and IT.

2.7   Knowledge Management in the Healthcare Domain: Related Research

Several studies have focused on integrating the knowledge management domain with the healthcare domain. The concept of knowledge management is currently emerging in the healthcare sector (Kothari, 2011; Quinn, 2014).

Patients and physicians could benefit from the development of applying knowledge management strategies between organisations and family practices, and amongst family practices themselves (Orzano, 2008). It is beneficial to adopt the strategies of the knowledge management domain in the healthcare sector particularly for medical staff, patients, public health, and organisations (Chen Y. L., 2012).

Based on their studies, it can be concluded that their research could potentially have beneficial effects on diabetes management. This potential is assured by (Chen Y. L., 2012) who claims that infectious diseases can be managed by efficient knowledge management within the hospital.

A similar observation was found in the context of this thesis. Currently, medical staff in NHS practices report that knowledge is not being utilised efficiently to help diabetes disease management (Ferlie, 2006). Thus, the researcher in this thesis proposes a knowledge management framework in order to help the diabetic community which will result in a reduction of costs to the NHS.

This framework can enhance the performance of the healthcare systems by appropriate dissemination of medical knowledge so that information is managed and retrieved by others when required. Almuayqil 2017 and Al-Madi 2017 both focus on the domains of knowledge management in diabetes and consider the overload information in KM and healthcare. But unfortunately, they do not consider the benefit of KM in other types of illnesses, which is vital as in the medical domain there are a vast number of diseases (such as liver and heart disease) that could benefit greatly from KM (AL–MADI, 2017; Almuayqil, 2017).

A framework by Almuayqil, was proposed to help the National Healthcare of Saudi citizens effectively and can also support current government initiatives. However, this study is limited to one country only, Saudi Arabia, and the target audience for the framework is focused on diabetic patients only.
While our proposals will be made available for anyone in the world and have the potential to help a number of diseases and can be applied to commercial or other fields.

In previous studies, patient data could not be shared with members of a healthcare team, including GPs, as professionals had not been involved in the studies. The research in this thesis focuses on multi-nationals and improving communication between healthcare professionals and patients In addition it will create a hub for patients to share their reviews on certain topics. Those reviews will help in the process of generating new knowledge that will benefit the medical health care professionals.

The proposed framework in this study includes both healthcare professionals and the patient to communicate about their illness, with data and possible predictions of treatment, it will also identify the side effects of medicines from discussion between patients which they both can edit and use. This research is inclusive and includes a diversity of income, gender, age, race etc. This demonstrate that the research involves a range of patients and even includes doctors. It aims to be a helpful and valuable system in combating diabetes, with applications to further illnesses in the medical field, including updating doctors in the NHS, and helping to diagnose or decide what treatments are appropriate.

Additionally, Almuayqil successfully identified the barriers affecting e-health in Saudi Arabia, possibly due to there being many healthcare organisations, affecting the issue of information overload when one or more barriers emerge, However, their model has only been designed and constructed but has not been hosted, tested, publicised or revised.

As such, the Almuayqil e-Toolkit is insufficient for use as it has not yet been applied in the real world and not made available for use by diabetic patients. Additionally, the framework of the study was limited by lack of knowledge, challenge and issues affecting IT specialists. (AL–MADI, 2017). Their focus study also lacked information technology specialists, using only conceptual aspects, such as medical staff's perceptions and experiences.

There was no consideration of the technological aspects when structuring the proposed KM framework for reducing issues associated with diabetes information overload in hospitals. In a case study conducted in Jordan, Hussain, F. & Raza, S. has stated that a combination of conceptual and technological changes are necessary to improve healthcare (Hussain, 2004). In conclusion, the research study by Al-Madi was applied to a small number of clinics in Jordan, only seven hospitals, which cannot be generalised as it is limited to only a small group of patients. Similarly, the work by Mirza aimed to highlight the processes and impact of knowledge management and improve healthcare delivery within two developed countries, the USA and Sweden (Mirza, 2009). However, due to the incomplete structure of their knowledge management framework, only a small number of medical staff respondents and an insufficiently large sample of participants, the work was very limited and not widely used as a research case study and cannot, therefore, be considered reliable information.

The e-tool kit presented in this work is to be applied in the NHS in the United Kingdom which hosts a variety of people from different ethnicities, races and backgrounds. 12.8% are non-white, compared to only 10% non-Arab in Saudi Arabia, and a mere 2% non-Arab in Jordan. Studies from these countries only represent people with illness from the same background and cannot be applied elsewhere unlike the work presented here (CIA World FactBook, 2018). This diversity strengthens the basis and framework of this study.

Summarising the above review, it can be seen that when there are many flaws in a research study, the proposed work becomes limited and cannot be generalised. However, it can be stated that there are several valuable points that can be concluded from the review of the above research:

- Some research studies were limited to the technological aspects, and other studies were limited to the conceptual aspects.
- Some research studies were limited to knowledge management as it interacts with the general healthcare domain, while other studies were limited to how knowledge management interacts with a particular type of illness.
- Some studies were only applied to one country and have not considered different cultures, backgrounds and ethnicities which are all relevant to the UK.
- Some of these studies were limited to investigating the benefits of improving knowledge management in healthcare industries and did not focus on the barriers and consequences that are likely to occur within these industries.

Other studies took barriers and consequences into consideration, as these issues are likely to affect any healthcare organisation.

While some conducted frameworks do combine both knowledge management and healthcare, these frameworks were only partially completed and lacked a full structure.

Regardless of all the identified limitations, these studies demonstrated significance in that they incorporated the areas of KM and healthcare, within the environment of a medical organisation. It can be concluded from these studies that the aim of this significant integration of knowledge management and healthcare is to improve the effectiveness of any healthcare industry in order to provide proper and efficient healthcare services to any medical organisation.

2.8   Current Available Monitoring Systems to help diabetics

This first part the literature review addressed the medical aspects of diabetes, the second part covered the use and application of KM, KD and SECI.

Consequently, the following sections discuss, in detail, existing diabetes monitoring systems and highlights their limitations within the context of the e-Toolkit in helping patients monitor their diabetes regardless of what device they are using, accessible anywhere in the world as well as disseminating knowledge collected from the e-Toolkit using data mining techniques.

2.8.1   Old-style Methods

The more medical professionals and scientists find out about this disease, the clearer it is that recording and monitoring essential health data is important to managing diabetes. This is stressed and explained to the patients at every medical appointment (Anderson, 1992; Resnicow, 2007). Every diabetic patient learns how to achieve this task in different ways. Most learn which system is best for them by giving it a try and see if this system suits them.

The most known methods are recording and writing readings in a scratch pad, blood testing record book, or a paper journal (American Academy of Family Physicians, 2018). Some people record their data electronically using a word processor, spreadsheet or some sort of diabetic monitoring software (Jon, 2018).

Others check their readings using a timetable or day organiser, some don't bother writing any readings down and rely on the past three-weeks of readings that are recorded in their glucose monitoring device / systems (diabetessupportsite.com, 2018).

2.8.2   Software and Mobile Apps

Diabetic patients are now able to use improved technology as it has become more widespread and easily accessible, assisting them in controlling their condition. Applications, software and websites are continuously being established, aiding patients whilst tracking and recording their readings.

The majority of glucose meter manufacturers, for example, Freestyle, Bayer and Accu-Chek, provide software or website access. It is possible for it to be simple e.g. a log accessed through their commercial website. Smart phone applications are also available, for example, Diabetes Tracker UK, Lenny and Life Scan.

Once a specific, sponsored pump or meter has been purchased, the patients will be able to access and use the website. Many patients use multiple meters as they are unable to merge all their data under one specific system. Current systems require external hardware or specific software, restricting patients to using one device to record their information, or carry additional equipment. The majority of systems add additional charges for subscriptions to access websites for extra software and equipment such as cables.

Not all of these systems allow users to transmit recorded data to medical professionals. However, if there is a small chance that this feature is available, then diabetic patients can only email those readings to medical professionals. Despite this, it is common that the patient prints their recordings and has to visit their health centre to deliver them in person. These systems only record insulin measurements, insulin pump and blood glucose readings. However, there is no software to enter vital data that a patient must track in maintaining their condition.

Below is an explanation and discussion of the most well-known systems. These system details are included for the purpose of relating their system features against the proposed system. They were chosen as they offer an overview of devices that are currently available to diabetic.

2.8.3   Roche Accu-Chek System

Accu-Chek is a product by Roche Diabetes Care based in the United Kingdom, the best known, global producer of diabetic equipment (Roche.co.uk, 2018). As mentioned on their website - "the Accu-Chek portfolio offers people with diabetes and healthcare professional innovative products and impactful solutions for convenient, efficient and effective diabetes management. It encompasses blood glucose meters, insulin delivery systems, lancing devices, data management systems and education programs" (accu-chek.com.ph, 2018).

The Accu-Chek Smart Pix system has further advantages including the ability to scan and import the data readings, but this can only be done from an insulin pump or glucose meter from Accu-Chek and those readings are shown on the user's PC.

The information exchange is filtered and organised automatically which increases the accuracy and decreases the amount of time wasted. Rather than the patient importing the data readings in manually, the system can do this automatically into ordered outlines, which reduces the possibility of any errors being introduced through mistyping. Parameters can be set to focus on specific sets of readings or information.

There are disadvantages to this system that are worth noting. To use this system, patients will need an external mass storage device (MSD) which must be purchased.

Figure 10 – Accu-Chek Smart DM System – Source: (accu-chek.co.uk, 2018)

Although the organisation guarantees, no specific requirements for the application to be installed on any pc, however, they do have specification on which operating systems it

requires for the software to operate. These include an up-to-date operating system higher than Windows XP, Vista, 7 or 8 as well as a USB port, internet access and a minimum resolution on the computer monitor.

These conditions force limitations on the device's usage. For example, why does the system specify web availability, as everything in this system is stored in the device itself and not stored online? Additionally, there is no option to email the information to medical professionals which can be a limitation, as diabetic have produce paper report of their data with this system (Darwish, 2012).

In the UK it will cost £24.99 to use this device and the software that comes with it, although this cost fluctuates dependent on the UK market. However, in the US, the same device is valued at $99.00 which is approximately 300% higher than the price in the UK.

Additionally, the Accu-Chek system only stores a diabetic patient's blood glucose readings but there is no apparent option to include other factors such as diet, changes in mood, carbohydrate consumption, exercise, etc (Diagnostics, 2010; accu-chek.co.uk, 2018).

To summarise, The Accu-Chek Smart Pix system is finely established but is quite limited, with an excessive number of restrictions on its usability. Although the software transmits the insulin and blood glucose readings to the computer and assembles the data into diagrams, it disregards the various further factors a diabetic patient must check for good health. Most significantly, the data that has been recorded is not made easy to communicate to medical professionals for their use.

2.8.4 OneTouch Reveal app

OneTouch (known as LifeScan) is described as "a mobile / smart phone application made by a Division of JNJ, a well-known pharmaceutical company in America" (Fahad Darwish, 2019). It was presented during the preview of Apple iPhone Operating System 3 launch. Even though LifeScan produce a range of medical products they have a focus on diabetes through their LifeScan Diabetes Institute based in the USA (LifeScanInstitute, 2018).

Figure 11 - OneTouch (iPhone app) – Source: (LifeScan, 2018).

OneTouch was described as an app that would connect with the blood glucose device via either an iPhone cable or Bluetooth. Unlike Accu-Chek which only reads blood and insulin readings, OneTouch has the capacity to label notes to readings to show, for example, if the reading was taken before or after a meal (CBE, 2019; Fahad Darwish, 2019).

OneTouch also offers a feature called "Meal Builder", to assist diabetic users in calculating the correct foods to eat. Glucose readings are transformed into charts, and an algorithm is created to distinguish and assess the required intake of insulin a diabetic must take. Additionally, OneTouch stated that the app can share readings if diabetic patient wanted to share their readings, including their doctors. (LifeScan, 2018). Unfortunately, they never successfully executed this feature (Dolan, 2009). The app can only share data by emails, text messages or print out but not directly to the medical health care as our system is proposing (LifeScan, 2018).

Another huge disadvantage of LifeScan was the recall and replacement of the device in 2013 because of a problem whereby "at extremely high blood glucose levels of 1024 mg/dL and above, the meter will not provide a warning that the blood glucose is extremely high and will shut off" (LifeScan.com, 2013). Unfortunately, it was due to lack of money, time and resources and did not in any way assist the diabetic community.

To summarise, The LifeScan application is a good example to show the risks of an interface with an over-limiting scope. Regardless of the prestigious introduction it was given during the Apple's iPhone iOS-3 launch, it was not effective for the majority of diabetics.

2.8.5   Diabetes UK App

During early 2012, the then Health Secretary Andrew Lansley made the proposal that GPs should urge their patients to expand their use of free smartphone apps with a specific end goal to help patients better manage their health. A particular app that was mentioned was Diabetes Tracker UK, developed by Diabetes UK, one of the largest health charities in the country. (CBE, 2019) (Diabetes-UK, 2012) (newscientist.com - Matt Reynolds, 2017) (Darwish, 2012)

This app helps patients to not only record their blood glucose & insulin readings. They can keep track of their calories (food consumed), carbohydrates, body weight and ketones - (American Diabetes Association, 2012). As mentioned in the above examples, readings can be charted in graphs. This application also shows day by day averages and enables the diabetic to record notes which can be shared at their next meeting with their doctor.

To conclude, this app goes further than other applications. It is free, and not attached to a particular meter, it is accessible to significantly more patients. However, it does not allow patients to record all their important information in one place and does not connect directly to medical professionals.
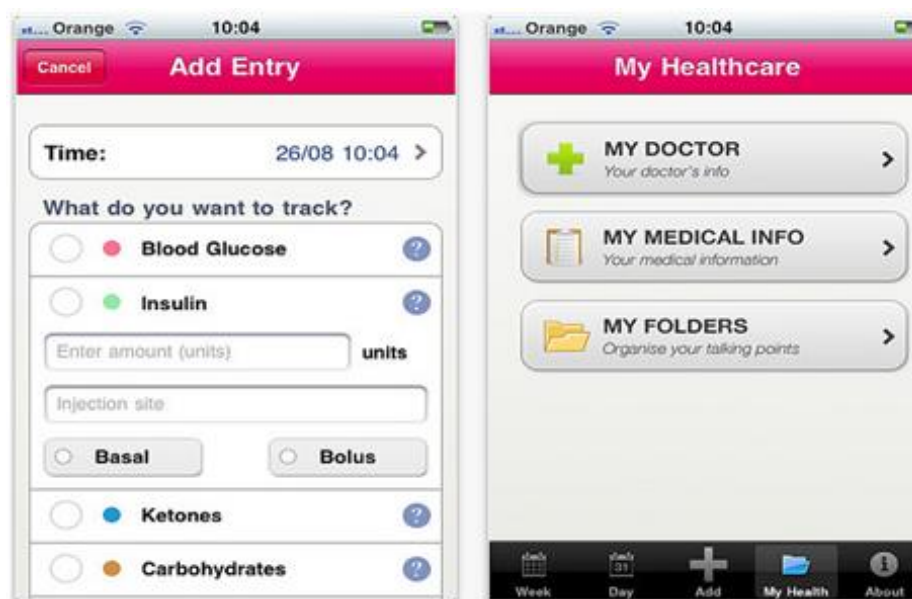
Figure 12 - Diabetes Tracker UK – Source: (Diabetes-UK, 2012)

In spite of the fact that this app claims to have the choice of emailing the information to the medical staff, the main obstacle with this emailing scheme is, as was noted at the end of chapter 1, because there are no schemes to make sure that medical professionals can obtain, analyse, and record the emails.

The obligation still rests with the diabetic to send the charts by printing out the emails and physically taking those data to their healthcare provider when required. This is can be considered wasteful, when data could be sent electronically and only required the use of resources employed when booking appointments. Finally, the application is no longer available to download on Apple store and Diabetes UK have decided to stop supporting it even though there are some patients using it.

2.8.6   Glooko Diasend

Diasend was at first made by the Swedish organisation Aidera AB. The name was changed in May 2012 to Diasend AB (Bloomberg.com, 2018). In 2017 Diasend merged with Glooko, a company that was founded in 2011 by the internet pioneer Yogen Dalal. Glooko claimed to offer a unified diabetes management platform for the Diabetic community (Moukheiber, 2014) (Diasend.com, 2018).

Figure 13 – Diasend System – Source: (Diasend.com, 2018).

The main function is to transfer data from blood glucose devices or insulin pumps to the server. All things considered, Diasend is not designed to be a system for patients to record all readings, symptoms and more, or where medics can go to gain a better understanding and overview of their patient's wellbeing.

Diasend functionality relies on a 'transmitter', which is a device that the user must physically connect to their computer. The purpose of this is to allow readings to be detected and imported automatically by the transmitter, from the user's insulin pumps and glucose meters. All of the information obtained is saved online using Diasend's' website where patients can make their own accounts they can gain access to using any computer with an internet connection. Although this method improves accessibility for patients to view their readings from anywhere, they can only submit readings if they carry the heavy and bulky transmitter with them (Diabetes.co.uk, 2018b).

The measurements of this gadget make it unfeasible for a client to carry with them, so the patient is restricted to transferring data most likely only once per day, when they have access to both the transmitter and computer at home. Another limitation is that specific software for Diasend (Diasend SOFT) will need to be installed onto the patient's computer before they can begin. They will not have access to view, add or edit data from other devices such as tablets and mobile phones. Therefore, it is not likely that patients will use this system especially if on holiday or travelling etc.

Aside from the transmitter and its cables, the patient will need a USB device and its cables, unless their equipment (glucose meters and insulin pumps) have Bluetooth or infrared technology. Most diabetic devices lack this level of connectivity. Nonetheless, once the data is transferred to Diasends website the results can actually be viewed in diagrams, charts and graphs, as similar to devices mentioned above.

Diasend is not accessible to just anyone who owns the devices. For a patient to have access to an online account a medical professional must already have enrolment in the system and can then authorise the creation of the patients account. This medical professional would then have access to the patients account and be able to see all the readings the patient submits. As mentioned, the readings are limited to insulin and glucose results only, and in order to benefit from Diasends' features the patient must purchase a specific device that is compatible (glooko.com, 2018). Finally, Diasend requires users to purchase a subscription. The exact cost of the subscription is not clear however according to Diasend's terms and conditions the subscription cannot be cancelled, and the charges cannot be undone. It is also not evident if the Diasend SOFT software is free to install, potentially making this a very pricy system to use (Diasend, 2017; Lamb, 2008; Aidera; glooko.com, 2018).

To summarise, Diasends' functionality is limited in terms of the data that can actually be recorded and it is financially prohibitive due to the costs of equipment and software. The success of this system also depends on participation from medical professionals who firstly must be willing to be enrolled on the system and then must be able to commit to accessing the patients updated data frequently enough. Therefore, it can be said that Diasend is the most restricting system reviewed in this research and does not address the requirements of the diabetic patients for convenient centralisation, where ideally all important data would be

documented and shared. This system was used in our comparison as it was one of the few offering the feature of allowing medical professionals to gain access to patient's accounts, viewing their readings easily in real-time without having to contact the patient via telephone or email etc to receive an update. Aside from this the system does not appear to have any similarities with the proposed e-Toolkit.

2.9   Summary

Despite the fact that innovation is frequently developing and there is a conscious effort to make applications and services for the specific use of diabetic patients, it can be concluded that it is difficult to identify a system which genuinely satisfies the requirements of diabetics, where all important data can be easily accessed and recorded.

The difficulty of finding any sort of system which gives healthcare professionals ongoing access to those data in real time is evident. The objective is to help diabetics with the volume to log in and share all their vital data with their medical team, while having a hub to share their views or opinion, in an organised and effective way is yet to be fulfilled. Equally critical is the ability of medical professionals to provide feedback to the diabetic patient in case of an emergency. These features will be some of the main functions of this e-Toolkit (Fahad Darwish, 2019).

The table 1 shows the differences between the above systems and the proposed system:

System	Special Hardware Required	Subscription required	Connects to Doctor	Data Mining tools to discover new knowledge with Knowledge Disseminating
Our Proposed Web System	No	No	Yes	Yes
Diasend	Yes	Yes	No	No
Diabetes Tracker UK	No	Yes	No	No
Lifescan	Yes	Yes	No	No
Accu-Chek	Yes	Yes	No	No

Table 1– Systems Comparison

As shown above on table 1, our proposed e-Toolkit will meet the needs that the other systems have not, most importantly the connectivity to the Doctor and the knowledge discovery from data collected from within the e-Toolkit.

CHAPTER THREE: SURVEY DESIGN, DATA COLLECTION AND ANALYSIS

3.1   Introduction

To achieve the aim of this thesis conducting a survey is necessary to provide the basis for the research. This chapter will carry out a survey in order to determine if the proposed e-Toolkit would be generally welcomed and to identify which aspects may make it engaging. The survey will also investigate which data is necessary to be disseminated to healthcare professionals in an aim to bridge the gap that the existing systems lack, as shown in table 1 of the previous chapter. The first part of this chapter investigates the survey techniques to be used, followed by an example of the survey questionnaires used. A comprehensive analysis of findings from the survey results will be broken down in the conclusion of this chapter.

3.2   Survey data design, collection and analysis

This part of the thesis includes the findings and breakdown of the surveys which were given to the medical professionals and diabetic patients. As discussed above, the results were used to establish if the proposed e-Toolkit would be generally welcomed and which aspects may deem it more appealing. The initial priority was to see which aspects potential users of the service rendered as crucial, and the perceived importance of those features.

3.2.1   Survey Research process

The survey information that was gathered will be presented with the following structure: firstly, a template of the survey sheet used (survey development), secondly an account of the results (survey administration) and thirdly a dissection of those results as illustrated in figure 14 (Fahad Darwish, 2019).

This methodology is described in Angela M's book titled "Survey methods for educators: Analysis and reporting of survey data." The process is outlined in figure 14.
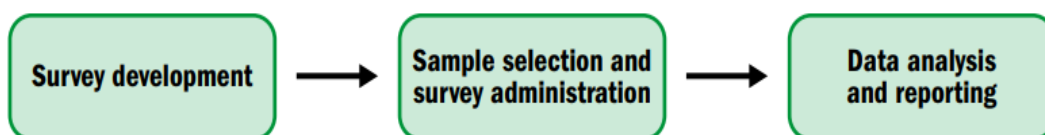
Figure 14 – Survey Process - Source: (Angela M. Pazzaglia E. T., 2016).

3.2.2   Ethical Consideration

This research has followed ethical guidelines as set out in the Ethical and Research Governance by the University of Salford. An ethics class was also taken in order to fully understand how the data collected in the surveys should be anonymous and coded (University Of Salford Website, 2019). Physical copies of the data, for example consent forms, were kept in a secure cabinet inside a sealed room, with access available solely to the researcher. All published data was written in a specific way in order to protect the participant's identity of the study unless prior consent had been obtained and permission had been gained to release that individual's identity. If data sharing by another member of the research team was required, for example, to allow secondary analysis, this release of data was consented for and was fully anonymised. On every sheet given, there were contact details of the researcher if the participants had any issues including withdrawing during the study. The safety and rights of participants was a priority in this research study and were clearly communicated and agreed upon prior to the commencement of the study.

Individuals were approached in the UK and abroad who were either diabetic patients or healthcare professionals to validate that the research goals were essential and had the potential to assist in managing this growing illness. These individuals were approached via online forums, such as Survey Monkey, JotForm, and by physically visiting hospitals to approach healthcare professionals. Participants were provided with verbal and written explanations on their invitation letter, (see Appendix D) and were offered the chance to ask questions at any time during the research process to make sure they were completely aware of how their data would be used before they consented to taking part.

Information Information sheets were provided upon agreement of participation (see Appendix E). These sheets stressed the importance of patient confidentiality and stated that all information given is subject to the General Data Protection Regulation (GDPR) 2018 and the Data Protection Act 2018. The sheets laid out the aims and objectives of this research project, the possible benefits and risks of the study, the significance of obtaining informed consent, reassurance of confidentiality and anonymity and lastly an explanation as to how information will be used.

Each individual who agreed to participate in this study was first asked to sign a consent form, see Appendix F, which stated that participation was optional and that participants had the right to withdraw at any point of the study and remain anonymous. Participants could also request a copy of the study should they desire. The researcher had then been gained an approval certificate from the ethics committee at the University of Salford, see Appendix G.

3.2.3   First Review of Survey

As the e-toolkit grew, more surveys and data collection information were created. This section summarises the responses from the latest series of surveys.

Two surveys were released, aimed at different audiences; one was focused at medical healthcare professionals and the other aimed at diabetic patients with both types of diabetes. The surveys were issued via paper and online through the online survey website 'Survey Monkey' with participation from diabetic patients and medical personnel from all over the world. These included responses from diabetic forums on the web and via email from organisations and diabetics from various places including the UK, Middle East, North America and Australia (Fahad Darwish, 2019).

3.2.4   Hypotheses and Assumptions

During the development of this thesis, a number of people were contacted who suffer from diabetes or who work in the medical community and understand there are many problems with medical testing and communication, websites and software applications. There were many reasons contributing to these issues, but the primary reason for lack of commitment to the applications and websites was low user-friendliness. There were limitations and restrictions. For example, some were only available to patients who have paid for their monitoring equipment, did not consider the needs of the diabetic patient and lacked the ability to communicate information to healthcare personnel.

Taking this information into account, surveys were produced in order to improve existing technology. The focus was specifically on the issues that existing system lacks, which are the connectivity between the patients and medical healthcare professionals and having a hub where patients could discuss issues between themselves. These issues would then be applied to our proposed model in chapter 4, to convert tacit to explicit knowledge.

3.2.5 Survey for Healthcare Professionals

The medical professionals' surveys are illustrated in table 2 below. The surveys were issued to the health care professionals to distinguish the need of such an e-Toolkit and the likelihood of them using it.

University Project

PROPOSAL:

The research survey aims is to create a global e-Toolkit, not constrained by situation, income level or diabetic equipment the patient uses, to provide and support diabetes education and management. There will be numerous features including blogs where patients can talk amongst themselves. The e-Toolkit will act as a hub for patients and health care professionals to communicate as well as provide a space for recording daily measurements such as glucose readings, food diary, insulin dosage and more relevant results and information. The patient's health care providers would then be able to view this information in real time, which would in theory enhance precise and convenient communication between the patient and the expert.

If the patient's readings are abnormal or a negative movement is detected, a feature may be incorporated within the e-Toolkit to alert the health care provider of this latest information and allow them to help the patients in such an emergency. This system could possibly help doctors and other health care providers stay up-to-date with tracking and monitoring the patient's healthcare.

HOW YOU CAN HELP:

You have been selected to participate in this short survey as you are a healthcare professional. The survey is comprised of yes or no questions, multiple-choice questions and the opportunity to add any further suggestions or leave feedback.

Participation is voluntary. You may withdraw from the study at any time and remain anonymous as stated on the consent form which must be signed prior to the study.

If you have any questions, concerns or complaints about the study at any stage, you can contact: the lead researcher Fahad Darwish at *f.darwish@edu.salford.ac.uk.*

Q1 - Which of the following best describes your position in the medical field?	
PCP (Primary Care Physician or Doctor)	Educator of Diabetes
PCP's Assistant	Nurse
Optometrist / Ophthalmologist	Dietician
Dentist	Endocrinologist
Mental Health Provider	Pharmacist
Other	
Q2 - Would you be likely to use an e-Toolkit like the proposed one above?	
Yes	No
Q3 - Which of the below would possibly discourage you from using the e-Toolkit? (Please check all that apply)	
Time consumption	Possibility of inaccurate or slow alerts
Inability to frequently use the e-Toolkit	Security issues

Unsure if the patients would use system properly/regularly	Lack of differences from other current systems
Possibility of patients abuse of e-Toolkit to seek attention	Others

Q4 - In your opinion, which of the following characteristics is important for the patient to record on the e-Toolkit? (Please check all that apply)

Weight	Symptoms
Blood Glucose Levels	Changes in Mood
Urine Glucose Levels	Foot sores/injuries
Intake of insulin	Adverse reactions to treatment
Diet	

Q5 - If the e-Toolkit was applied to your medical centre, which staff member, in terms of position, would be allocated to patient's readings and responding to messages and alerts? (Please check all that apply)

Doctor	Nurse
Doctor's Helper or Assistant	Nurse's assistant
Diabetic Point of Contact (DPOC)	Admin/Secretary Helper
None of the Above / Other	

Q6 - Would you approve of a feature which allows patients to input live medical test results (intention of trend tracking)

Yes	No	Maybe

Q7 - Would you find it appropriate to have access to the patient's medical file from the e-tool kit to store and utilise this information?

Yes	No	Maybe

Q8 – Would you find it helpful if a feature would allow you to share information and notes with other healthcare professionals, would you use it?

Yes	No	Maybe

Q9 - Would an option of converting a patient's readings and data into easy-to-read graphs and charts be beneficial to you?

Yes	No	Maybe

Q10 - Would an option to disclose a patient's confidential data to others be appealing, i.e the patient, their relatives or other medical personnel?

Yes	No	Maybe

Q11 - Do you think it would be useful if the e-Toolkit had a notification feature to notify yourself or other medical staff when a patient's record shows a high or abnormal reading?

Yes	No	Maybe

Q12 - In addition to the above feature to be an option in the proposed e-Toolkit, would you prefer a default notification feature that would be applied to all or to be switched on or off and adjusted only for specific diabetics with high-risk?

Default notification feature	Advanced notification feature

Q13 - Would you support a blog feature on the e-Toolkit which would allow patients to discuss a specific medication or problems?

Yes	No

Q14 - In your opinion, what features do you think that it has to be included or excluded for this e-Toolkit to be effective to the diabetic community?

Q15 - Do you have further suggestions or concerns that have not been mentioned?	
Yes	No
Q16 - If Yes, please indicate what you feel is important or suggestions you would like for us to take into account.	
Thank you for taking the time for completing the survey	

Table 2 - Diabetic Medical Professional Survey

3.2.6   Survey Results for The Healthcare Professionals

As mentioned by Angela M Pazzaglia book – To calculate the survey % rate, "it is necessary to divide the number of individuals who responded to the survey by the number of individuals to whom the survey was administered (the sample size)" (Angela M. Pazzaglia E. T., 2016).

The survey response rate is key and ideally should exceed 85 percent, to limit any bias in the results and to eliminate the need to take any further steps. It has been reported by Angela M Pazzaglia that "it is still advisable to check the characteristics of the respondents and non-respondents" (Angela M. Pazzaglia E. T., 2016).

Note: In the medical professional survey it was noted that not all the doctors had answered every single question. This may be due to a disagreement with the questions and answers, or perhaps they were unsure and left the questions out or had limited time to complete the survey.

1.   Which of the following best describes your position in the medical field?
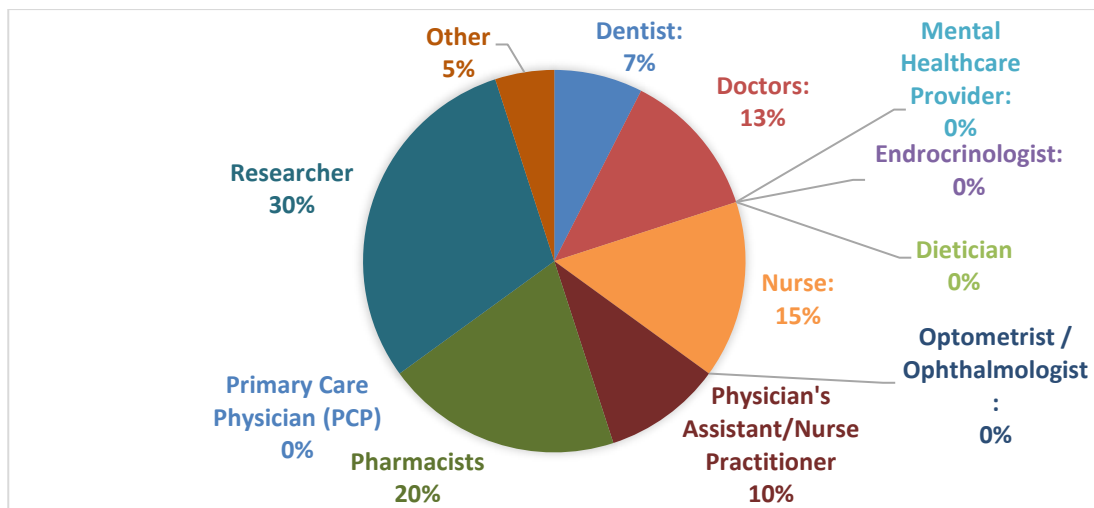
Figure 15 – Medical Professional Survey – Responses to Question 1

This survey was completed by 40 participants with varying medical qualifications; the most common being researchers, pharmacists, nurses and doctors.

2. Would you be likely to use an e-Toolkit like the proposed one above?
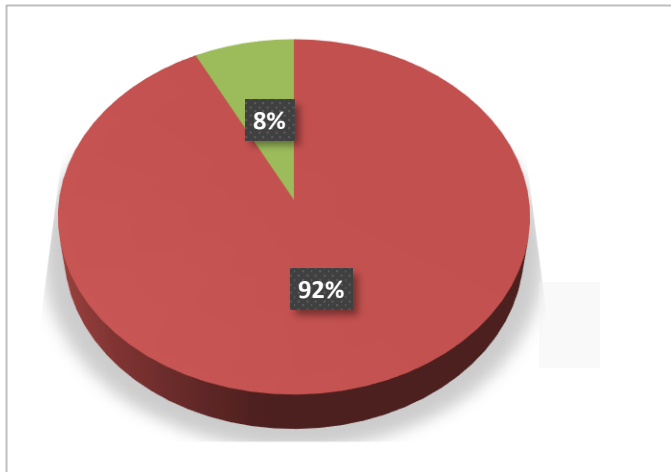
Figure 16 – Medical Professional Survey – Responses to Question 2

Out of 40 respondents, only 8% said they would not use the e-Toolkit which is great feedback as 92% of participants were in favour of the concept. An unexpected response was received by the pharmacists who liked the idea for this e-Toolkit. These answers indicate that it would be beneficial if the diabetic health care specialists can work with each other for the first time observing real life updated information.

3. Which of the below would possibly discourage you from using the e-Toolkit? (Please check all that apply)
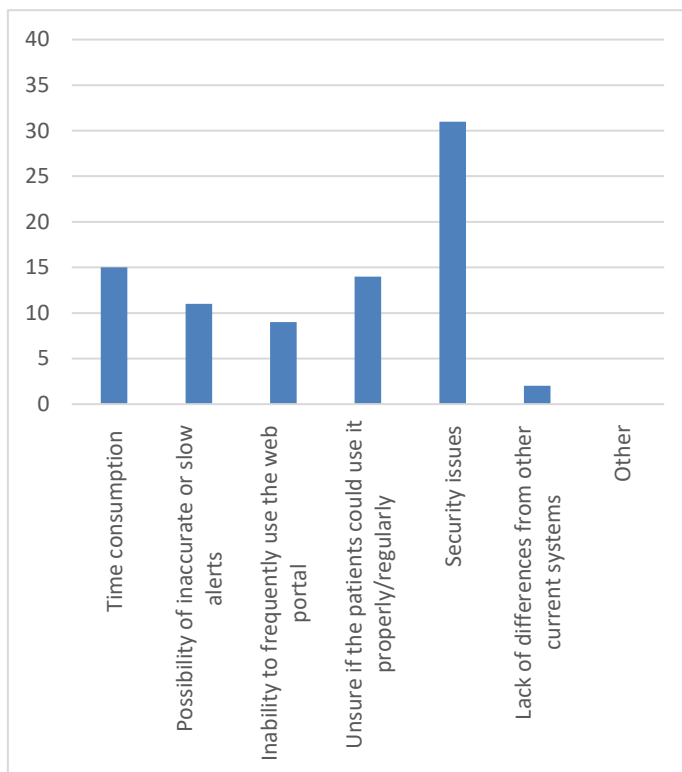
Figure 17 - Medical Professional Survey – Responses to Question 3

The results of question 3 can be seen in figure 17 which shows security issues was the most common answer, following time consumption and unsure if the patient could use it properly/regularly. The researcher has implemented many security measures on the proposed e-Toolkit as discussed in the chapters below.

4. **In your opinion, which of the following characteristics is important for the patient to record on the e-Toolkit? (Please check all that apply)**
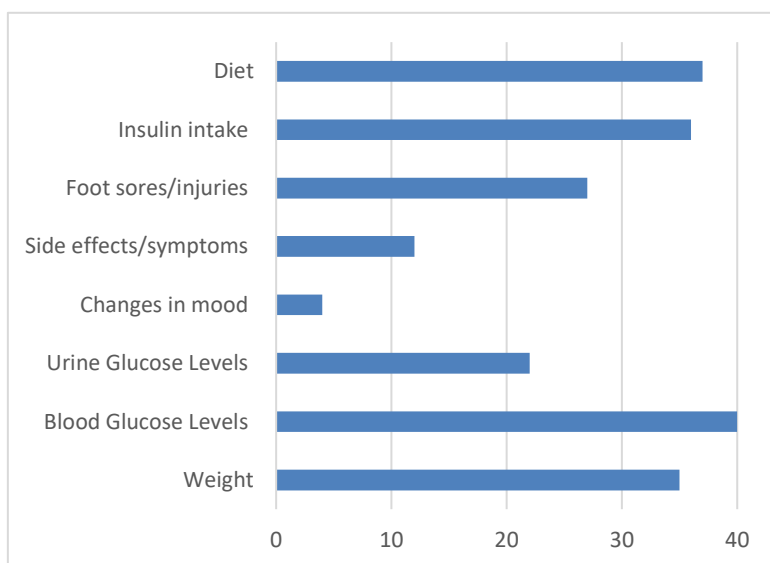
From the results shown in Figure 18, the top four important characteristics responses from doctors were the blood glucose levels, diet, weight and insulin intake.

Figure 18- Medical Professional Survey – Responses to Question 4

5. **If the e-Toolkit was applied to your medical centre, which staff member, in terms of position, would be allocated to patient's readings and responding to messages and alerts? (Please check all that apply)**
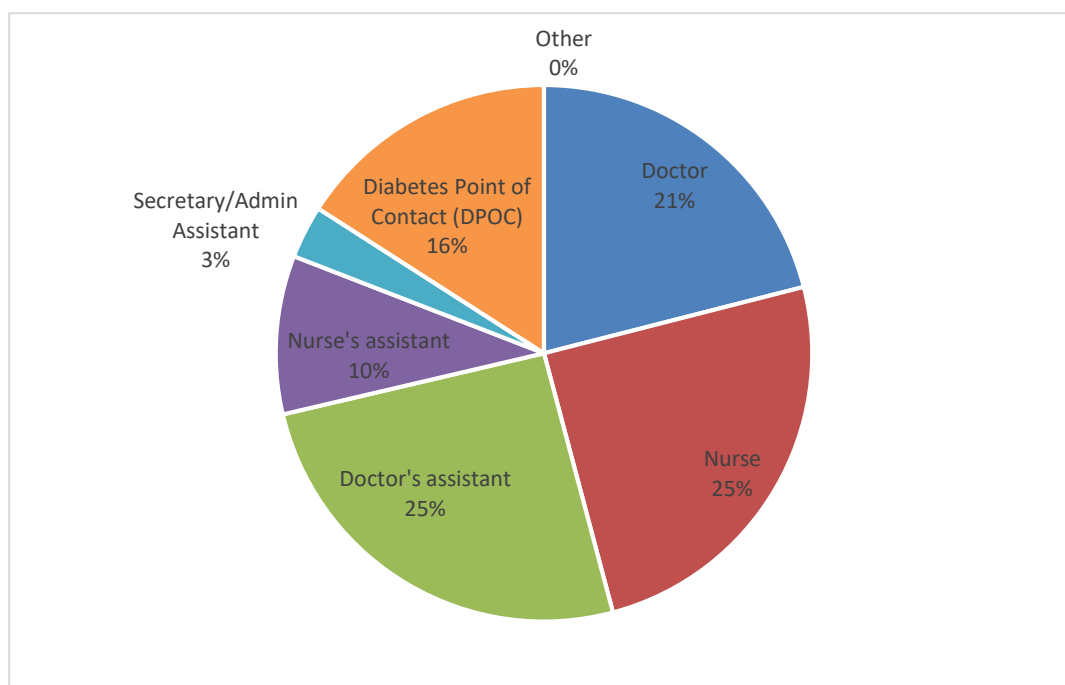
Figure 19 - Medical Professional Survey – Responses to Question 5

Since 'too much time' was one of the main choices given for medical staff not using this system, it was surprising to see that doctor was the third most likely profession to use the e-Toolkit.

6. Would you approve of a feature which allows patients to input live medical test results (intention of trend tracking)
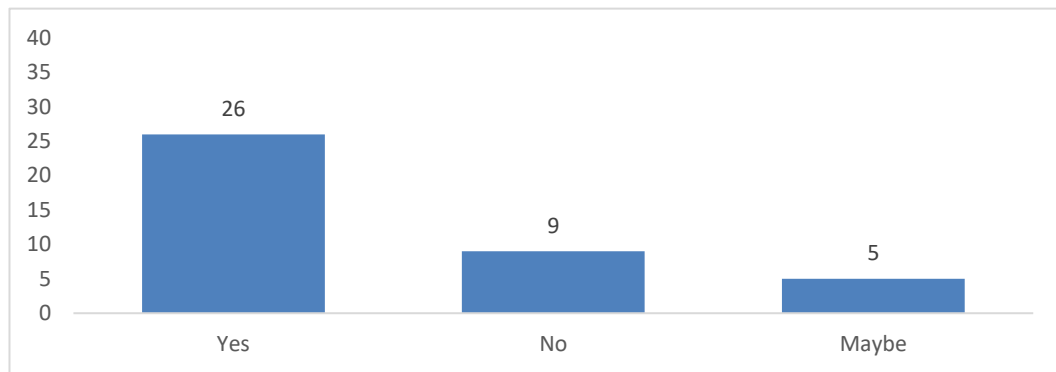
Figure 20 shows that 5 respondents who entered 'Maybe' in this question did clarify they will only back this idea/feature if it were limited to participation from only medical professionals. However, 65% of participants stated that it is useful to import diabetics readings, such as cholesterol, A1C (type of blood glucose test) etc. They explained this further stating the information will be recorded in one suitable site, whereas the 'Maybe' group said that patients may incorrectly input these results and could not be trusted.

7. Would you find it appropriate to have the ability to have access to the patient's medical file from the e-Toolkit to store and utilise this information?
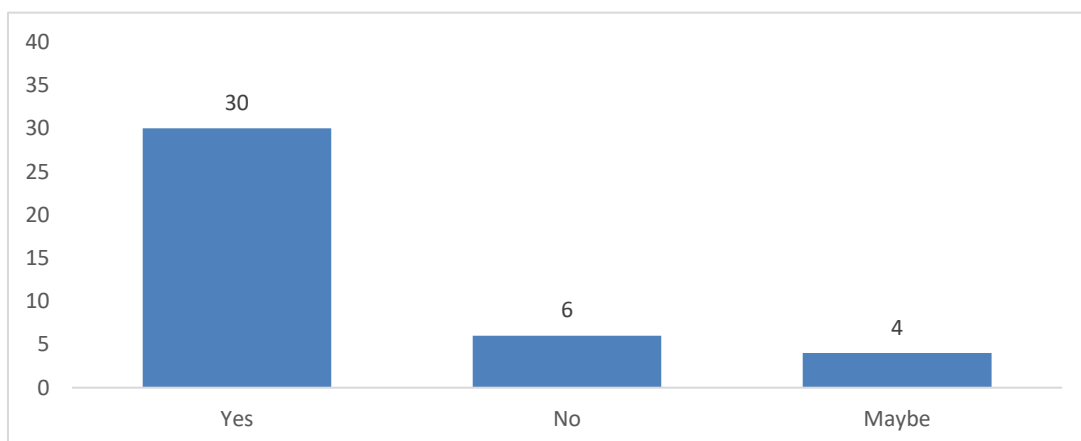
As shown on figure 21, 30 participants said yes to having access to patient data. So, this feature should be taken into account when developing this e-Toolkit. The response to this question was vital to form the concept for the e-Toolkit as this feature was not part of the beta version. However, this feature may be implemented in a future version of the e-Toolkit.

8. Would you find it helpful if a feature would allow you to share information and notes with other healthcare professionals, would you use it?
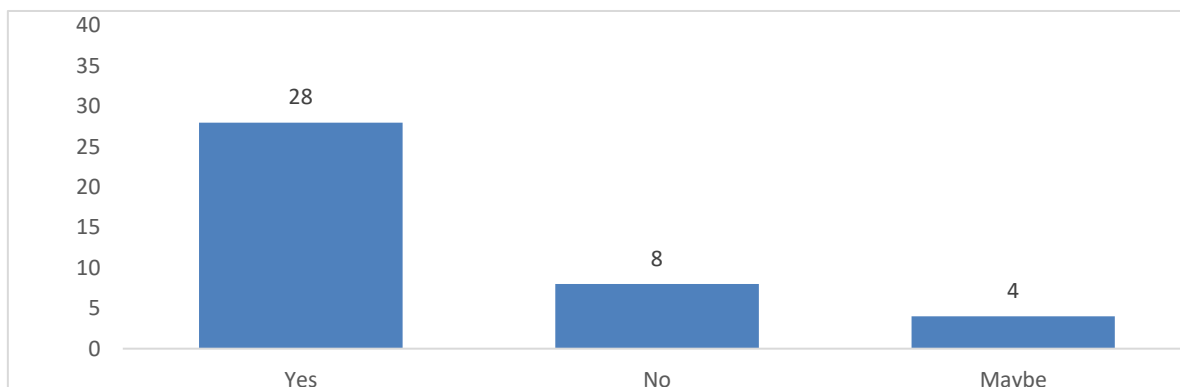
Figure 22 - Medical Professional Survey – Responses to Question 8

At the time when question 8 in the medical professional survey was formed, the original concept was that doctors from different clinics and offices could send information and communicate about the patient, with existing registration with the system. This would allow on-line consultations or referrals take place. But as the concept developed, it was seen that it would be useful for the patient to be able to include and register each member of their health care team, allowing the same information to be viewed. The four respondents who answered 'maybe' raised concerns about the security of any system which allowed the feature.

9. **Would an option of converting a patient's readings and data into easy-to-read graphs and charts be beneficial to you?**
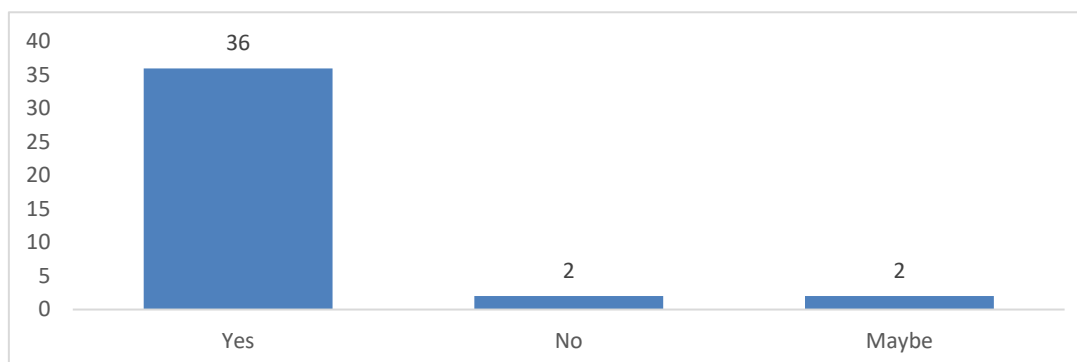
Figure 23 - Medical Professional Survey – Responses to Question 9

Most applications and software about diabetics have been successful when they include the data reviewed in visual forms like charts to allow comparison with another value or to see a spread over a period. However, if the charts are difficult to read or unclear there will be little patient compliance.

10. **Would an option to disclosing a patient's confidential data to others be appealing, i.e. the patient, their relatives or other medical personnel?**
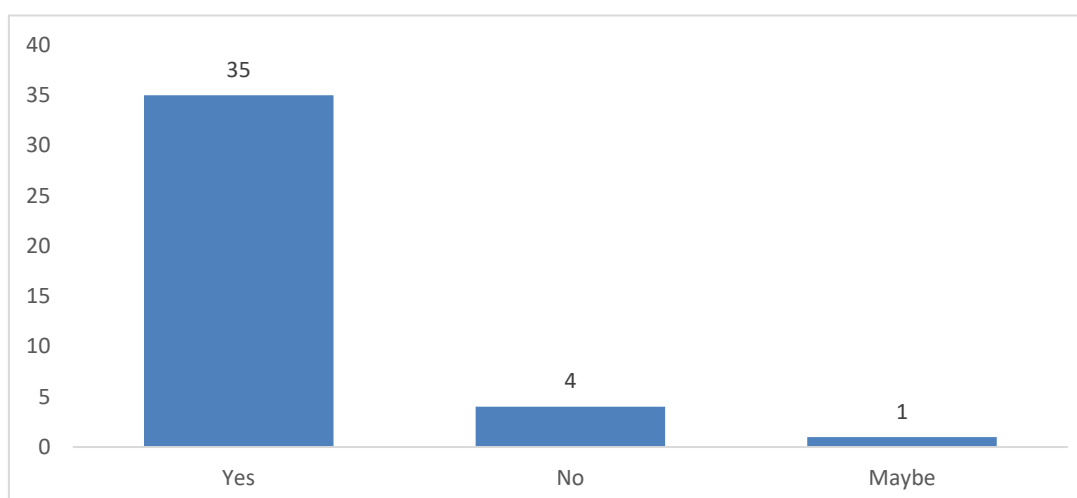
Figure 24- Medical Professional Survey – Responses to Question 10

The main reason for this question was to gain more information on how the graphs and charts would be used, and if they would be beneficial when used to explain relevant information to the patient or their family members etc.

11. Do you think it would be useful if the e-Toolkit had a notification feature to notify yourself or other medical staff when a patient's record shows a high or abnormal reading?
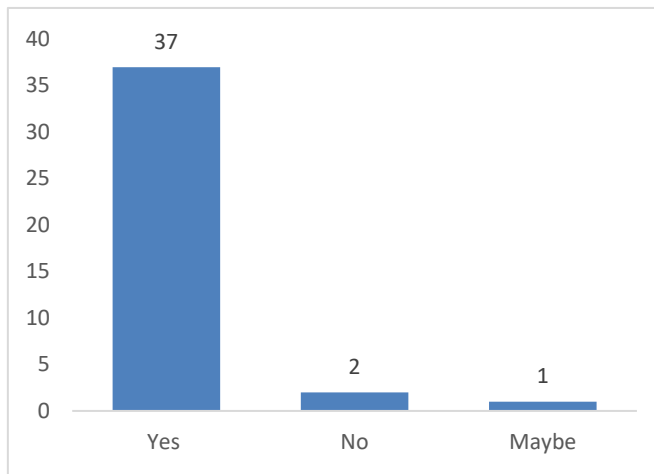
In spite of the fact that this inquiry received a larger part of positive reactions, the remarks made it clear that it would not be generally welcomed, and would in actuality discourage utilisation of the site, if the clients were not able to:

1) Switch the feature off and on

2) Alter the parameters of the alert to suit the particular necessities of the individual patient.

Figure 25- Medical Professional Survey – Responses to Question 11

12. In addition to the above feature to be an option in the proposed e-Toolkit, would you prefer a default feature that would apply to all diabetics or an advanced feature that can be switched on/off and adjusted for specific high-risk diabetics?

■ Default notification feature

■ Advanced notification feature
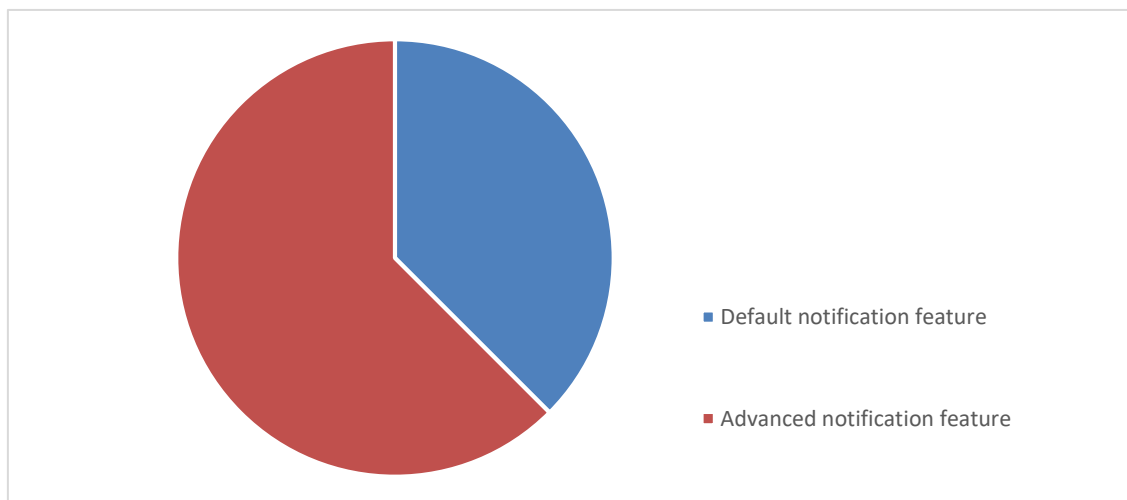
Figure 26 - Medical Professional Survey – Responses to Question 12

The response received was interesting to see the majority, 26 to 14, would prefer the advanced notification feature targeted at only high-risk diabetic patient.

13. Would you support a blog feature on the e-Toolkit which would allow patients to discuss a specific medication or problems?
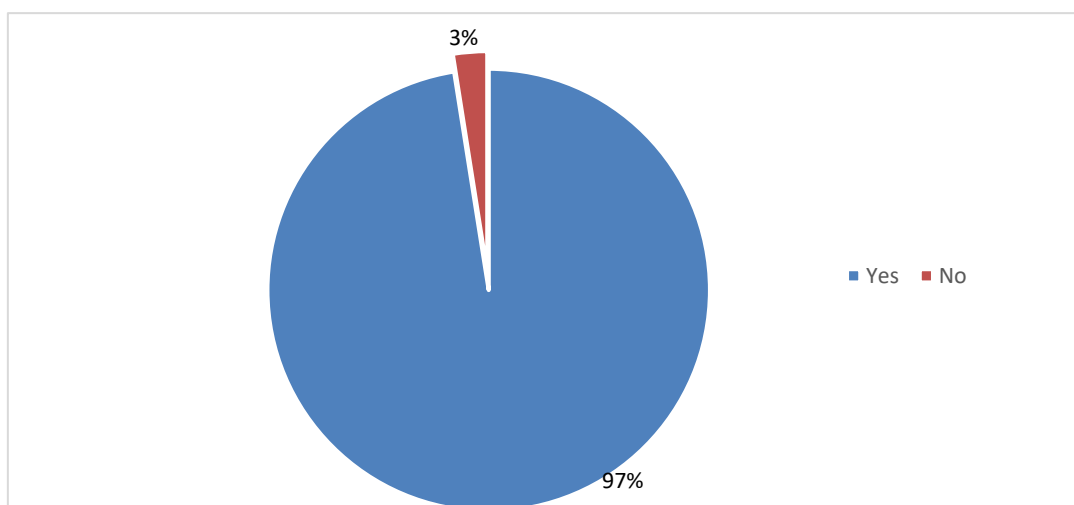
Figure 27 - Medical Professional Survey – Responses to Question 13

39/40 said yes to the Blog feature which would allow registered patients and their assigned doctors to interact between themselves, and then for the knowledge to be disseminated through our SECI model.

14. In your opinion, what features do you think that it has to be included or excluded for this e-Toolkit to be effective to the diabetic community?

The main feature that the professional health care team wanted to be included was a hub for everyone to communicate which is very beneficial for the goal of this project, what they also asked was to have that hub only available for patients who have registered on the e-Toolkit.

15. Do you have suggestions or concerns that have not been mentioned?
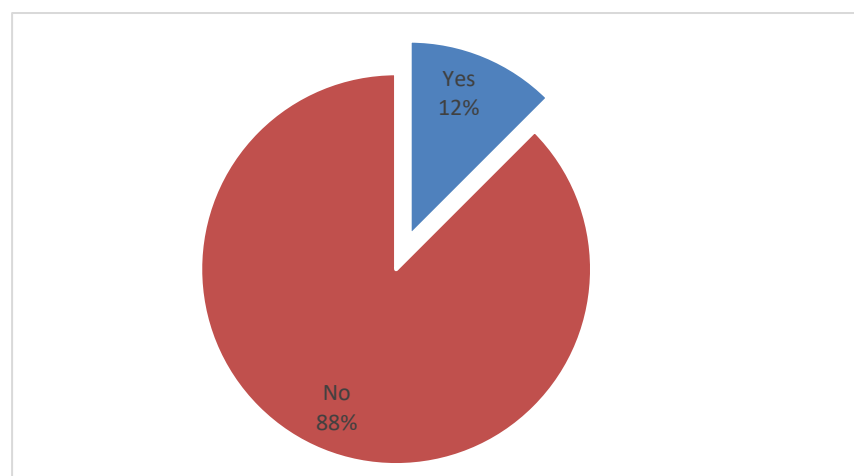
Figure 28 - Medical Professional Survey – Responses to Question 15

16. If Yes, please indicate what you feel is important or suggestions you would like for us to take into account.

The top two answers were security and easy to use as many of the older generation might find it difficult to use such an e-Toolkit. This is why the proposed e-tool kit will be accessible from a smart phone device.

3.2.7   Survey for the Diabetics

<table>
<tr><td colspan="2" align="center">

University Project

</td></tr>
<tr><td colspan="2">

PROPOSAL:

The research survey aims is to create a global e-Toolkit, not constrained by situation, income level or diabetic equipment the patient uses, to provide and support diabetes education and management. There will be numerous features including blogs where patients can talk amongst themselves. The e-Toolkit will act as a hub for patients and health care professionals to communicate as well as to provide a space for recording daily measurements such as glucose readings, food diary, insulin dosage and more relevant results and information. The patient's health care providers would then be able to view this information in real time, which would in theory enhance precise and convenient communication between the patient and the expert.

If the patient's readings are abnormal or a negative movement is detected, a feature may be incorporated within the e-Toolkit to alert the health care provider of this latest information and allow them to help the patients in such an emergency. This system could possibly help doctors and other health care providers stay up-to-date with tracking and monitoring the patient's healthcare.

HOW YOU CAN HELP:

You have been selected to participate in this short survey as you are a healthcare professional. The survey is comprised of yes or no questions, multiple-choice questions and the opportunity to add any further suggestions or leave feedback.

Participation is voluntary. You may withdraw from the study at any time and remain anonymous as **stated on t**he **consent form which m**ust be **signe**d pri**or** to **the study.**

If you have any questions, concerns or complaints about the study at any stage, you can contact: the lead researcher Fahad Darwish at *f.darwish@edu.salford.ac.uk.*

</td></tr>
<tr><td colspan="2">

Q1. How long have you been diagnosed with diabetes?

</td></tr>
<tr><td>< 3 years</td><td>Between 3-13 years</td></tr>
<tr><td>Between 13-23 years</td><td>23+ years</td></tr>
<tr><td colspan="2">

Q2. Do you have difficulty keeping a record of your day-to-day diabetic readings?

</td></tr>
<tr><td>Yes</td><td>No</td></tr>
<tr><td colspan="2">

Q3. Are you currently monitoring your diabetic readings? If so how?

</td></tr>
</table>

Written (Physically on paper i.e diary)	Stored Digitally
Generic diabetic website	Smart Phone App (General app)
Personal Blood Glucose Monitoring system	iPhone or Samsung app for your monitoring device
I do not track my readings	Other
Q4. Would you consider using an e-Toolkit like the one proposed?		
Yes	No	Unsure
Q5. If you were presented with an app for the proposed e-Toolkit, would this interest you?		
Yes	No	Unsure
Q6. In your opinion, which of the following characteristics are important to record on the e-Toolkit? (Please check all that apply)	
Weight	Blood Glucose Levels
Urine Glucose level	Changes in Mood
Side effects/symptoms	Foot related problems
Insulin	Diet
Additional / other features to include	
Q7. What is it that you currently dislike about the system /app/monitoring software that you currently using? (Please check all that apply)	
Difficult	Hard to access
Poor layout / not intuitive	Not able to track the information I want to
Readings not presented as graphs or charts	Does not track trends
Healthcare does not have access to my details	Does not have feature to share readings/other ways to share it i.e. E-mail
Security issues	
Any further suggestions? (both positive and negative)	
Q8. Sending data and alerts to your selected healthcare professionals is one of the main features of this e-Toolkit. Would you recommend this e-Toolkit to your Healthcare professional, in order to help you better self manage yourself?	
Yes	No
Q9. Would you support a blog feature on the e-Toolkit which would allow patients to discuss a specific medication or problems?	
Yes	No
Thank you for taking the time for completing the survey	

Table 3 - Diabetic Patient Survey – The questionnaire above was given to Patients around the world to fill it in.

3.2.8   Survey Results for Diabetics

1.   How long have you been diagnosed with diabetes?
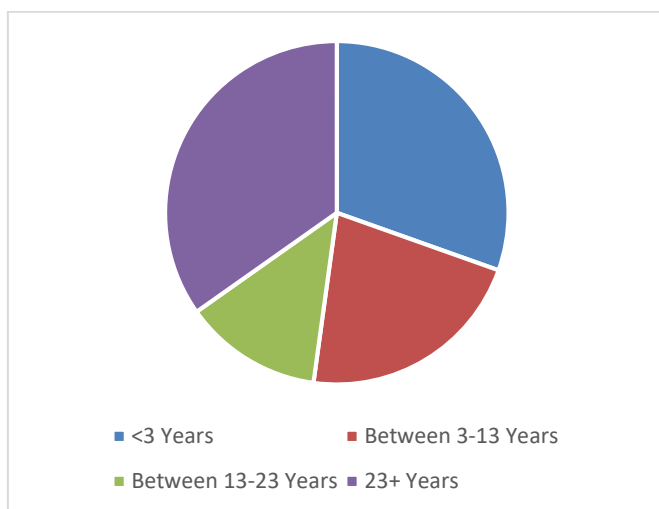
This survey got reactions from 60 diabetic patients' participants from around the globe from all ages.

Two-thirds of the patients have been diagnosed with diabetes either for less than 3 years or more than twenty-three years.

Figure 29- Diabetic Patient Survey – Responses to Question 1

2.   Do you have difficulty keeping a record of your day-to-day readings?
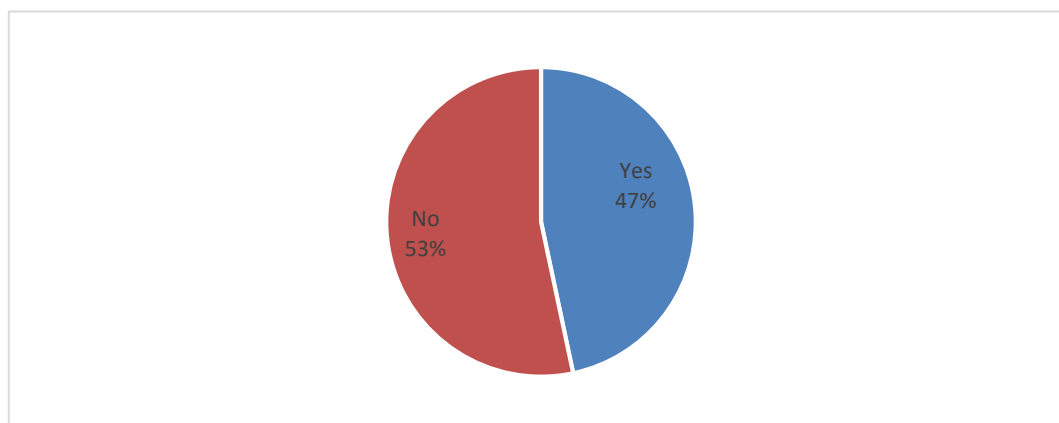
Figure 30 - Diabetic Patient Survey – Responses to Question 2

The responses to this were very surprising. It was largely expected that even those patients who perceived their condition as being well controlled would still consider the task of recording their day to day readings a challenge, especially the recently diagnosed. One important finding from the surveyors, is that there seem to be two particular classes of diabetics. There are those individuals who make the effort to understand their condition, remain in control, and manage themselves well. By contrast, other individuals may dismiss medical guidance and continue to carry forward the same eating and drinking habits that they have already acquired prior to their diagnosis and rely heavily on their medications. These are coincidentally the individuals who would not take part in such a study or use the proposed e-Toolkit.

3. Are you currently monitoring your diabetic readings? If so how?
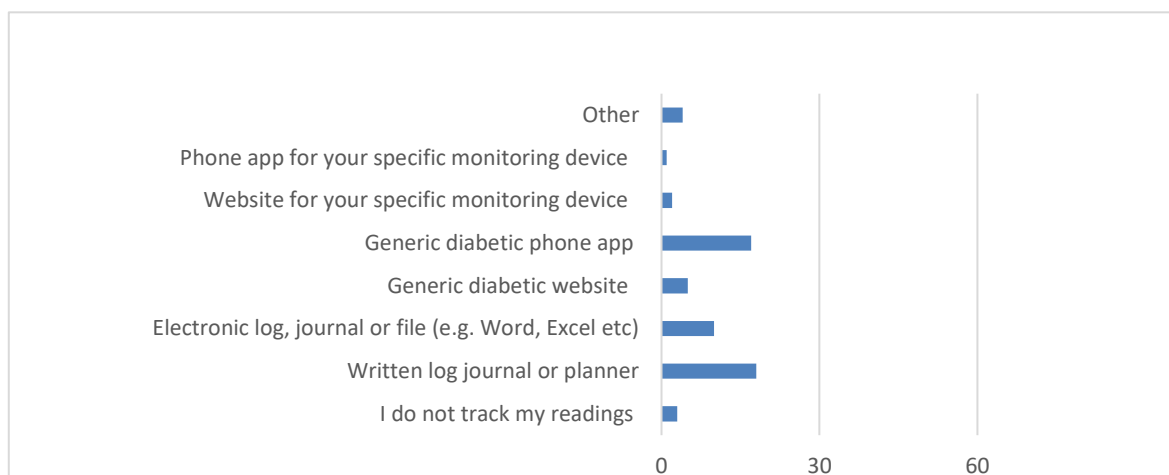
Figure 31- Diabetic Patient Survey – Responses to Question 3

Numerous patients have said that they spend a copious amount of time creating their own recording tools such as spreadsheets and journal pages. Some patients admit to having abandoned such records in favour of recording the two weeks of readings that are consequently stored in their monitoring device. A few patients said that they will just do this if they have an issue or their specialist required a more itemised history to decide how to treat them.

4. Would you consider using a e-Toolkit like one proposed?
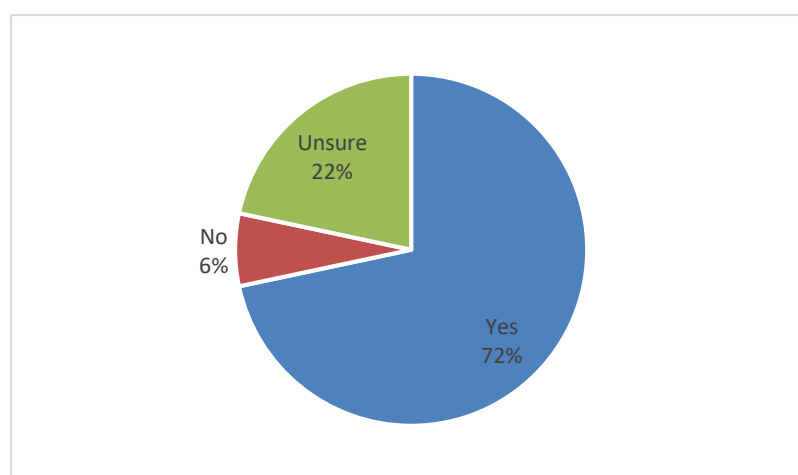
Figure 32 - Diabetic Patient Survey – Responses to Question 4

Initially it was believed that such a e-Toolkit would be welcomed mainly by diabetic groups. Seeing the medical professionals overwhelming support and encouragement was very encouraging as it was expected to see some uncertainty or doubts to e-Toolkit proposed.

The responses to Q4 demonstrated those with diabetes are significantly more reluctant, mindful and worried to use such e-Toolkit, prior to understanding it fully.

5.   If you were presented with an app for the proposed e-Toolkit, would this interest you?
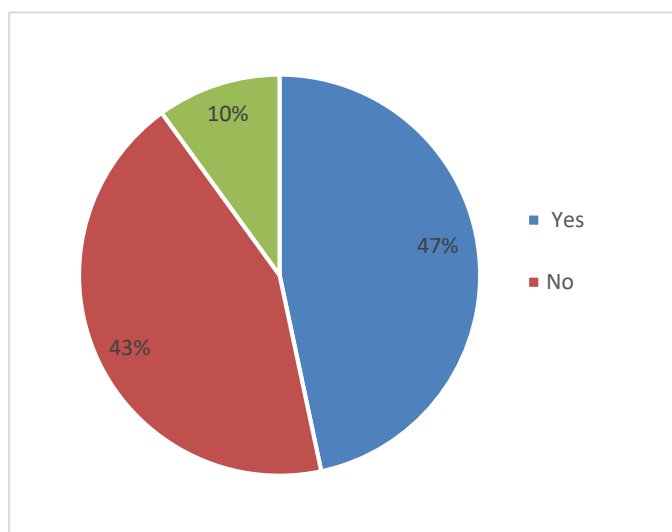
It is intriguing to see the connection in responses between this question and the one before.

One would have thought that people might be interested in mobile apps more than a PC e-Toolkit purely for the reason of mobility.

Figure 33 - Diabetic Patient Survey – Responses to Question 5

6.   In your opinion, which of the following characteristics is it important to record on the e-Toolkit? (Please check all that apply)
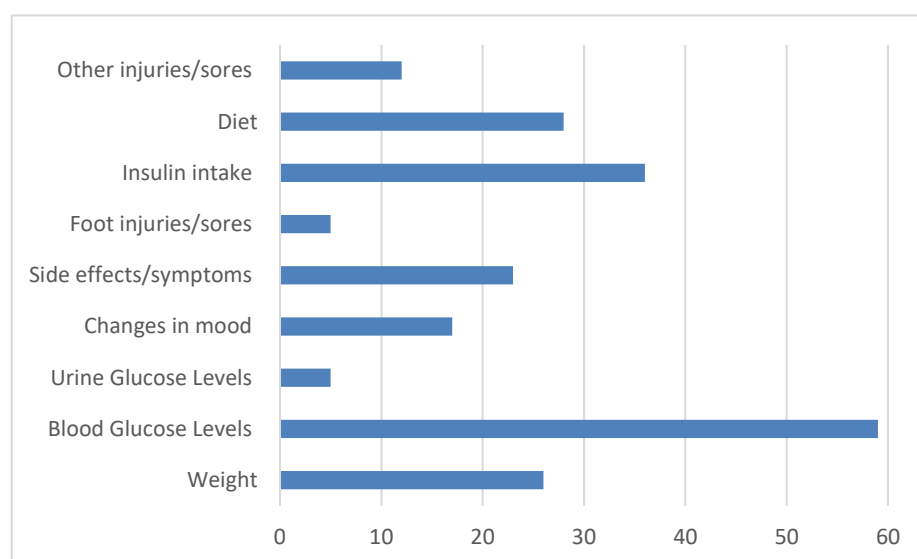
Figure 34- Diabetic Patient Survey – Responses to Question 6

Reponses from this question on both diabetic patient and medical professional surveys were used to conclude the most significant characteristic to record to the e-Toolkit. The highest rated characteristics for the diabetic patient survey are blood sugar levels, insulin intake, diet and weight.

7. **What is it that you currently dislike about the system /app/monitoring software that you currently using? (Please check all that apply)**
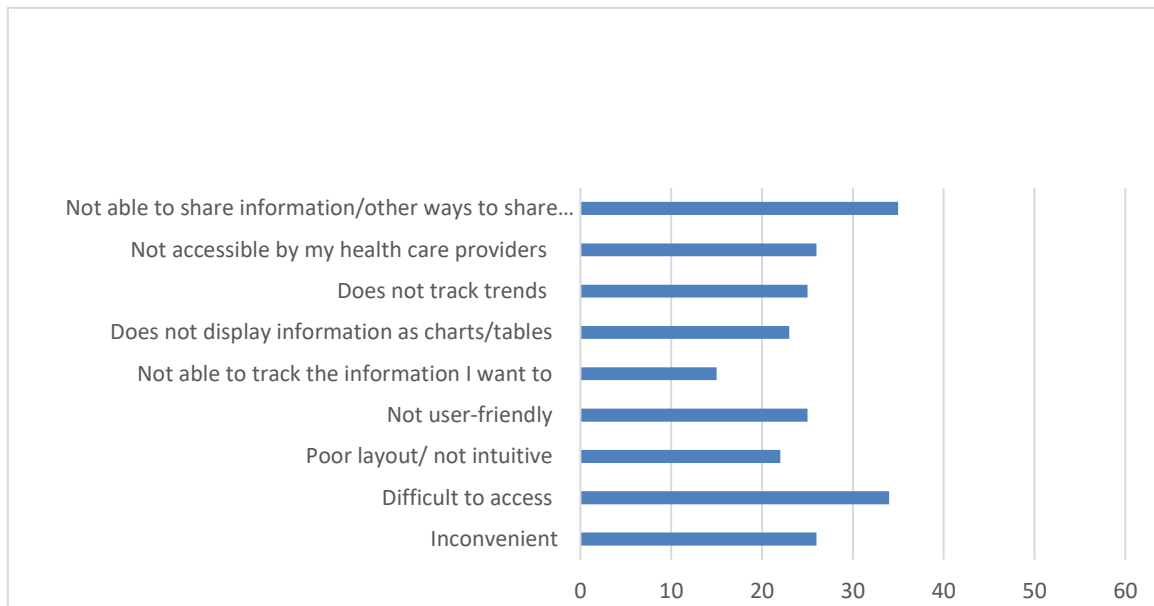
Figure 35 - Diabetic Patient Survey – Responses to Question 7

Some of the features the diabetic patients were wary about were not being able to share information, access difficulty, inconvenience and if their healthcare provider did not have access to the system.

8. **Sending data and alerts to your selected healthcare professionals is one of the main features of this e-Toolkit. Would you recommend this e-Toolkit to your Healthcare professional, in order to help you better self-manage yourself?**

Out of 60 responses, 35 diabetic patients agreed to wanting to ask their doctors to use this e-Toolkit due to the biggest feature of sending data and alerts when necessary for a more accurate and efficient way of using the data.
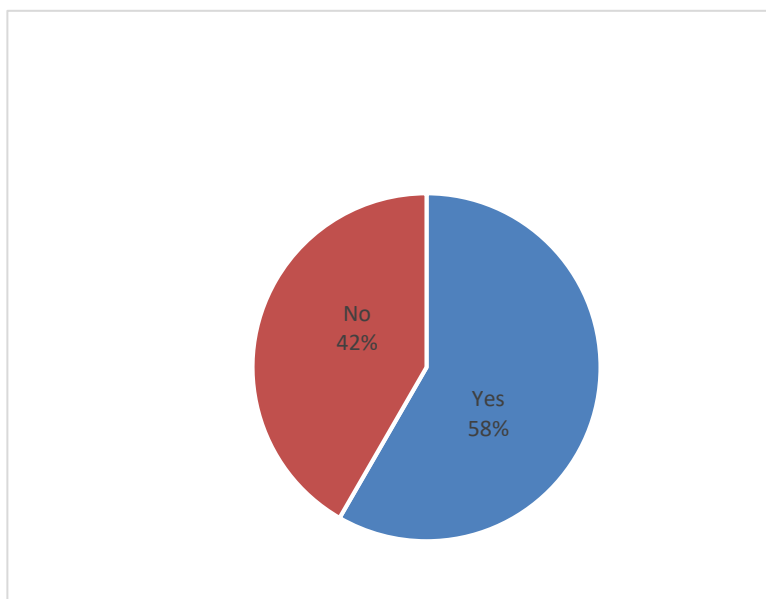
Figure 36 - Diabetic Patient Survey – Responses to Question 8

9. Would you support a blog feature on the e-Toolkit which would allow patients to discuss a specific medication or problems?
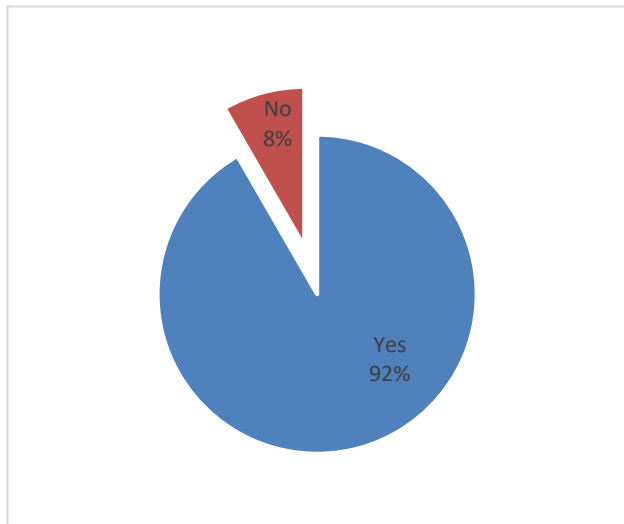
Out of 60 responses, only 5 patients disagreed with a blog feature to allow interaction between medication users to share successful stories and any difficulties.

Figure 29 - Diabetic Patient Survey – Responses to Question 9

3.2.9   Survey Comparison

Below are side-by-side comparisons of the most prioritised questions to the researcher.

1.   Would this e-Toolkit be something you would use?
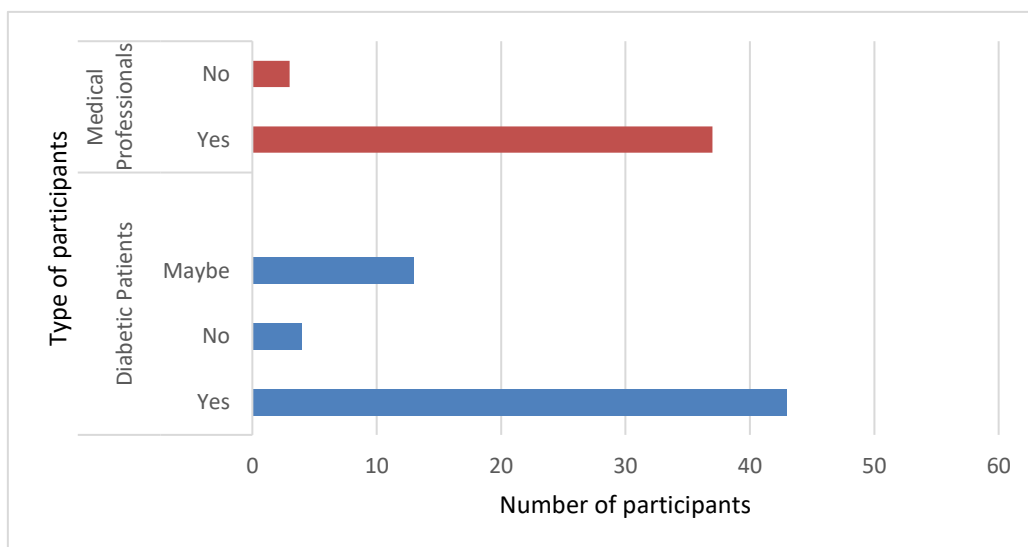
Figure 37 - Side by side comparison - Type of participants

This graph shows a two-sided comparison of the two surveys issued to medical staff and diabetics. This proves that medical professionals were more likely to approach the proposed e-Toolkit with extraordinary eagerness as opposed to patients who were more apprehensive.

2.   Please indicate the readings that the e-Toolkit should record?
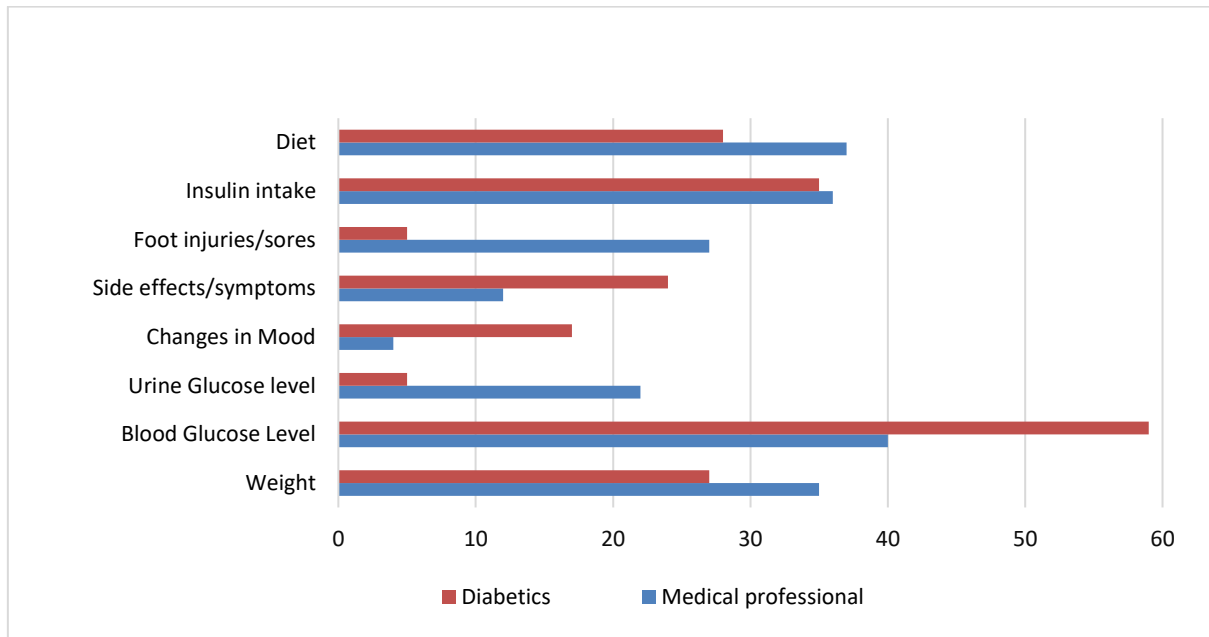
Figure 38 – Side by side comparison – Readings that the e-Toolkit should record.

The comparison demonstrates that medical professionals and diabetics are both happy with which data ought to be recorded on the proposed e-Toolkit. The top health factors to be record were (Glucose readings, Insulin readings, weight, symptoms if there is any and carbohydrate intake).

3.   Would you recommend a feature where the e-Toolkit sends notifications to your healthcare professionals recording up normal readings?

The main feedback gathered from the side by side comparison was that both medical professionals and patients were in agreement to be notified as long as the doctors are able to set the level of the up normal reading for each patient as this will vary from one person to another.
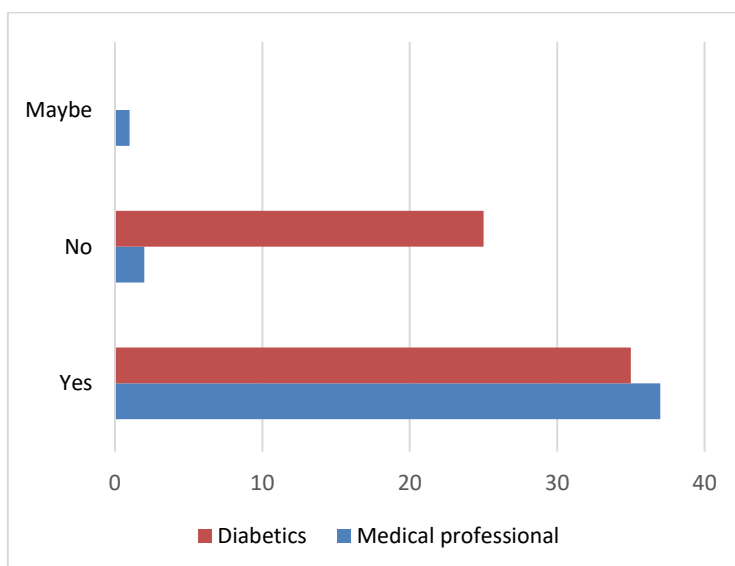
Figure 39 – Side by side comparison – Notifications feature

3.3   Summary

When carrying out the aforementioned research, some aspects such as the impact on the people themselves throughout their lives, physically and emotionally may be have been missed. Whilst collecting these findings many ideas were discovered from the diabetic community. Throughout the depth of the numerous subjects in this research, the researchers were reminded constantly that diabetes is a very serious disease affecting millions of people's wellbeing and everyday lives around the world. However, creating an e-Toolkit to try support a huge community is neither easy nor simple.

When conducting a research project like this it can be easy to get carried away with the theoretical aspect and neglect how it may serve people in reality. It is clear that diabetes is a condition with global implications affecting millions of people worldwide. Formulating a tool which can enrich the lives of such people is no easy feat. The most striking outcome from this research was the discovery of the communication gap between healthcare professionals and diabetic patents. This was more exaggerated in North America when compared to the UK. Many questions have arisen from this revelation, many beyond the scope of this thesis. As previously touched upon, it appears that the diabetic patients in this study can be divided into two general categories, as shown in the following table.

GROUP 1	GROUP 2
This group takes account of the diagnosis and management of their disease by paying attention to their symptoms, testing frequently and maintaining records. This makes them feel they can control their health and manage their diet and exercise. As a group they particularly expressed some feelings of doubt and distrust of their healthcare professionals including doctors. They have stated that the doctors have a lack of communication with the patients and are not sympathetic towards their treatment. A recurring theme is present in which patients repeat that healthcare professionals were not helpful and did not impact the diabetic patient's health in tackling the illness. Instead they felt like the doctors made the situation worse. Patients from group 1 were fearful of civil liberties being revoked, including driving, as the doctor can decide this if they wish. Therefore, this group did not agree with the unlimited and unrestricted access of the e-Toolkit with their records and information.	This group seems to be the people that medical professionals are most concerned about They are most likely to be the main people to benefit from this e-Toolkit however are unlikely to use it. Group 2 are the patients that are most neglectful of their disease. These are patients who rely heavily on their medications or insulin intake as they do not want to change their diet or daily lifestyle. Patients in this group have an increased likelihood of hospitalisation or to experience major complications, such as limb amputation, for not looking after themselves.

Table 4 – Summary of results - Two groups of Diabetic community

CHAPTER FOUR: PROPOSED FRAMEWORK

4.1   Introduction

In the previous chapter, the research problem was identified, and a survey was conducted to see what the e-Toolkit should include in terms of diabetes readings, what the patients and the medical health care thought of the idea proposed and what would be the benefits to the community. This chapter will propose a framework to overcome some of the problems identified in the previous chapters. The framework integrates KD and KM to help patients in managing themselves. In this chapter, the two parts of the thesis are discussed as two layers. The first layer focuses on the improvement of medical data collection and communication methods that can be applied to different diseases, as well as the creation of an e-Toolkit. The framework aims to be applied to the field of diabetes mellitus in the United Kingdom. The framework seeks to improve self-management techniques for patients and to improve the effectiveness of the medicines provided by the healthcare professionals, related to diabetes mellitus. Layer two of the framework is linked to KD and Data Mining in specific. This is where the reviews of the medicines provided for diabetic patients which was obtained from layer one is data mined. In addition, layer two acts as the KM layer where the SECI model is used to capture these reviews and disseminate this knowledge on the e-Toolkit in the form of guidelines or documents. This chapter will also introduce the proposed SECI model where knowledge management and knowledge creation processes are taking place, which can assist in sharing the knowledge among individuals. The SECI model is considered to be the most documented knowledge model which converts and creates knowledge between tacit and explicit forms (Ikujiro Nonaka, 2008).

4.2   Proposed solution & Framework

The first goal of this research study is to develop an e-Toolkit that is globally available and that will fill the communication gap between medical and healthcare professionals. The second goal of this research is merging KM and KD within the e-Toolkit to support diabetes education and self-management in the UK. Knowledge generated from the reviews will then be disseminated using the developed SECI model. This knowledge can also be generated from reviews or the discussion of specific topics between patients and the medical team within the blog section of the e-Toolkit through the conversion of tacit knowledge to explicit knowledge as explained in Chapter 2. The e-Toolkit needs to be user friendly and accessible to medical teams and patients enabling them to record individual risk factor data that is important in monitoring an individual patient's condition.

The patient will have an e-Toolkit which they can use to consistently record readings, and which can be accessed from home, work or from any other location desired using a mobile phone, laptop or any other device with internet connectivity and a browser. Patients will no longer have to experience the inconvenience of carrying a journal and writing notes.

The e-Toolkit will also be accessed by the patient's doctor allowing vital patient/doctor interaction. Doctors and patients are able to interact using the direct messaging system. As the system requires a log in, procedures will be implemented to allow regular monitoring with the aim of the e-Toolkit becoming integrated to the daily routine of the patient. Additional features can potentially provide displays of graphs and charts, notifications to high-risk patients as well as enabling data from lab results to be recorded and shared (Fahad Darwish, 2019).

The e-Toolkit aims to be accessible to all patients and is not limited by expensive apparatus or subscriptions. This will be of significant help to the diabetic community as it has been reported that the cost of treating diabetes and its complications amounts to 93bn Euros in Europe alone (OECD, 2012). Therefore, having an e-Toolkit such as the one proposed will help to manage the resources necessary to combat the disease.

In summary, the limitation of the current methods and technologies discussed in chapter three can be overcome by the proposed e-Toolkit as the proposed framework will solve the everlasting broken relationship between doctors and patients. Therefore this will allow doctors and patient to resolves issues related to diabetes more effectively and efficiently. This will have an effect on the doctors having a better understanding of each patient and have a clear thought to aid in reducing the effects of diabetes mellitus by advising solutions best for individuals. This can benefit the reduction of the rate of diabetic patients in the UK and possibly around the world. Having this e-toolkit available and free of charge to all patients with internet access and can become an important tool for diabetic patients and healthcare professionals. It aims to reduce the cost of the budget currently spent on diabetes by preventing patient complications via close monitoring.

4.3   The proposed solution and framework Components

As mentioned above the proposed solution and framework is described in two layers, those layers are illustrated in figure 40.

Layer 1: The Improvement of

Medical Data and Communication

Methods

Layer 2: SECI Model (Creation and

Dissemination of Knowledge)
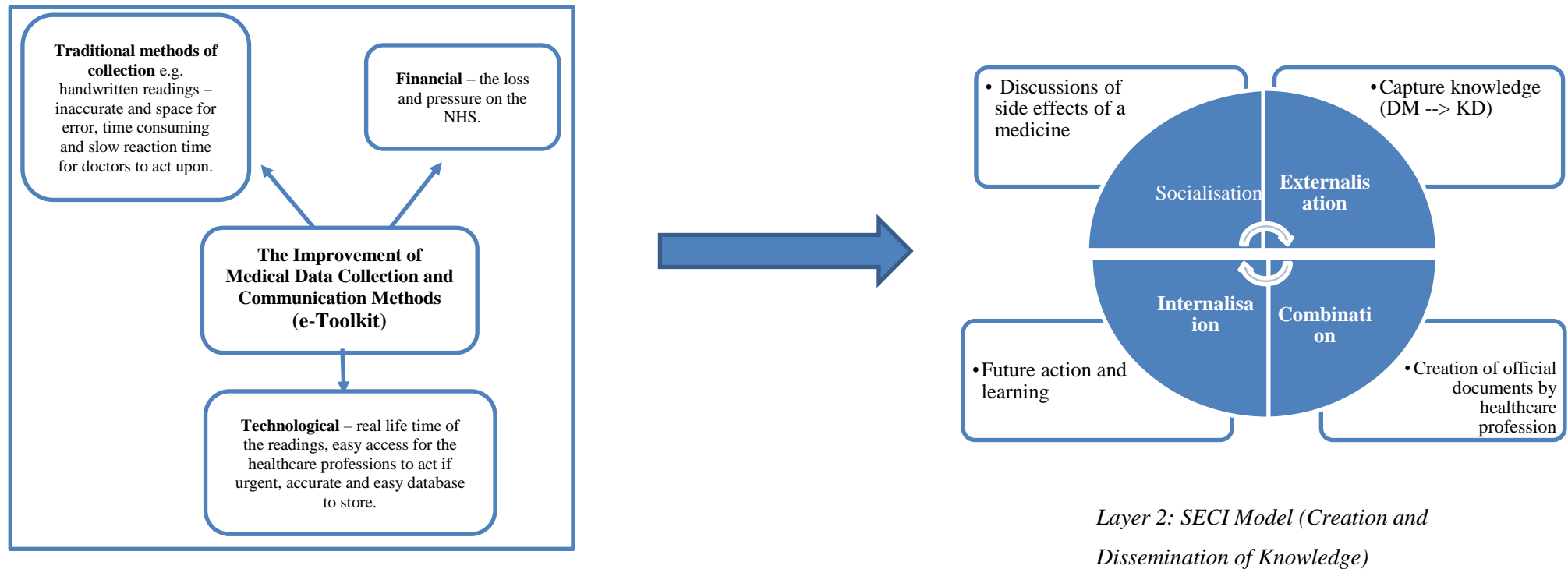
Figure 40 – The Proposed Solution

4.3.1   The First Layer: Method and Development of a Diabetics Self-management System

In this layer, there are two categories: the improvement of medical data collection methods and the communication with the healthcare team. The framework is applied to improve the interaction between doctors and patients and to improve upon the current methods of capturing data. It has been reported that daily self-management for diabetic in the UK is a concern (Shayban, 2017). This layer aims to become innovative for patients to discuss topics amongst themselves. It will also capture readings and log them into a database for the healthcare profession to analyse and make a choice from a selection of medical responses.

The Diabetes Management e-Toolkit System of the Medical Health Care Home Page is illustrated on figure 41. A more detailed explanation of the e-Toolkit is explained in chapter 6.
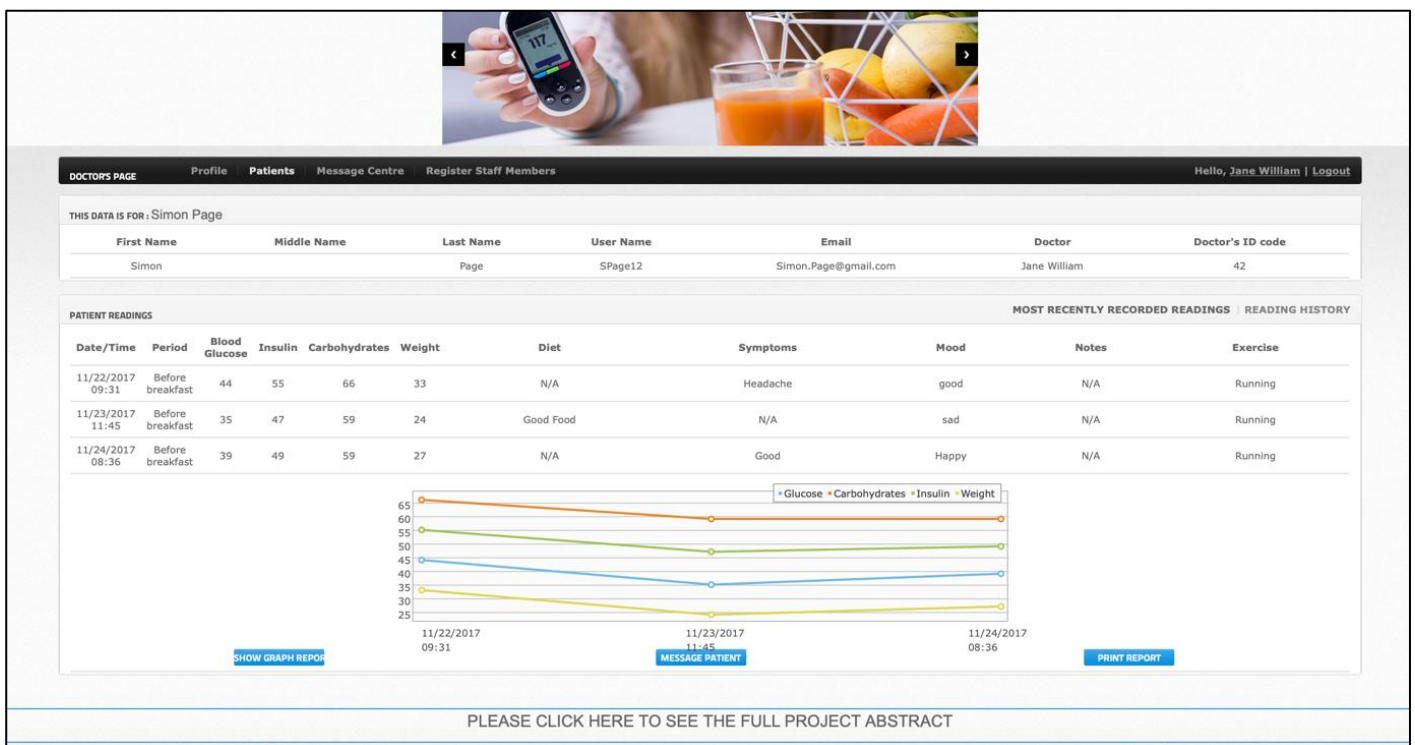
Figure 41 – The Diabetes Management System e-Toolkit

4.3.2 The Second Layer: The Discovery and Dissemination of Knowledge via the developed SECI model

The aim of this layer is to mine the data acquired from the reviews that were collected from the first layer via the blog within the e-Toolkit and to then elicit new knowledge that can improve the medical field. The DM method employed to achieve this is 'Sentiment Analysis' which is appropriate to the structure of our reviews which contains ratings of 1 for good review and 0 for a bad review (Lin Zhanga, 2014). The knowledge discovery approach is shown on figure 42.

In order to develop the recommendation system within the SECI proposed model, different machine learning algorithms were reviewed to compare their performance on different datasets (as discussed in Chapter 5). These techniques have applications in text classification, sentiment analysis, pattern recognition and data mining problems. In addition, layer two helps in the conversion of tacit and explicit knowledge between patients and the health care team to provide useful knowledge that can help patients to manage and control their diseases. As explained in Chapter 2, the developed knowledge creation model (SECI) shown in figure 42 is applied in this layer to provide effective self-management and education for sufferers of diabetes mellitus, to facilitate effective sharing of guidelines and best practices to diabetics. Strategies, discussions, guidelines and best practice are useful recommendations to be shared between patients and healthcare professionals (Almuayqil, 2017). These practices should provide solutions to help predict the right treatment or side effects of certain medicines depending on patient location, age etc. Details about the proposed SECI model are discussed in section 4.4 below.
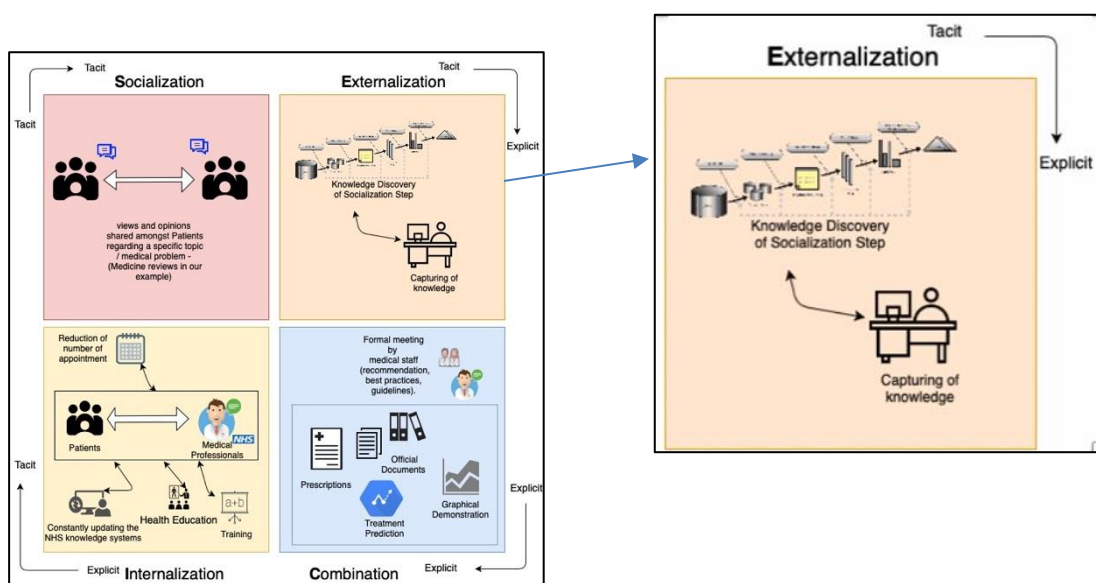
Figure 42 – The proposed knowledge discovery part of the SECI model

4.4   The proposed SECI model

The proposed SECI model is illustrated in figure 43, the process and the merging stages of KD & KM happen within one model. Knowledge will be discovered and managed through the SECI four modes starting from Socialisation where views and opinions are expressed within the e-Toolkit blog section then to the Externalisation mode where knowledge is captured through data mining techniques. Followed by Combination where knowledge is discovered and passed to the professional personnel to be verified and to create official documents. Finally, in the Internalisation mode, documents that were generated from the Combination mode is published into the e-Toolkit as illustrated in figure 44 this is to help the diabetic and health care professionals.
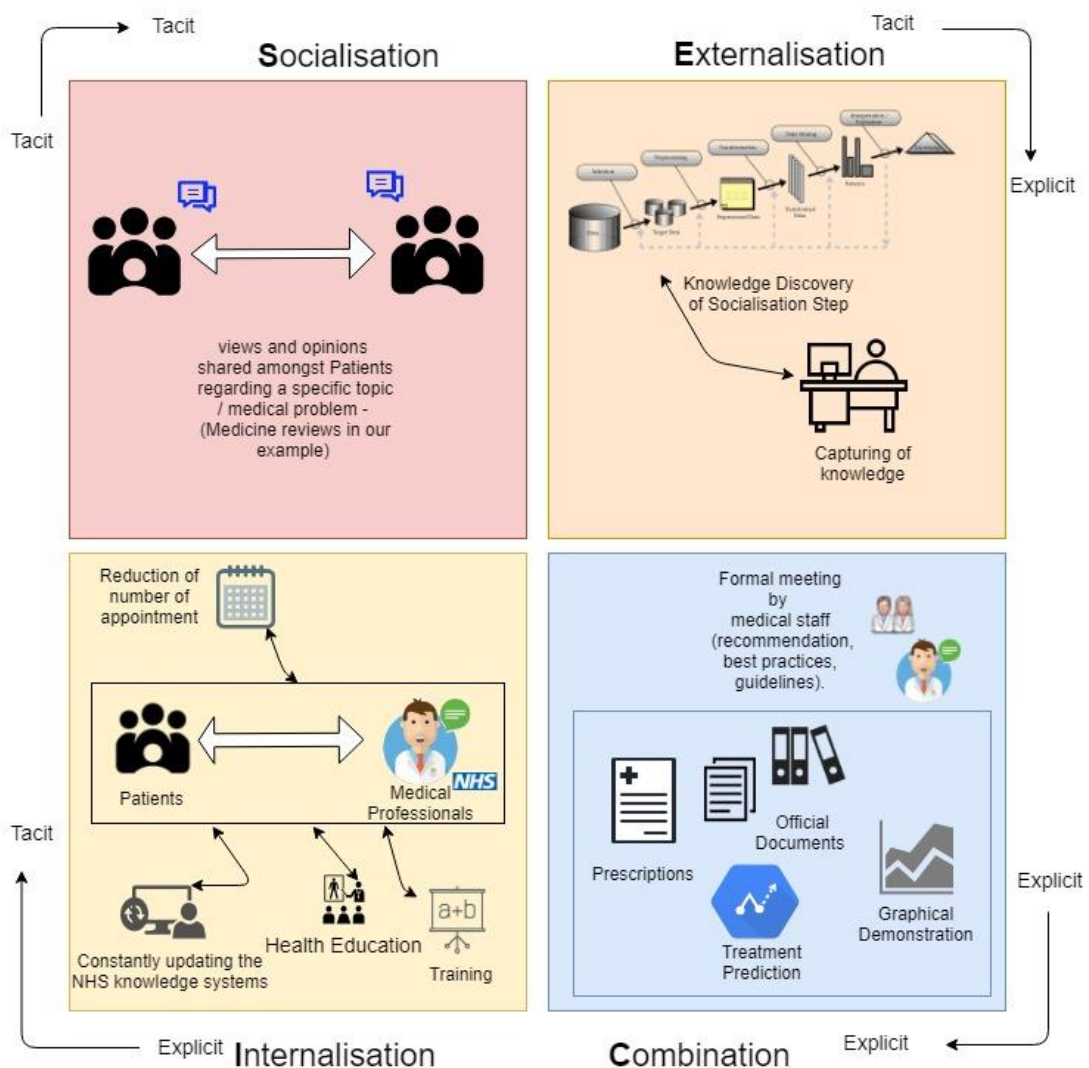
Figure 43 – Illustration of the Proposed SECI Model process for Knowledge Creation and Dissemination

Each of the modes is explained in detail below:

Socialisation- Tacit to Tacit. As stated by Nonaka "Knowledge is passed on through practice, guidance, imitation, and observation" (Nonaka, 1994). This is where the patient's community and medical staff will have meetings e-Toolkit features Blogs, as it was indicted by (M.A., 2007). This mode in the proposed SECI model can overcome the complication of matters such as medication side effects. Diabetic patients can socialise amongst themselves within the blog provided in the portal. Patients whom might be taking the same medication can discuss any other issues that they might share.

Externalisation- As stated by Nonaka – "Tacit knowledge is codified into documents, manuals, etc. so that it can spread more easily through the organization" (Nonaka, 1994). The externalisation mode in the proposed SECI model will be the mode where knowledge collected from the blog will then be converted from tacit to explicit knowledge. This is where the knowledge discovery part takes place using sentiment analysis and other data mining techniques that are explained in chapter 5.

Combination- Explicit to explicit. As stated by Nonaka "This is the simplest form. Codified knowledge sources (e.g. documents) are combined to create new knowledge" (Nonaka, 1994). The combination stage is where formal meetings by medical professionals to discuss findings will take place which will then results in (recommendation, best practices, guidelines) all those then put into an official document to help the community/NHS etc.

Those official documents will be published on the portal. As shown on Appendix K & Figure 44. As stated by (Natek, 2016) "One of the effective web technologies for combination is a document system". As indicated by (Natek, 2016) a document system helps members to upload and download important documents of interest. It facilitates access to information. As it was mentioned by (Nonaka, 1994) that "an open source to information is an important tool for combination".
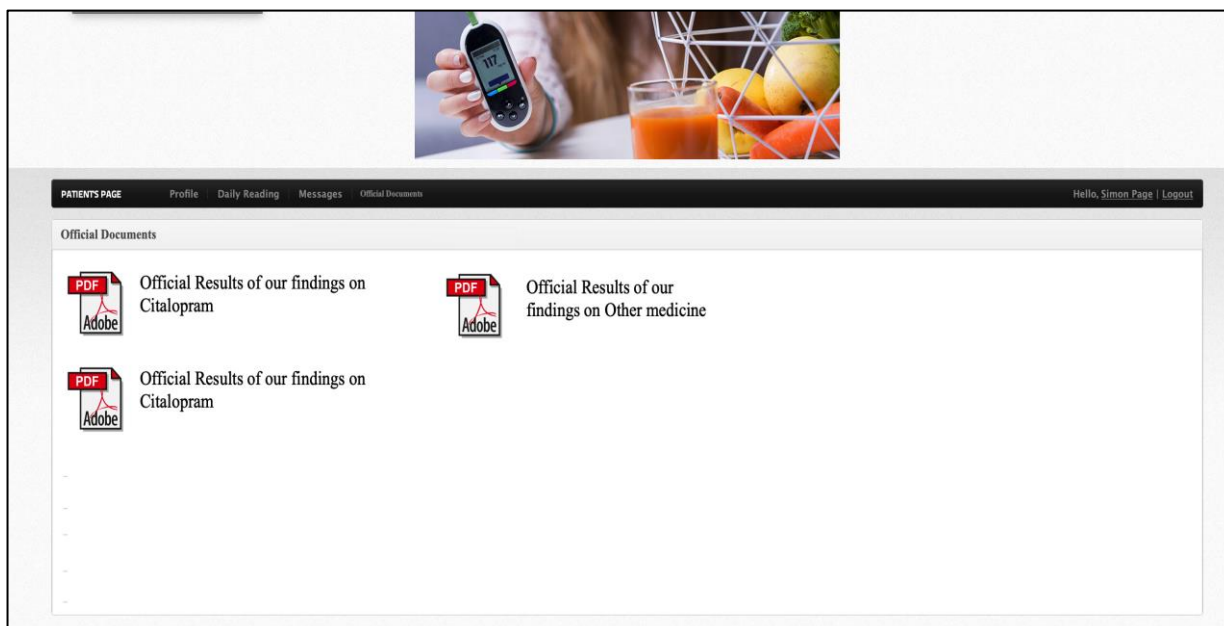
Figure 44 – Official Document generated from our Knowledge

Internalisation – As indicated by (allkm.com, 2018) "internalisation of explicit knowledge is created using tacit knowledge and shared across the organisation. When this tacit knowledge is read or practiced by individuals it broadens the learning spiral of knowledge creation". This is where tacit knowledge is converted to explicit knowledge. Everything concluded from the previous stages is presented to the patients and medical community as explicit knowledge. The e-Toolkit will have the features to share this knowledge including user training through instructional videos.

The knowledge discovered could contribute in finding side effects of treatments. The e-Toolkit could have the option to organise that information and potentially update NHS knowledge systems such as 'nice.org.uk'. The "National Institute for Health and Care Excellence" system which as described by (NICE.org.uk, 2018) a system that "delivers access to particular and authoritative evidence in health, social care and public health. with their latest findings".

4.5   Chapter Summary

In this chapter, a solution has been discussed and proposed that integrates KD and KM. An e-Toolkit is also proposed to bridge the gap in the NHS and help patients and healthcare professionals better communicate with each other within one platform. The two layers of the solution and framework are also described. The proposed e-Toolkit system will not only fill the issues mentioned in chapter 2 but it will also provide a tailored KM approach to help diabetics to self-manage their health condition, likewise the e-Toolkit will help in cutting down GP appointments which will results in budgeting the NHS budget.

There are several contributions associated with our integrated framework. These are listed below:

1. The Framework is the first to integrate DM and KM in the healthcare domain in the UK.
2. The use of DM to extract useful knowledge in relation to medication side effects and other non-health related diseases.
3. The dissemination of useful knowledge using the proposed SECI model first of its kind in the UK.

Having discussed the solution model and the associated techniques, the experimental implementation is discussed in detail in the next chapter.

CHAPTER FIVE: EMPIRICAL IMPLMENTATION OF EACH OF THE PROPOSED LAYERS

5.1   Introduction

This chapter covers the implementation of the framework described in Chapter 4, in particular layer 2 of the framework "knowledge discovery part" of the SECI model. In order to implement and validate layer two of the framework, three datasets are used. This is because time was a limiting factor and sufficient data was not collected from the e-toolkit. Therefore, the researcher decided to utilise datasets that are available online for the purpose of validating this framework. These datasets are explained in detail below. This chapter will also discuss the evaluation process, the datasets used, the knowledge extraction implementation for KD as well as comparing the available algorithms and analysis of the results. Accordingly, to achieve these tasks, the following software and programming languages tools were used:

- Anaconda (Spyder): Python IDE software written in Python and it is used for data mining analysts (anaconda.org, 2019).
- Python
- Natural language toolkit (NLTK), scikit-learn and regex

5.2   Framework Structure

As mentioned in the previous chapter the proposed framework is designed in two layers (see figure 45). The first layer focuses on filling the gap in helping patients and the medical care team communicate freely by creating a web portal as discussed from survey findings in Chapter 3. This layer also acts as a hub to help diabetics discuss their concerns or complications that they are currently facing within the blog section of the web portal.
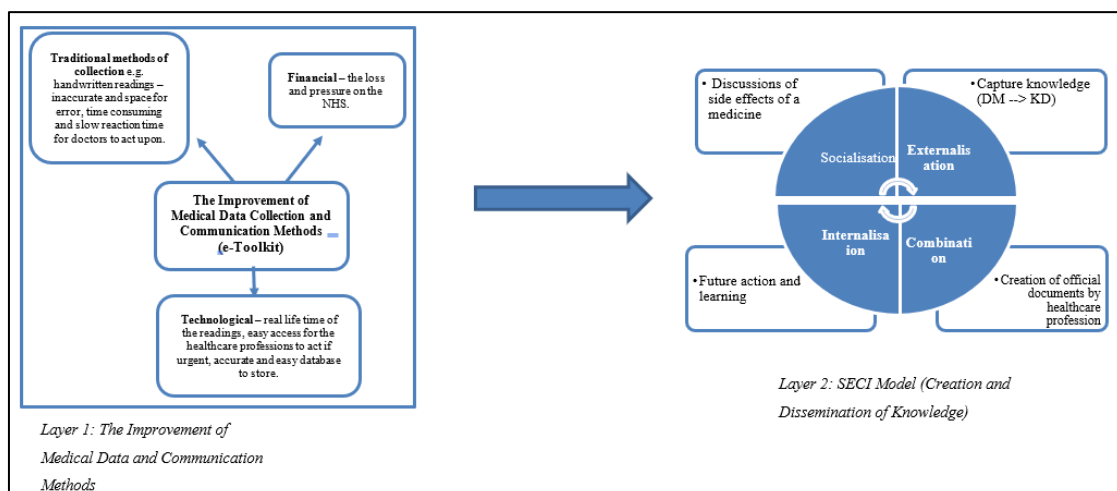
Figure 45 – The Layers of the proposed Framework as discussed on chapter 4

The second layer is related to KD and DM in particular. This is where the developed SECI model (illustrated in figure 45) is used and where the concerns relating to a specific medicine / topic are discussed in the blog provided to patients within the e-toolkit. This will result in the sharing of useful diabetes self-management tips and provision of education, guidelines and best practices to the patients as a group.

5.2.1   Dataset Search Strategy, explanation and preparation

To achieve the goals of the second layer, a dataset is required. The dataset has reviews and ratings in order discover new knowledge to validate layer two of the framework, which is knowledge discovery. The process of discovering new knowledge through the SECI model is shown in figure 43 in the previous chapter.

Sentiment Analysis, in the context of reviews, as described by (Haddi, 2015) "Is the process of exploring product reviews on the internet to determine the overall opinion or emotions about a product. Reviews that represent the so-called user-generated content is currently attracting increased attention". It is a useful resource for marketing teams, sociologists, psychologists and others who might be concerned with opinions, views, public mood and general or personal attitudes. In our case this would be the medical teams who would aim to establish the side effects of medication through the use of the proposed framework, which converts knowledge which has been extracted from blog reviews into useful knowledge. Therefore, the knowledge conversion from tacit to explicit is achieved and by using sentiment analysis helps to convert reviews to official documents to support the diabetic community (Tang, 2009; Haddi, 2015).

5.2.1.1   Dataset

Three opensource datasets were used to verify the proposed model, including two medical datasets and one commercial dataset. The datasets are split into two columns, the first column is reviews and the second column is the ratings, the ratings are "0" for a negative review and "1" for a positive review.

An explanation of each dataset is shown below in the form of case studies to help in the validation of the framework:

5.2.1.1.1   Case Study 1 – Reviews on Citalopram Drug

Citalopram, which has the common trade name of Celexa in the US. As described by (fda.gov, 2019) "is an FDA approved antidepressant medication that is often used to treat other mental illnesses, such as anxiety disorders. It may also be found under the brand name *Cipramil* in other parts of the world" (fda.gov, 2019).

Figure 46 – Citalopram Drug – Source: (fda.gov, 2019)

Dataset Source: " https://www.webmd.com/drugs/drugreview-1701-citalopram+oral.aspx?drugid=1701&drugname=citalopram+oral&pageIndex=47&sortby=3&conditionFilter=-1 "

296 Reviews on the above medicine with ratings, 1 = good review / 0 = bad review

5.2.1.1.2   Sample of Citalopram Drug Dataset

	A	B	C
1	comment	rating	
2	Can t stay awake always tired and doesn t seem to work for me I ve had up to 40mg need to try a different tablet	0	
3	My doctor told me 10% of people feel an improvement in the first drug they take.	0	
4	No improvement after 3 months @ 40 mg. TERRIBLE dry mouth especially at night. I waken throughout night cuz my mouth feels like it s full	0	
5	I was on this for two years for depression/anxiety. At first it worked but I had to keep increasing my dose and finally it just quit working comple	0	
6	Depression has greatly increased	0	
7	I ve been told by doctors through the years that I ve been depressed all my life and don t know it. Was prescribed Wellbutrin but it seems I	0	
8	Too many side effects to warrant its use in my case was slightly effective but my mild depression didnt warrant use of what I see as quite a pow	0	
9	makes me very tired and nauseated	0	
10	I was given this generic actually two generics (both new to me) were mixed in one prescription container. I had been doing fine on another citalo	0	
11	They prescribed this drug because of the depression I experienced when my testosterone went low. I should have had the testosterone issue tre	0	
12	Against my wishes my pcp prescribed this to me because my mother wanted a quick fix instead of some therapy (@ age 07). It gave me worse d	0	
13	This medication doesn t work at all.	0	
14	After taking this medication for only 7 days my brother-in-law commited suicide. He was a kind gentle man who never talked of suicide was de	0	
15	i couldnt get past the first day of it going to try someting else	0	
16	i took this medication and within about 0 or 3 days i was gaining bounds of weight and shaking so violently i could not write. by the 4 week mark	0	
17	I was taking lexapro but stopped due to expense and loss of insurance.I have been taking Celxa for 1 months.I feel more depressed irritable and	0	
18	Used for 30 days didn t see any results. Made me extremely tired during the day. Would have to take a nap during the day most days.	0	
19	I do not have very much energy and have sleep problems	0	
20	Prescribed for depression and anxiety. Seemed to help with anxiety not with depression. Sexual side effects. Horrible withdrawal symptoms wh	0	
21	I just started on this medication	0	
22	My doctor put me on 00ml of this with phentermine so I can lose the 00 pounds I have gained. After reading how these two drugs should not b	0	
23	very weak	0	
24	My P.C.P. the following day sent me to E.R. at one of our illustrious hospitals here at the Medical Center Houston. This was the same hospital th	0	
25	Very bad results. First two months was great. Chronic pain subsided. Then went into deepening depression.	0	

Figure 47 – Sample of Citalopram Drug Dataset

5.2.1.1.3   Case Study 2 – Reviews on Amazon Echo Show Alexa-enabled Bluetooth Speaker with 7" Screen

The Echo Show has both a display and a front-facing camera. It can be used to video call people. Specifically, it allows the user to video call people with either an Echo Show or the Alexa app on their phone (amazon.co.uk, 2018). As mentioned by Amazon website "The Echo users could also watch video flash briefings, access the Amazon Video, see music lyrics, security camera images, photos, weather forecasts, to-do and shopping lists, browse and listen to Audible audiobooks, and much more" (amazon.co.uk, 2018).

<u>*Note:*</u> Amazon Echo Show had been used to validate the layer two of the Framework which is KD. The dataset from Amazon Echo reviews shows evidence that the framework can be applied to monetary products and is not limited to medicinal purposes.

Figure 48 – Amazon Echo – Source: (amazon.co.uk, 2018)

Dataset Source: kaggle.com – " https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products#Datafiniti_Amazon_Consumer_Reviews_of_Amazon_Products.csv "

846 Reviews on the above Device with ratings, 1 = good review / 0 = bad review

5.2.1.1.4   Sample of Amazon Echo Dataset

Figure 49 – Sample of Amazon Echo Dataset

5.2.1.1.5   Case Study 3 – Reviews on Gilenya Drug

The primary use of this drug is to treat specific types of multiple sclerosis. The principle action of the drug as described by (webmd.com, 2018) is "to prevent the immune system cells (lymphocytes) from attacking the nerves in the brain and spinal cord. As a result it may help decrease the number of episodes and may prevent or delay disability" (webmd.com, 2018).

Figure 50 - Gilenya Drug – Source: (webmd.com, 2018)

Dataset Source: " https://www.webmd.com/drugs/drugreview-154676-Gilenya+oral.aspx?drugid=154676&drugname=Gilenya+oral&pageIndex=17&sortby=3&conditionFilter=-1 "

73 Reviews on the above medicine with ratings, 1 = good review / 0 = bad review

5.2.1.1.6   Sample of Gilenya Dataset

	A	B	C	D	E	F	G	H
11	i took Gilenya for around 00 months. During that time my walking improved dramatically. I had been falling around	1						
12	I have been taking the pill since nov. 0000 .. I m walking much better (though I still can t wear heels) all MS sympto	1						
13	been taking this drug since 0009 and haven t had a replase since.	1						
14	have been on Gilenya for 0 years and I felt it was working well until I ended up with cancer in my spine and lung. Not	1						
15	uper easy pill! However I developed shortness of breath when on drug. Started about one to two months afternoon drug. Dr did not relate shortness of breath to expiration/ asthma type problems until after a
16	izziness nausea hot flashes tiredness	1						
17	Gilenya has made a big positive difference in my life! Diagnosed with MS in 0999 I was on the injectible drug Avone	1						
18	My husband took Gilenya from November 0000 until December 06 0000. He began having severe depression anger i	0						
19	I was on Avonex for 9 years I begged my Dr. to take me off of the injections he refused. I switched to a new Dr. he t	1						
20	My neuro called and wanted me to come in to see him ASAP he had my blood work back. I ve been on Gilenya since	1						
21	I have been on Gilenya since 6/0000. It had been working wonders for me until 0/01. This is when my immune syster	1						
22	I have been on G for 00 months. I have been on Avonex Rebif and Copaxone. I was dx MS 9.1 years ago... I am 16. Us	1						
23	I have been taking Gilenya about 9 months now. The first 6 months went ok after blood tests for my liver the dr. too	0						
24	I am very happy to be free on needles. Gilenya has made treating my MS a lot easier. I set up to take my daily dose v	1						
25	o far the only side effect I have had (besides the expected drop in lymphosites...if that s the right word) is that I have	1						
26	Really bad side effects. Confusion perceived heart attack depression everything that could have happened happen	0						
27	My daughter has MS was diagnosed 1 years ago. She stopped all injectable medications due to side effects. With t	1						
28	have been on gilenya for 0 0/0 years and my overall function has significantly improved since taking this drug - and w	1						
29	I have not started this medication yet.I received it on Fri.I wanted to do a little more research on it before I started it	0						
30	I started Gilenya in January. It took a long time to figure out what to go on since I was on Tysabri..taking it every two	1						
31	I just turned 10 and after discovering my MS over 01 years ago am doing better than ever. All I wanted after my dia	1						
32	After 00 years of injection site lumps and pain this is pure heaven. I actually cried when my dr told me that a pill for	1						
33	Seems to be working can t beat a once a day pill.	1						
34	Started taking thedrug in October of last year (0000) no side effects. Walking is much better very noticable by other	1						
35	I been on it for over a year and have had no side effects! I also had an increased libido great side effect. I am feeling	1						
36	ave been on medication for a couple weeks. No major side effects and I am glad to say that I think I have some impr	1						
37	I took my 0st dose on Nov 10th 0000. For the month of December I experienced some lower back pain extreme fatig	1						
38	After some initial nausea and related G/I stuff I am thrilled to report that my walking has improved significantly. F	1						
39	I have been taking it for 1 weeks so far so it s hard to put a effectiveness rating but at least I haven t had any new	1						
40	So far so good. I m not experiencing any side effects no exacerbations. :) An answer to my prayers...	1						
41	ve been on Gilenya for 0 year this week and have had no new symptoms or side effects AT ALL. I ve been extremely	1						
42	They really downplay possible serious side effects. I have almost gone blind taking this. I developed Macular Edema	0						
43	I was a bit hesitant of taking Gilenya as I already have bad eye issues and asthma. I,Äôve been on it for 6 months no	1						
44	I started Gilenya in Feb 0000 after failing all the injectable MS meds. I was diagnosed in 0007 and initially I felt th	0						

Figure 51 – Sample of Gilenya Dataset

5.2.2   Data Mining Methodology

In order to generate new knowledge from the discussed reviews on specific topic, the CRISP-DM method was used to extract the most common words. CRISP-DM stands for cross-industry process for data mining. This methodology produces a well-structured approach in six phases when developing a data mining project (Daniel T. Larose, 2015).
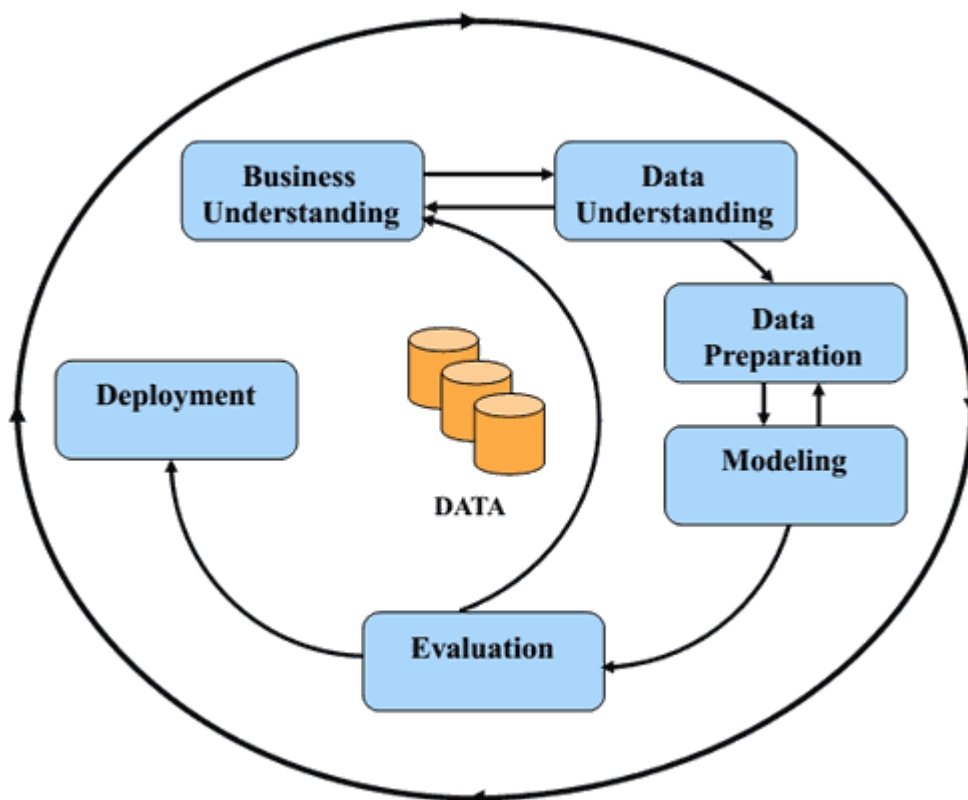
Figure 52 – CRISP-DM Diagram – Source: (Daniel T. Larose, 2015)

The first phase is the business/research understanding phase. In this phase, the project's aims, and terms are clearly outlined and translated into a data mining problem definition. A fundamental strategy to accomplish the requirements is then prepared. The second phase of the CRISP-DM method is phase of understanding data. This stage consists of data collection and using investigative data analysis to discover fundamental observations, evaluate the quality of the data and then, if necessary, choose subsets that may include actionable patterns (Daniel T. Larose, 2015).

The third labour-intensive phase is the data preparation phase which converts the initial raw data into the final data set. Suitable variables and cases are chosen and analysed, transformations are performed on specific variables, if necessary, and the raw data is cleansed and prepared for the modelling phase which is phase four.

The modelling phase applies the most suitable modelling techniques and calibrates the settings of the model to improve the results. Frequently, various techniques may be used for the same data mining problem. In some cases, it may be necessary to repeat the third phase (data preparation) to bring the data form in line with the explicit requirements of a certain data mining technique. The model(s) in this phase are delivered to the fifth phase, the evaluation phase. The model(s) are checked for quality and efficiency before deployment in the field and to ascertain if the model achieves and meets the aims that were set in the first phase. Additionally, checks are made if some important parts of the business/research problem have not been included (Jen Stirrup, 2017).

The final, sixth phase of the CRISP-DM method is the deployment phase. Even though the model has been created, the project is not completed. An example of a simple deployment is carried out with a report being created' An example of an intricate deployment could be an implementation of a parallel data mining process used in an alternative department. For the business industry, usually the customer carried out the deployment based on the model (Jen Stirrup, 2017).

5.2.3   Data processing

The flowchart used to extract the most common words to help knowledge creation is shown in figure 53.
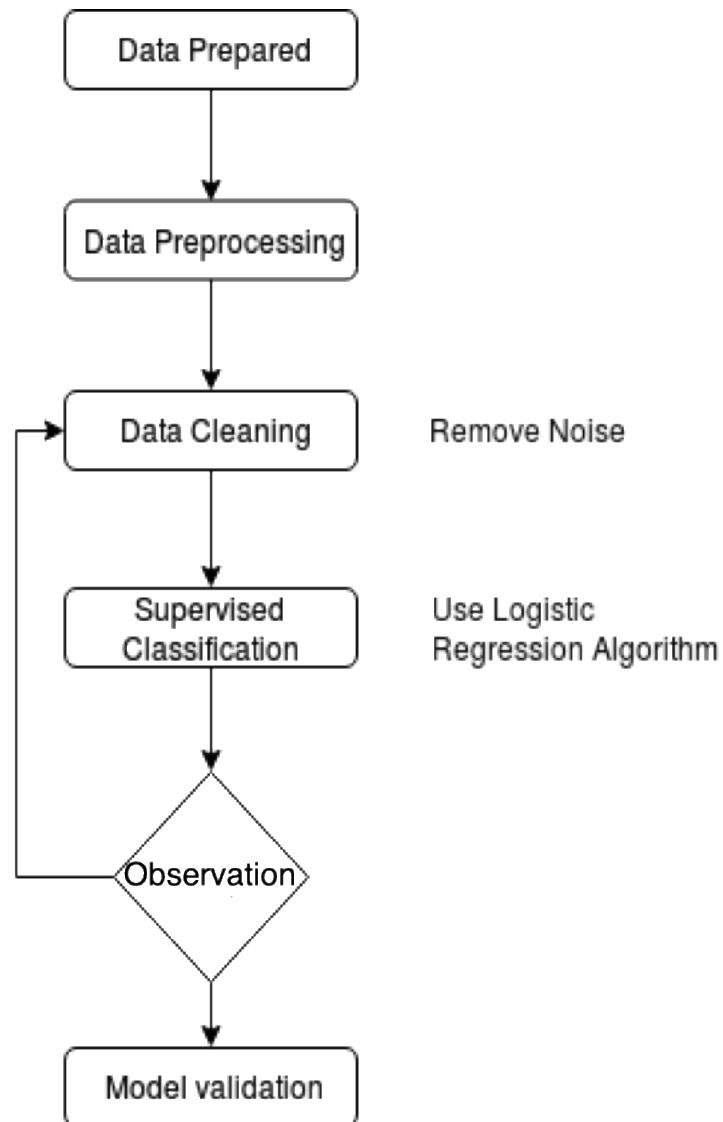
Figure 53- Flowchart of Knowledge extraction process

5.2.4   Data preparation

Pre-processing of data is essential in machine learning tasks. It is a data mining method used to convert raw data into a beneficial format. Pre-processing also plays a huge role in text mining, natural language processing (NLP) and information retrieval (IR).

Our datasets are selected, shuffled and randomly divided into training and test sets in an 80:20 ratio respectively. As stated by (Oded Maimon, 2005) "To measure the accuracy of the model, the data needs to be split into these two parts - the training set to fit and tune the model and the testing set to create predictions and evaluate the model at the very end".
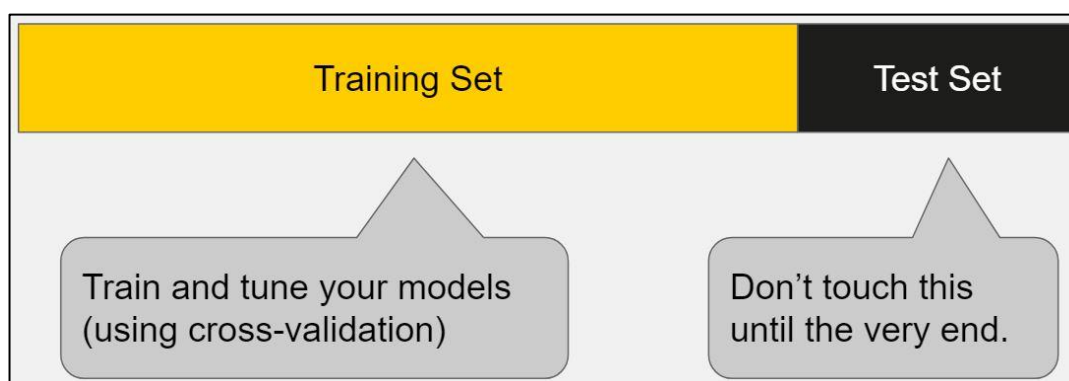
Figure 54 - Train Test Split - Source: (elitedatascience.com, 2018)

Information retrieval (IR) is defined by (Gurusamy, 2014) "as the process of determining which documents should be recovered to meet a user's need for information. The information is represented by a query or profile containing one or more search terms, plus some additional information such as the weight of the words". (Gurusamy, 2014).

5.2.5   Pre-processing and NLP techniques

The general pre-processing and NLP techniques used in this thesis are as follows:

5.2.5.1   Punctuation removal.

The words are kept the same but with without any punctuation, i.e. no commas or quotes and   contractions are kept together. "One way would be to split the document into words by white space (as in *2. Split by Whitespace*), then use string translation to replace all punctuation with nothing e.g. remove it" (Brownlee, 2017).

Python provides a constant called string.punctuation as mentioned by (Brownlee, 2017) " string.punctuation it provides a list of punctuation characters". Figure 55 shows an example of punctuation removal being used in Python programming language.

```
Python code:
import string
input_str = "This & is[an] example ? { of
}
string.with. ? punctuation!!!!"#Sample string result =
input_str.translate(string.maketrans("", ""), string.punctuation)
print(result)

Output: This is an example of string with punctuation
```

Figure 55 - Sample of Python code for punctuation characters – Source: (Brownlee, 2017)

5.2.5.2   Html tag removal.

As described by (Brownlee, 2017) "Unstructured text contains a significant amount of noise, as stated by especially if retrieved using techniques such as web or screen scraping". HTML tags are typically one of these components which add little or no value when understanding and analysing text (Brownlee, 2017).

```
def strip_html_tags(text):
soup = BeautifulSoup(text, "html.parser")
stripped_text = soup.get_text()
return stripped_text

strip_html_tags('<html><h2>Some important text</h2></html>')
```

Figure 56- Sample of Python code for HTML tags removal – Source: (Brownlee, 2017)

It is quite evident from the above output that we can remove unnecessary HTML tags and retain the useful textual information from any document (Brownlee, 2017).

5.2.5.3   Convert text to lowercase.

An example is shown on figure 57 showing the conversation from upper to lowercase.

```
Python code:
input_str = "The 5 biggest countries by population in 2017
are China, India, United States, Indonesia, and Brazil."
input_str = input_str.lower()
print(input_str)

Output: the 5 biggest countries by population in 2017 are
china, India, united states, Indonesia, and brazil.
```

Figure 57- Sample of Python code for converting to lowercase – Source: (Brownlee, 2017)

5.2.5.3 Lemmatization.

Lemmatization is to cut down words, as described by (Brownlee, 2017) "Lemmatization is to cut down inflectional forms to a common base form. As opposed to stemming, lemmatization does not simply remove inflections. Instead it uses lexical knowledge bases to obtain the correct base forms of words".

Lemmatization tools are available in Python packages which was used on this thesis such as NLTK (Brownlee, 2017).

```
Code:
from nltk.stem import WordNetLemmatizer
from nltk.tokenize import word_tokenize
lemmatizer=WordNetLemmatizer()
input_str="been had done languages cities mice"
input_str=word_tokenize(input_str)
for word in input_str:
    print(lemmatizer.lemmatize(word))

Output:
be have do language city mouse
```

Figure 58 - Python code showing how Lemmatization works – Source: (Brownlee, 2017)

5.2.5.4 Stop words removal.

Words recur in any document very frequently but some of those words have little meaning as they are used to join words together in a sentence. As described by (Gurusamy, 2014) "Words such as 'the', 'to', 'and', 'are', 'this' etc. are not useful in the classification of documents, therefore they need to be removed. Nevertheless, the development of a 'stop words' list is difficult and inconsistent between textual sources". Python has its own stop words package which the researcher has used in this thesis. As stated by (Silhavy, 2018) "This process will help reduce the text data and improve the DM performance".

5.2.6   Results Analysis and Discussion

5.2.6.1   Word Clouds

A word cloud is a collage of randomly arranged words where the size of each word is proportional to its frequency in the corpus. It provides an idea of what words are represented in the corpus of each class. However, they do not clearly indicate accurate information especially when comparing each class. (Robinson, 2017). A sample of the code that generates a word cloud is shown in figure 59.

```
01.    def build_wordcloud(reviews):
02.        img = io.BytesIO()
03.        text = " ".join(line for line in reviews)
04.        stopwords = set(STOPWORDS)
05.        stopwords.update(['started','gilenya','ms','drink', 'now', 'wine', 'flavor', 'flavors', 'side', 'effe
06.        wordcloud = WordCloud(stopwords=stopwords, max_words=100, background_color="white").generate(text)
07.        plt.imshow(wordcloud, interpolation='bilinear')
08.        plt.axis("off")
09.        #plt.show()
10.        plt.savefig(img, format='png')
11.        img.seek(0)
12.        graph_url = base64.b64encode(img.getvalue()).decode()
13.        plt.close()
14.        return 'data:image/png;base64,{}'.format(graph_url)
15.
```

Figure 59 – Sample of Python code to generate word cloud

5.2.6.2.1   Word Cloud for Citalopram Dataset

As shown below, the bigger the words are the most mentioned phrases including headache, depression and tiredness. At first glance it can concluded that the medication is used as an antidepressant from the responses to the Citalopram data. Furthermore, the phrase 'headache' could be a side effect from the medication.

Figure 60 – Word Cloud for Citalopram Drug Dataset

5.2.6.2   Word Cloud for Amazon Echo Dataset

As shown below in figure 61, knowledge can be gathered from what the consumer has been discussing about the product. There are positive responses in the Amazon Echo Dataset, like 'love', 'great', 'nice', 'fun' and 'easy', which have been mentioned multiple times due to the size of the word. The characteristics of the item have been mentioned, 'music', 'screen', 'sound', 'video' 'camera' 'call' and 'purchase'.

Figure 61 – Word Cloud for Amazon Echo Dataset

5.2.6.3   Word Cloud for Gilenya Dataset

Figure 62 is a Word Cloud of the dataset for the medication Gilenya. The phrases 'symptoms', 'better' and 'headaches' are examples of the patients' reviews about the use of medicine from which they have encountered side effects.

Figure 62 – Word Cloud for Gilenya Dataset

5.2.6.4   Feature Extraction

Once the text data is cleaned it is not yet ready for modelling. The text needs to be converted into a numerical form. Two different vectorisers were investigated to identify which one was best for the data. A logistic regression model was then used as it is a fast and simple classifier (scikit-learn.org, 2018). Figure 63 shows an example taken from the coding that was used on this thesis.

```
from sklearn.linear_model import LogisticRegression,
LogisticRegressionCV
from sklearn.metrics import accuracy_score

lr = LogisticRegression(random_state=1)
```

Figure 63 – Packages needed for Feature Extraction

5.2.6.4.1   TF-IDF Vectorizer

TF-IDF (term frequency–inverse document frequency) vectorises words by taking into account the number of a words in a given document and the frequency between documents. (Robinson, 2017). As stated by Olga Davydovaand "The importance increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus" (Davydova, 2019).

Term Frequency**:** The more frequent terms in a target article, the higher the score

$$\text{tf}(t,\, d) = N_{\text{term}}$$

Equation 1 – Term Frequency Equation – Source: (Robinson, 2017)

Inverse Document Frequency: The more common in other articles, the lower the score

$$\text{idf}(t,\, D) = 1 + \log\left(\frac{1 + N_{\text{Documents}}}{1 + N_{\text{Documents that contain term}}}\right)$$

$$\text{tf-idf}(t,\, d,\, D) = \text{tf}(t,\, d) \cdot \text{idf}(t,\, D)$$

Equation 2 – Inverse Document Frequency – Source: (Robinson, 2017)

If a word is common in most documents it is suppressed and rare words are given more influence showing they are highly specific for a particular document.

```
1. from sklearn.feature_extraction.text import TfidfVectorizer
   #vectorize our reviews so we assign numbers to each reviews using the
   ngram techinques
2.  #we started with uniGram
3.  vectorizer = TfidfVectorizer(max_features=1000, ngram_range=(1, 3))
4.  #vectorizeration is a process of converting words to number i.e did y
    ou see me so did = 1 you=2 and see = 3 me = 4
5.  X = vectorizer.fit_transform(reviews_train_clean)
6.  X_test = vectorizer.transform(reviews_test_clean)
```

Figure 64– Code for assigning numbers to words

5.2.6.4.2   Results for TF-IDF (Without N-Gram)

5.2.6.4.2.1   Results for Citalopram Dataset

WordCloud & Most used words in our reviews
Conclusion based on whats mostly talked about
The concerns that are commonly being discussed in our reviews is about our patients having :

- depression
- med
- week
- anxiety
- year
- better
- tired
- good
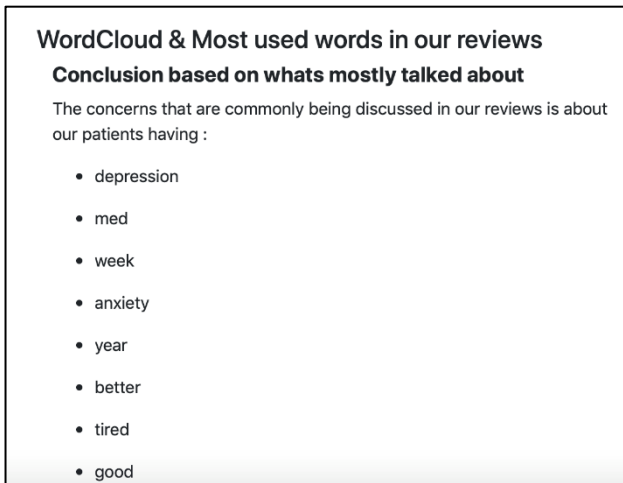
Figure 65 – TF-IDF Results for Citalopram Drug Dataset

As shown on figures 65,66 and 67 on the left these are the most mentioned words that have been discovered using TF-IDF on all the three datasets/ This is the initial step for knowledge discovery from the reviews feature.

5.2.6.4.2.2   Results Gilenya Dataset

WordCloud & Most used words in our reviews
Conclusion based on whats mostly talked about
The concerns that are commonly being discussed in our reviews is about our patients having :

- since
- year
- better
- pill
- symptom
- week
- hello
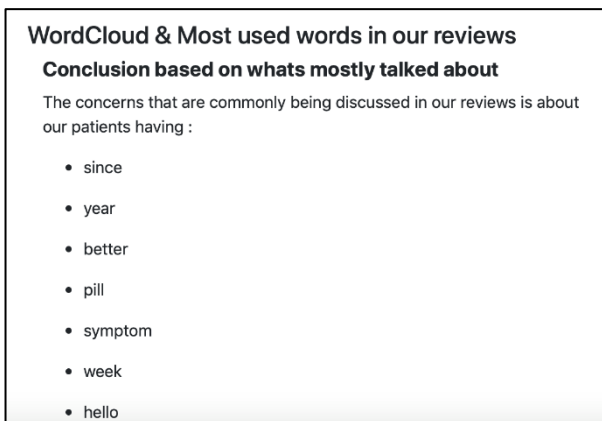
Figure 66 – TF-IDF Results for Gilenya Drug Dataset

5.2.6.4.2.3   Results for Amazon Echo Dataset

- love
- echo
- great
- product
- video
- screen
- music
- amazon
- easy
- bought
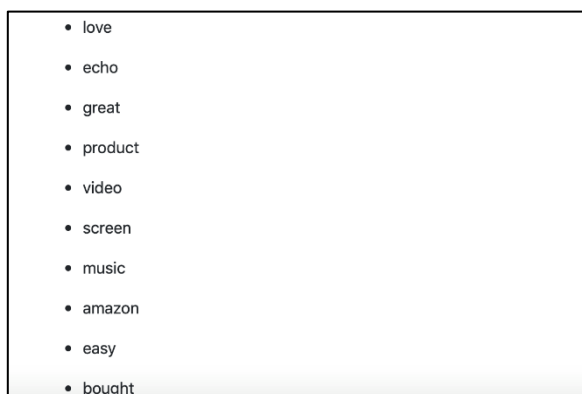
Figure 67 – TF-IDF Results for Amazon Echo Dataset

5.2.6.4.3   N-Gram

An N-gram is an N-token sequence of words: a 2-gram (bigram), is a two-word sequence of phrases like "please move", "move your", or "your homework", and a 3-gram (trigram) is a three-word sequence of words like "please move your", or "move your homework" (Jurafsky, 2014). Figures 65-67 have shown that it is not practical to rely on a single tokens word to hold as much information compared to double or triple combined.

N-gram will take a range of n tokens to use as features, for example:

(1, 1) default: using only a singular token

(1, 2) bigram: using a range of singular and double tokens

(1, 3) trigram: using a range of singular, double and triple tokens

As illustrated in figure 68, that the trigram sentence "severe constant headache" is what most patients have struggled with as a side effect, therefore the use of n-gram is beneficial in better understanding compared to TF-IDF. It can also be concluded from figure 68 that the accuracy increases with the range as it gives more information for the model to predict. Here (4, 4) range has the highest score.

WordCloud & Most used words in our reviews

Conclusion based on whats mostly talked about

The concerns that are commonly being discussed in our reviews is about our patients having :

- med began severe headache

- severe constant headache well

- abilify lamectal worked well

- experience good doctor mentioned

- caused hormonal inbalance menopause

- dosage adjusted completely person

- involuntary jerkinghands leg awhile

- depression changexanaxx never pill

- stay awake always tired

- awake always tired tablet

- bad pill dizziness weak

Figure 68 – n-gram (4,4) results
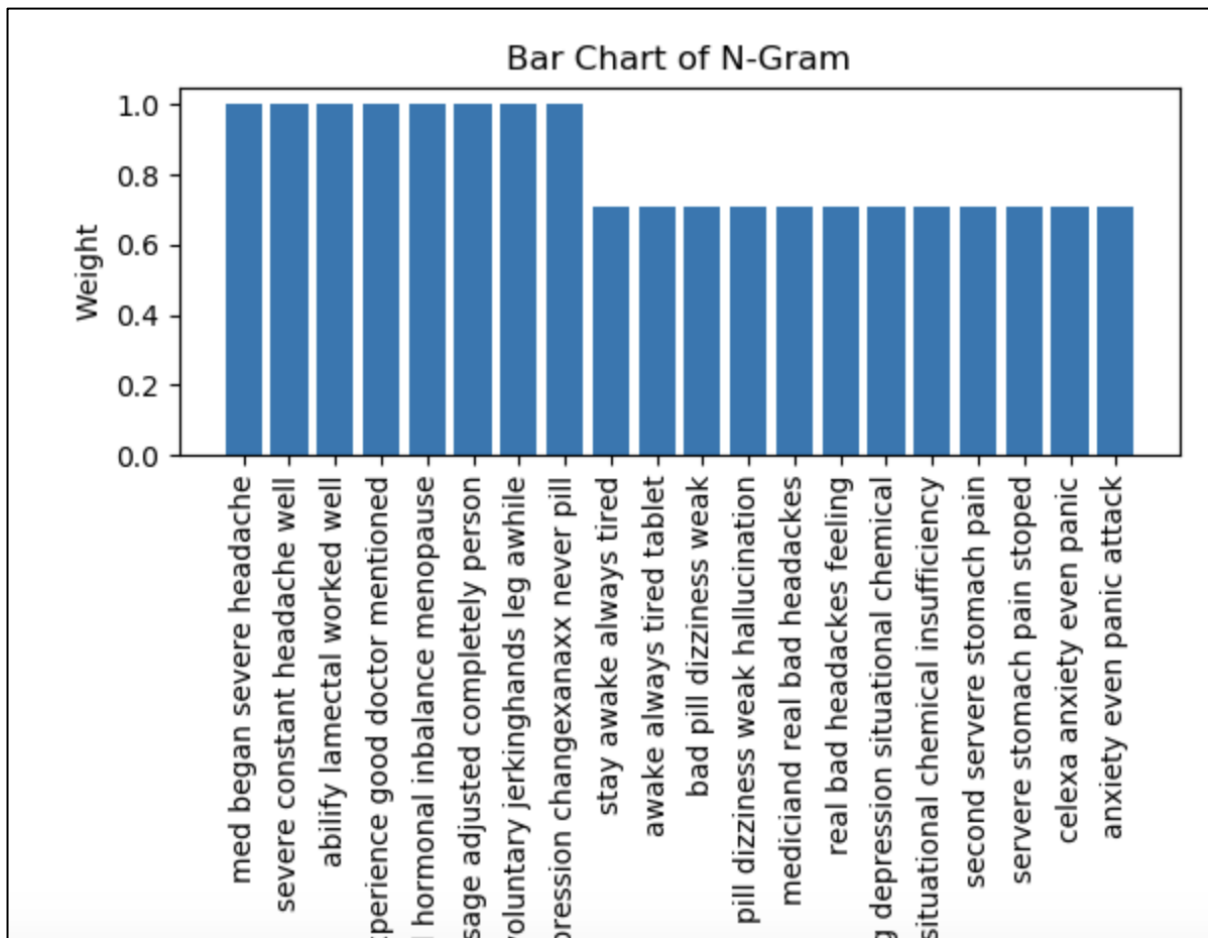
Figure 69 – Frequent word Vs Word Weight Bar Chart

5.2.5.5 Algorithms comparison and results

In addition to the word weight shown on figure 69, multiple DM algorithms were used and a code written to compare the highest F-score on each dataset.

Machine learning algorithms that were used and compared were support vector machine (SVM), logistic regression, decision tree classifier, random forest, naïve Bayes and neural network.

The models proposed was trained using these techniques and used F-score and accuracy metrics to compare the performance of each datasets. Whenever a user enters a new topic, the e-Toolkit will automatically take the reviews of this topic and generate knowledge as well as run those reviews through the algorithm comparison process, this helps in finding the best algorithm to be used on those specific reviews.

The above-mentioned algorithms are explained below.

5.2.5.5.1   Logistic Regression (LR)

LR is used for the supervised classification part. As mentioned by (Hilbe, 2009) "LR is a statistical method for analysing a dataset, which has one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes).   It is used to predict a binary outcome (0 or 1, Yes or No, True or False) given a set of independent variables. To represent a binary or categorical outcome, dummy variables are used" (Hilbe, 2009). Logistic regression can be considered as a special case of linear regression when the outcome variable is categorical, where log of odds as dependent variable is used. (Hilbe, 2009).

In easier terms, LR can predicts occurrence probability of an event by fitting data to a **logit** function.

$$\text{logit}(x) = \log(x \,/\, (1 - x))$$

<div align="center">Equation 3 – Logit function – Source: (Hilbe, 2009)</div>

A plot of the logit function is illustrated in figure 70. As stated by (Brixius, 2018) "The value of the logit function heads towards infinity as p approaches 1 and towards negative infinity as it approaches 0".
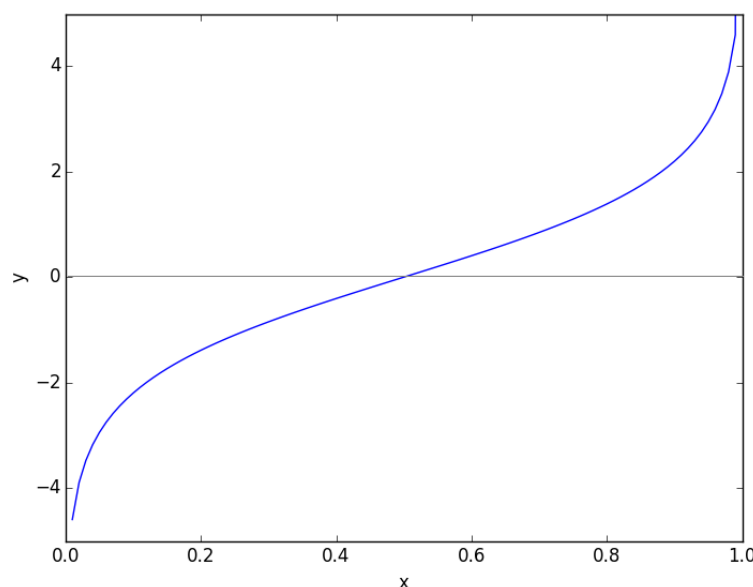
<div align="center">Figure 70 – Plot of the Logit function – Source: (Brixius, 2018)</div>

In this research, the binary outcome is either 0 or 1 which represent either having a negative review or a positive review respectively.

5.2.5.5.2   Support Vector Machine (SVM)

SVM is a supervised learning algorithm as described by (Hamel, 2011) SVM can "outputs an optimal hyperplane which categorises new examples. In two-dimensional space, this hyperplane is a line dividing a plane in two parts where in each class lay in either side". There are many variants of SVM. Polynomial fitting SVM is used for this research.

5.2.5.5.3   Decision Tree

A decision tree is a classifier expressed as a recursive partition of the in-stance space. It is described by (Lior Rokach, 2015) that "decision tree consists of nodes that form a rooted tree, meaning it is a directed tree with a node called "root", that does not have any incoming edges. All other nodes have exactly one incoming edge.

A node with outgoing edges is called an internal or test node. All other nodes are called leaves (also known as terminal or decision nodes)" (Lior Rokach, 2015). Decision Tree has been used on this thesis as a way to verify our results as shown on the next chapter.

5.2.5.5.4   Random Forests

Random forests (or random forest) are made of many decision trees, as stated by (Lior Rokach, 2015) "Random Forests are an ensembles of learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and modelling the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees".

Random Decision forest can help correct decision trees when overfitting to their training set (Lior Rokach, 2015).

5.2.5.5.5   Naïve Bayes

Naive Bayes classifiers are a family of simple "probabilistic classifiers" it is described by (Lior Rokach, 2015) as probabilistic classifier which is "based on applying Bayes' theorem with strong (naive) assumptions that every feature is independent of other features". Bayes' theorem calculates the output depending upon probabilities of features (Lior Rokach, 2015).

5.2.5.5.6   Artificial Neural Network

Artificial neural networks consist of a group of neurons working together to estimate an output. A single artificial neuron analyses a "weighted sum" input then adds a bias and make it makes a decision to whether it fired or not.

$$Y = \sum_{n=1}^{N}(weight * inputs) + bias$$

Equation 4 - Artificial Neural Network – Source (Giuseppe Ciaburro, 2017)

The value of Y is the output of a neuron.

5.2.5.6 Comparison of Algorithm Results

The calculated results from Citalopram drug, Gilenya drug and Bluetooth speaker (Amazon) are illustrated in figures 71 – 73 below:

5.2.5.6.1   Citalopram Dataset Results

Classification techniques	F score
SVM	0.9322
Naïve Bayes	0.9409
Logistic regression	0.9354
neural network	0.9834
random forest	0.9391
decision tree	0.9231

Table 5 – F-Score multiple DM results Citalopram Dataset
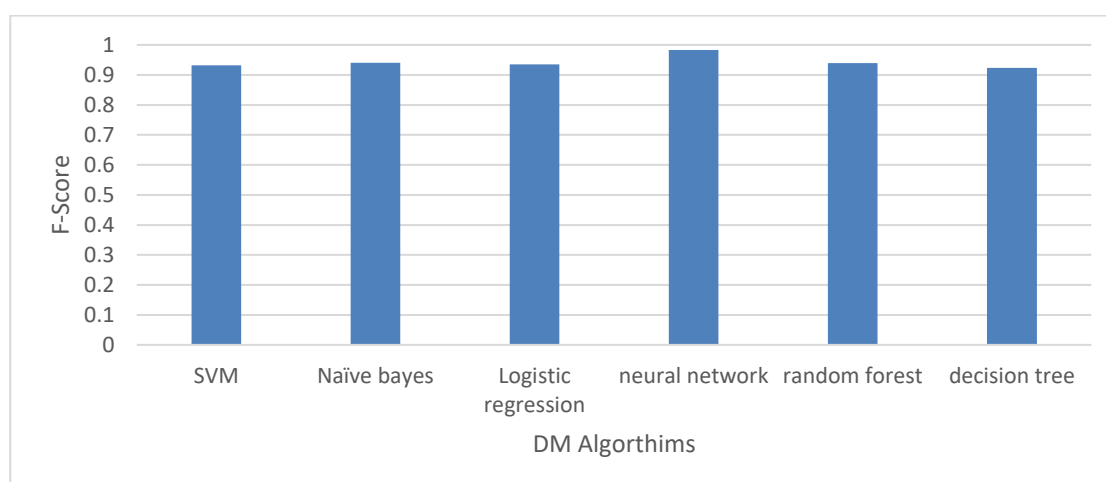
Figure 71 – F-Score Results for the Citalopram Drug Dataset

5.2.5.6.2   Gilenya Dataset Results:

Classification techniques	F score
SVM	0.923
Naïve Bayes	0.924
Logistic regression	0.925
neural network	0.923
random forest	0.923
decision tree	0.923

Table 6– F-Score multiple DM results Amazon Dataset
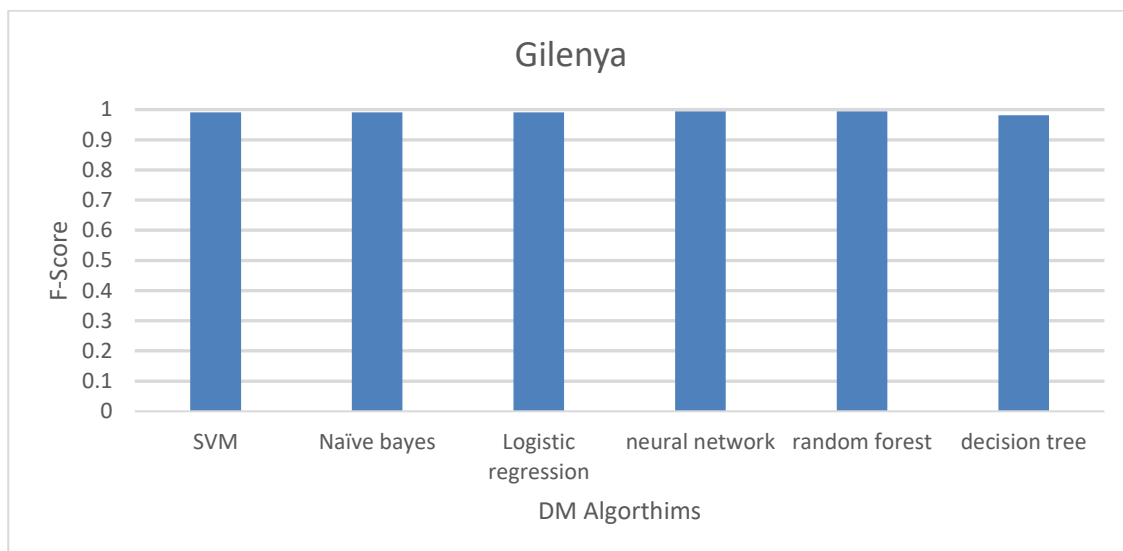
Figure 72– F-Score Results for the Gilenya Drug Dataset

5.2.5.6.3   Bluetooth Speaker (Amazon) Dataset Results:

Classification techniques	F score
SVM	0.991
Naïve Bayes	0.991
Logistic regression	0.991
neural network	0.9939
random forest	0.99401
decision tree	0.9817

Table 7 – F-Score multiple DM results Amazon Dataset
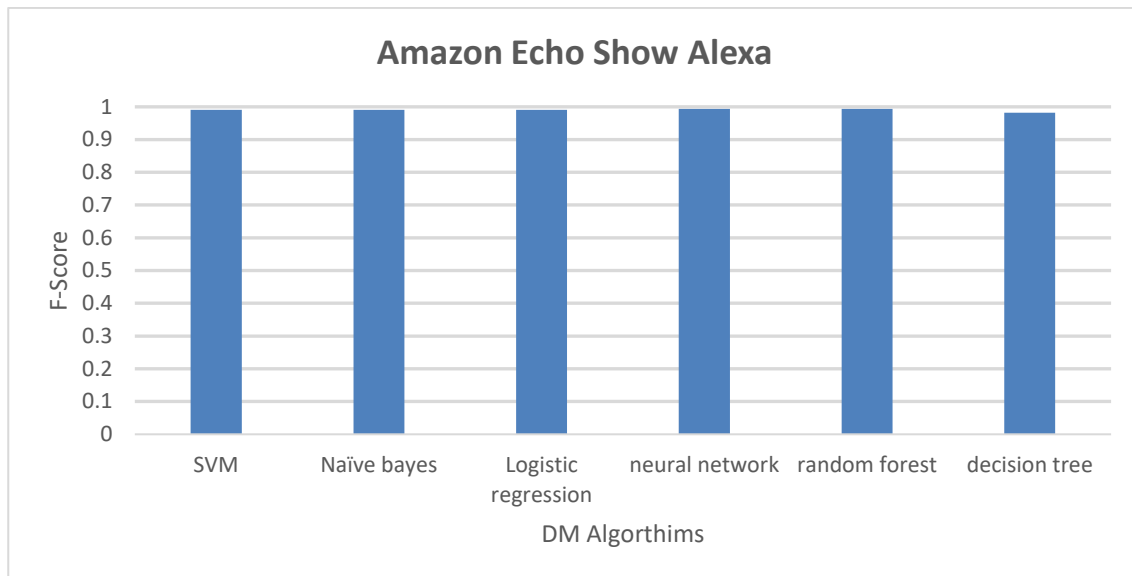
Figure 73 – F-Score Results for the Bluetooth Speaker (Amazon) Dataset

5.3   Results Discussion

The overall results presented in tables 5-7 and visualised in Figure 68 and 69 reveal that the proposed model demonstrated a very competitive performance as it shows the most common issues that the patients are currently facing from the medicine reviews dataset. This fulfils one of the main goals of this thesis, that is merging KM and KD in one framework and that converting Tacit knowledge to Explicit knowledge is achievable.

As shown in Figure 68-69 the top 3 highest concerns were "Severe Headache", "Caused hormonal imbalance", and "Always tired". Those concerns were passed to a medical team via an online meeting (consisted of 2 medical doctors and 2 pharmacist) to verify the combination mode within the SECI model. This is formal meeting by medical staff after discovering knowledge from the data mining techniques in the externalization mode. The questionnaires given to the participants are shown in Appendix K and were verified by all 4 medical professionals as accurate side effect of this medication and therefore an official document was issued on the e-Toolkit, a comprehensive validation process is shown in section 6.5 of this thesis.

Furthermore, using other datasets such as amazon reviews showed that it is possible for the implemented framework idea to be used in other non-medical fields and not only on Diabetes. Thus, further studies merit exploring to ascertain how this framework could be used on other fields.

Additionally, the comparison of different DM algorithms presented in Table 5-7 showed excellent F-Score results, in most cases an overall average of 0.8,

Overall, our framework investigated cutting-edge concepts in knowledge discovery and knowledge management as well as the knowledge dissemination area.

5.4   Chapter Summary

This chapter explained in detail the experimental implementation and the obtained results of text mining analysis from the knowledge discovery part of the SECI model (blog) which was mentioned in previous chapters. It also discussed the evaluation process, datasets used, the knowledge extraction implementation for KD and the additional feature this chapter has proposed comparing four DM algorithms and analysis of the results.

The next chapter will discuss the process of Knowledge Dissemination using the SECI proposed Model on the e-Toolkit, as well as the validation of the e-Toolkit.

CHAPTER SIX: E-TOOLKIT KNOWLEDGE DISSEMINATION, IMPLEMENTATION AND VALIDATION

6.1   Introduction

In the previous chapters the SECI model was discussed and explained thoroughly. The SECI model can be used in the electronic place to create and share knowledge among individuals and groups. This knowledge can be managed using Web 2.0 is a technological tool that can be used to create and share knowledge via forums, blogs, Wikis, etc. As mentioned previously, most researchers have used the SECI model to disseminate knowledge. However, the proposed SECI model will create and disseminate the knowledge through the e-Toolkit.

Therefore, a website was designed to include these technological tools to create and sharing knowledge. This chapter will first discuss the development of the e-Toolkit followed by an overview of the e-Toolkit, followed by a validation of the e-Toolkit and a summary of work done so far.

6.2   E-Toolkit Development

This section, presents the design and implementation of our e-tool kit which we have called a Diabetic Patient Monitoring System. Part of our aims is to develop this e-Toolkit to support diabetic patients and healthcare professionals to better communicate and to provide effective creation and sharing of knowledge related to the control of diabetes mellitus.

The development of e-Toolkit is defined by (Diffily, 2006) as the procedure of building a new website, or adapting features of an existing one, involving eight phases illustrated as follows:

1. Planning: why is the site being planned and what will be built
2. Content: a list of the materials (content) to be included on the website
3. Design: make an outline to show how the content will be presented
4. Development: write the code and upload the content
5. Test: ensure everything works appropriately and correctly with all current browsers
6. Hosting: select a domain name and identify a web hosting provider
7. Publicity: build up activity by drawing attention to the site
8. Review: audit the site at intervals, to ascertain its usage and effectiveness.
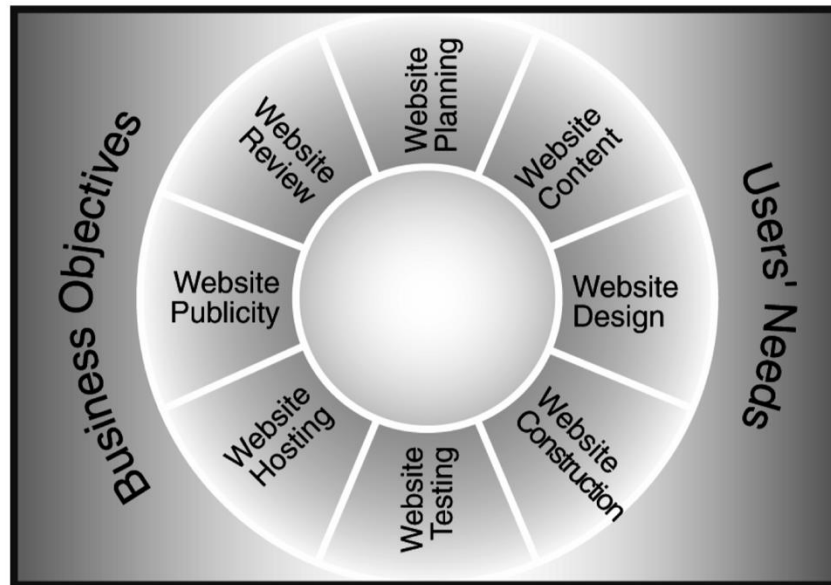
Figure 74 – The website development cycle. Source: (Diffily, 2006)

6.2.1 e-Toolkit Planning

Planning is considered to be one of the significant steps in the design of a website. It is a procedure to identify the users' needs and businesses' objectives, to start the Development Cycle (Diffily, 2006).

The objectives of building this website were determined, based on the survey conducted and the needs of healthcare providers and patients.

Thus, the objectives of this e-Toolkit are:

- To build an e-Toolkit to disseminate and create knowledge (SECI MODEL)
- To build an e-Toolkit to support diabetes education and self-management world-wide.

6.2.2 e-Toolkit Content

Website content includes most information and applications that need to be available online for Diabetes Management System e-Toolkit. It was essential to ensure that the information provided meets the needs of the patients (Diffily, 2006).

From the data gathered through the survey it was decided to use the MoSCoW method to define requirements by priority order. As Stated by (Stephens, 2015) "The most important requirements need to be met first for a greater chance of success" (Stephens, 2015).

6.2.2.1   MoSCoW

MoSCoW is a useful method used in software development to achieve a mutual understanding with investors on the "importance they place on the delivery of each requirement" (DSDM, 2018).

To decide on what content to be included on the e-Toolkit the MoSCoW approach is used. As Stated by (Ma, 2009), The letters MoSCoW stand for: "

- o "**M**UST have" - the requirements in this group must be contained in the project. Failure to deliver these requirements means the entire project would be a failure.
- o "**S**HOULD have" - the project should if at all possible, contain the requirements in this group.
- o "**C**OULD have" - the project should contain these requirements. But they are less important than the requirements in the "SHOULD have" group.
- o "**W**ON'T have" - is similar to a "wish list". It means that the requirements in this group are good requirements, but they will not be implemented in the current stage. They may be implemented in the next release" (Ma, 2009).

In a research like this where there is a time restraint, understanding the relative importance of all research related matters is important to have progression and meeting deadlines.

An illustration of the MoSCoW method is presented in figure 75.
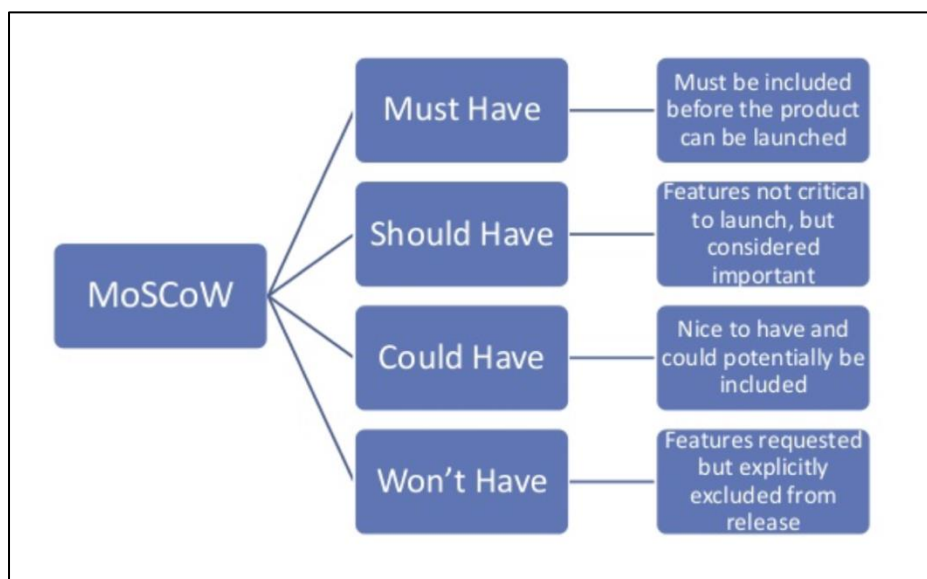
Figure 75 – Representation of the MoSCow Method – Source: (Stephens, 2015)

The reason why the MoSCoW method was used on the e-Toolkit is to set the importance and prioritise the features from High, Medium and Low. The specific use of "must, should, could or won't have" implies the outcome of failing to deliver that requirement. These features were decided from the results gathered by the surveys in chapter 3. Those requirements are illustrated below.

6.2.2.1.1 Must Haves

- The e-Toolkit must allow diabetics to enter all vital health information.

- E-Toolkit must be free to use for all diabetic and medical professionals.

- An ID must be assigned to every physician, nurse and medical staff member.

- Nurses and staff members must see the patient details the same way doctors do.

- An ID must be assigned to every diabetic.

- A secure communication messaging system between medical professionals and diabetic must be put in place

- Doctors who are web administrators, must be able to register and delete staff (such as nurses). Diabetic must be able to update their assigned medical professionals details.

- Blog/forum is required where patients could discuss certain topic such as medicines. Data mining techniques could then be used to create knowledge from those reviews

- Pages to present the knowledge discovered, generally in the form of PDF documents/guidelines. The documents to be published after approval from the medical team. This is to meet our research objectives in regards to documenting knowledge discovered.

- The e-Tool kit must have the capability to graphically present patient's data.

6.2.2.2 Should Haves

- The e-Toolkit should be accessible from any device and operating system.

- If this e-Toolkit is applied to the NHS, the e-Toolkit should check that every medical professional has a valid medical license and has no criminal convictions or any penalties on their medical license before registering to be part of the system.

- If this e-Toolkit is applied to the NHS, the e-Toolkit should have access and NHS medical systems to update new knowledge discovered by the e-Toolkit after approval by medical staff.

6.2.2.3   Could Haves

- Doctors could communicate with each other.
- The patients could be able to pick what data the doctors can see. For example, a patient who has written a personal note in the e-Toolkit's notes area, marking this note as a "personal diary", then the e-Toolkit should not allow their assigned health care professional to see this note.

6.2.2.4   Won't haves

Legality should be taken into account so therefore the e-Toolkit will not have the option to upload any personal information from the patient NHS record into the system.

6.2.3   e-Toolkit Design

In order to design a successful e-Toolkit, Jakob Nielsen defined *'ten good deeds'* in web design that increase the usability of virtually all websites. The ten elements was

Identified by (Jakob, 1999) as follows:

➢ "Allocate the name and logo of the organisation on every page and hyperlink the logo to the home page

➢ Add a search engine

➢ Describe the headlines and page titles as simply as possible

➢ Build the page so as to simplify scanning

➢ Use hypertext to structure the content space

➢ Display product photos

➢ 'Use relevance-enhanced image reduction'

➢ Use link titles

➢ Ensure that all essential pages are accessible for users with disabilities

➢ Adopt what the best websites do in an appropriate technique". (Jakob, 1999)

The above points were used on the proposed e-Toolkit, Figure 76 – 80 illustrates the initial e-toolkit wireframes.
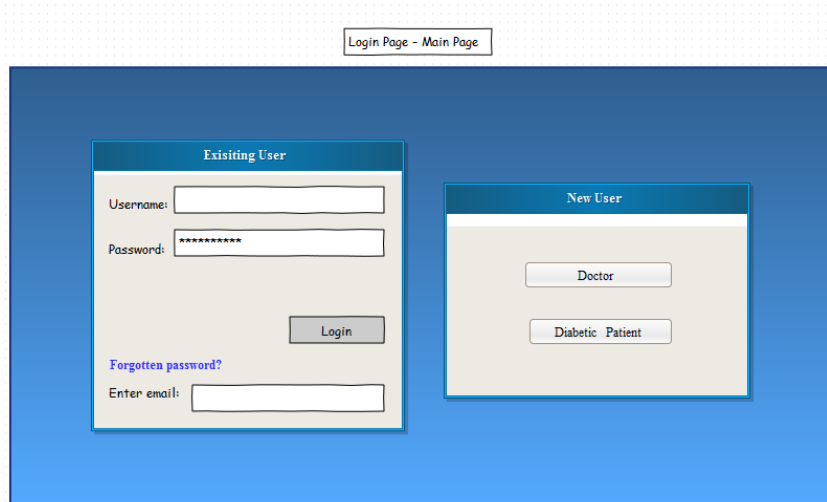
Figure 76 – e-Toolkit home page
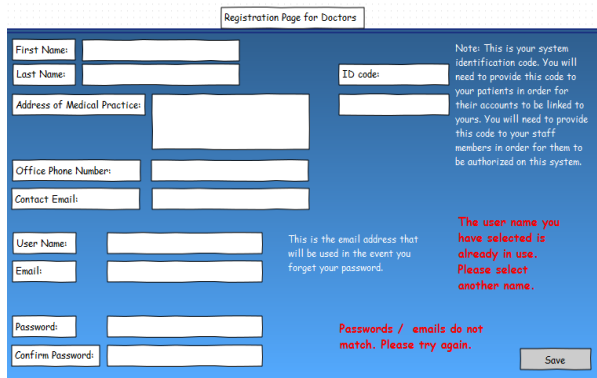
Figure 77 – e-Toolkit registration page
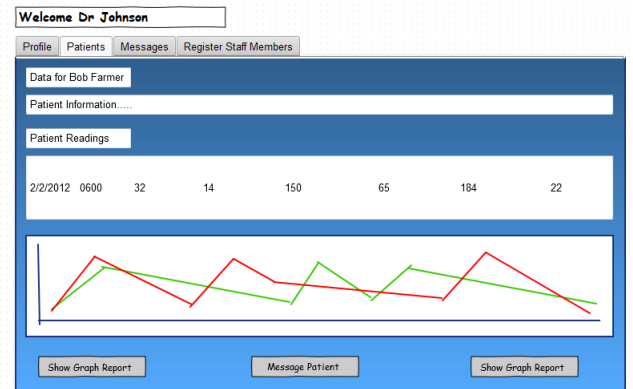
Figure 78- e-Toolkit doctor's account on looking patients' readings
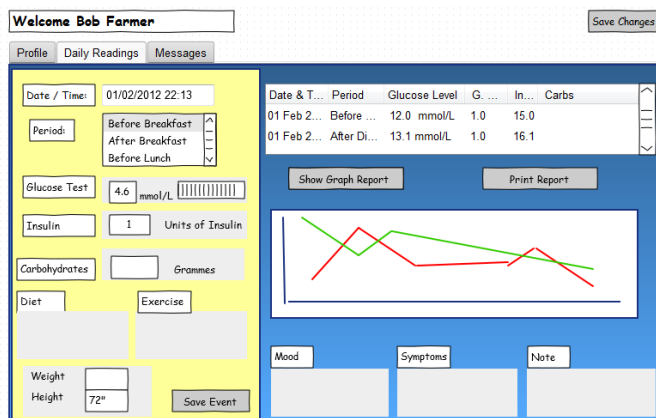
Figure 79- e-Toolkit patient's own readings and info
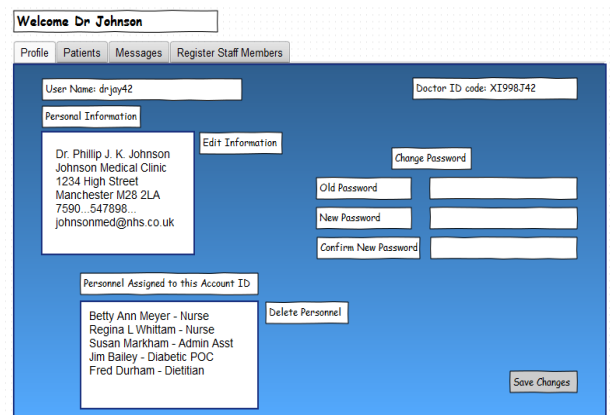
Figure 80- e-Toolkit doctor's profile page

6.2.4 e-Toolkit Development (programming codes to consider)

Selecting the right programming language to utilise for this project is an important choice that a developer faces when starting a new project. As it will have an extensive and profoundly established effect. Ideally, the advantages and disadvantages of all relevant programming languages should be considered to determine which is best suited for the project.   The selection process ought to be precisely weighed while considering the following:

- Is the procedure to determine a problem natural or does the language require additional documentation to comprehend why a line of code was made or what it expects to do?
- Is it easy to debug and test?
- Is there cost involve or is it free to use?
- Will the cost increment when connected to business project?
- Is it sufficient to achieve the research goals?
- How easy is it to get practical experience? E.g. open source projects? Easy to install and/or free to use?
- How much work is asked for in that language?
- Does the language have the basic features to construct the types of features you wish to include (if at all)?

6.2.4.1   Python

Python is a high-level programming language. It was introduced in 1991, it was created by Guido van Rossum. Python does not need compiling, which is a lot faster, making the test edit debug cycle substantially faster. As mentioned by David Kuhlman, "Python code is very easy to debug, since in case of a bad input, a bug will only raise an exception, once the interpreter discovers the error. If the program doesn't catch the exception it will print a stack trace" (Kuhlman, 2012).

The main disadvantage of Python is that it is slow to process, it is "much slower than C++, PHP and some other programming languages" (Onur R B, 2019). It is not suitable for multi-processor/multi-core work and has many limitations with database access. It is not a good programming language for mobile web browsing or app development. Any errors will only appear at the runtime and therefore it requires many more testing.

However, there are many advantages to Python, the code and syntax is much easier to read and write than many other languages (Kuhlman, 2012). Programmers can always add new modules that are implemented and compiled in a language such as C or C++.

6.2.4.2   Hypertext Pre-processor (PHP)

As it explained on php.net, "PHP is a widely-used open source general-purpose scripting language that is especially created for web development and can be embedded into HTML." (php.net, 2018), According to a web technology survey by the business W3Techs, 61.4% of the 10 million largest websites use PHP 5 on the server side (w3techs.com, 2019). "It was originally created by Rasmus Lerdorf in 1994, the reference implementation of PHP (powered by the Zend Engine) is now produced by The PHP Group" (documentation.online.net, 2018). Unlike Python and JSP, PHP is relatively straightforward to learn, enabling new programmer to start programming quickly to an acceptable standard. (php.net, 2018) (Whitehead, 2001).

6.2.4.3   JavaServer Pages (JSP)

JavaServer Pages (JSP) is a technology which provides a way to "dynamically generate web pages and simplifies the process of developing web-based applications. JSP-enabled pages use a combination of HTML, XML, and short Java scripts". JSP-enabled pages which is very similar in form and function to pages that was made by ASP, PHP or JSP the difference is the price they are very expensive to use and compile compared to PHP and Python (Indiana University, 2018).

6.2.5   Comparison of Programming languages

After reviewing the advantages and disadvantages of the above languages, it was decided to use PHP for the front end and Python for the Data Mining part.

	Python	PHP	JSP
Easy to understand?	X	X	
Easy to troubleshoot?		X	
Intuitive?	X	X	
Requires extensive notes and explanations to understand code?			X
Quick to create code?		X	
Executes quickly?		X	
Tests easily?		X	
Method dependent?			
Simple syntax?		X	
Framework dependent?	X		
Rigid?			X
Easily adaptable to changing criteria or future needs?		X	
Easily / widely supported by ISPs?		X	
Open source?	X	X	
Commercially expensive or complicated?	X		
Portable and agile?	X	X	
Serve project purpose?	X	X	
Lots of online support?		X	X
Lots of resources?		X	
Personal experience and knowledge?		X	

Table 8 – Comparison of E-Toolkit Programming Languages

6.2.6 Test

Website testing is the process used to assess the e-Toolkit, based on the previously agreed requirements. It ensures that the website meets the minimum standards to achieve its goal and objective (Diffily, 2006)

The researcher has carried out many tests on the e-Toolkit such as the following.

1. The first method used website testers such as website-grader.com
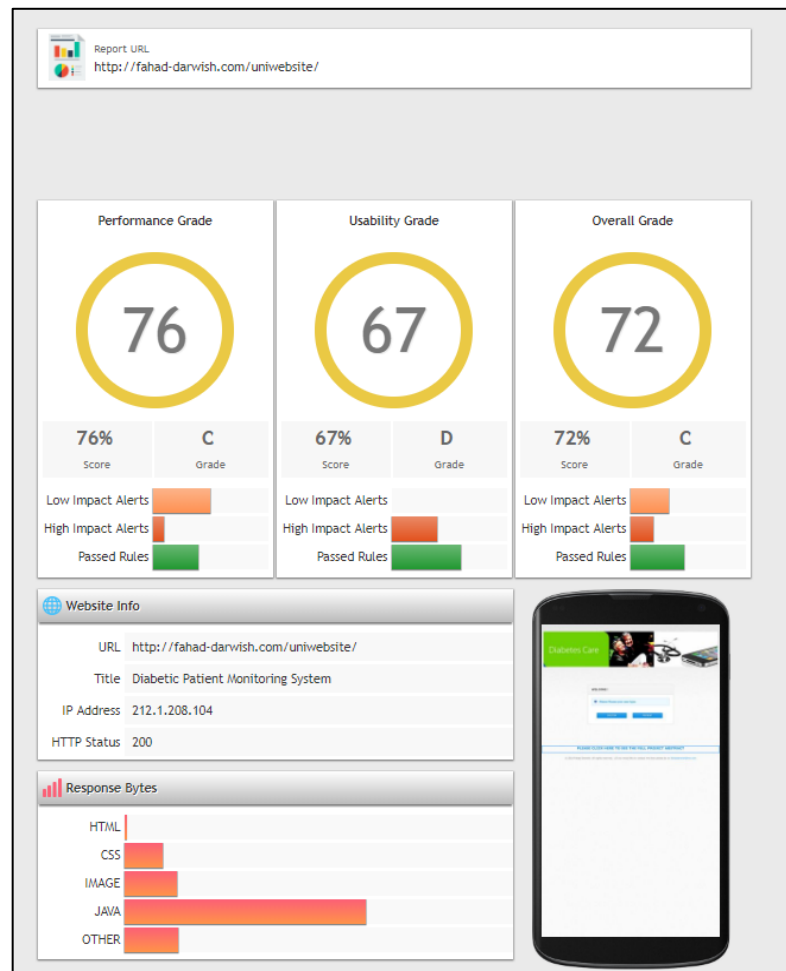
Figure 81 – The Test Results by website-grader.com

As shown above the initial results for the e-Toolkit was an overall grade of 72% for performance and usability.

2. The other tests were carried out as part of the evaluation where patients and doctors completed a website test form. See Appendix I for more details.

3. The final testing was achieved using Acceptance Testing – which is a test that determines whether the e-Toolkit had achieved its requirements or not. It includes running a number of "pass or fail" tests on the final product (Brian Hambling, 2013). See Appendix H for example of the Acceptance Testing that was used.

6.3.7   Hosting

Hosting is a space where the e-toolkit will be stored and accessed via any device that is connected to the internet from anywhere in the world (Brown M, 2016). Hosting24.com was used as they claim 99.9% availability and are a secure host. The link to our e-toolkit is " http://fahad-darwish.com/uniwebsite/ ".

6.3.8   Publicity

For future work, if the e-Toolkit is accepted by medical professionals and diabetic users and the NHS, then it can be advertised via social communications, such as Twitter, Facebook, LinkedIn, and using Instant Messages such as WhatsApp. Furthermore, hosting24.com provides Search Engine Optimisation (SEO).

However, an introductory video of the e-Toolkit was created that was used in presentations & conferences, a screenshot of the video is shown on Figure 82. The video can be found on:

" www.fahad-darwish.com/uniwebsite/University_of_Salford_Video.mp4 "

or

" https://www.youtube.com/watch?v=SxGvnzSCxAA&t=6s "

Figure 82 – An introductory video to present the e-Toolkit

6.3.9   Website Security

6.3.9.1   Registration and longing process

Users wanting to have access to the e-Toolkit will have to register, the user will create a username and a password. The e-Toolkit will automatically check to see if this username and password is unique and not already been used. Email and passwords must be entered twice, this will ensure that the email and passwords entered are correct.

The Patients information will be stored on the backend database for authentication and login, this will be managed by an administrator.

6.3.9.2   Credential Recovery

Users can request reminder of their username or password in the event that they have forgotten their credentials, they can request a password recovery by entering the email address in which they registered with, an email will then be sent for the user to change their password to a new one.

6.3.9.3   Encryptions and Password Protection

The SHA (Secure Hash Algorithm) encryption is used so the passwords are not stored as a plain text, instead every password is encrypted. Therefore, the developer or anyone else does not have access to the user password.

SHA-256 is a type of algorithm used in cryptographic hash algorithm, which was created by the NSA (United States National Security Agency). This algorithm generates a 256-bit unique text size hash. Password hashing is highly used due to its one-way function, it can cannot be decrypted back. This is appropriately used for password validation, challenge hash authentication, digital signatures and anti-tamper (Yue-Long OU, 2010; Singhal, 2018; Aumasson, 2018).

An SHA-256 example is shown on figure 83 below.

patient_password

```
31e28ba0744955ad6e808d05462ed72d80dc12e8139ee30065b267e52
c71a595
```
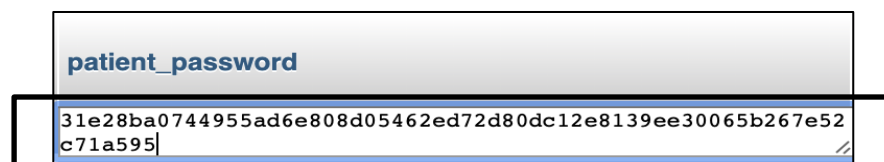
Figure 83 – SHA-256 Example of password encryption

6.3.10   E-Toolkit Review

A review of the website's plan indicates that the goals and objectives have been accomplished, this section explains the contents of the e-Toolkit.

6.3.10.1   e-Toolkit site Map

Sitemaps are a hierarchical graph it shows the construction of a website or application. as described by (MEARS, 2018) "They are used by user experience designers and information architects to define the taxonomy by grouping related content." Sitemap is an important part in web designing as it is a user centred process to make sure that the web site content is correct and in places where users would expect to find it. "They can also be used as a reference point for wireframes, functional specifications and content maps." (MEARS, 2018)

Figure 84 shows the sitemap of the proposed framework. The two layers are clearly shown. Layer one as the Diabetic Management System and Layer 2 as the SECI web portal.
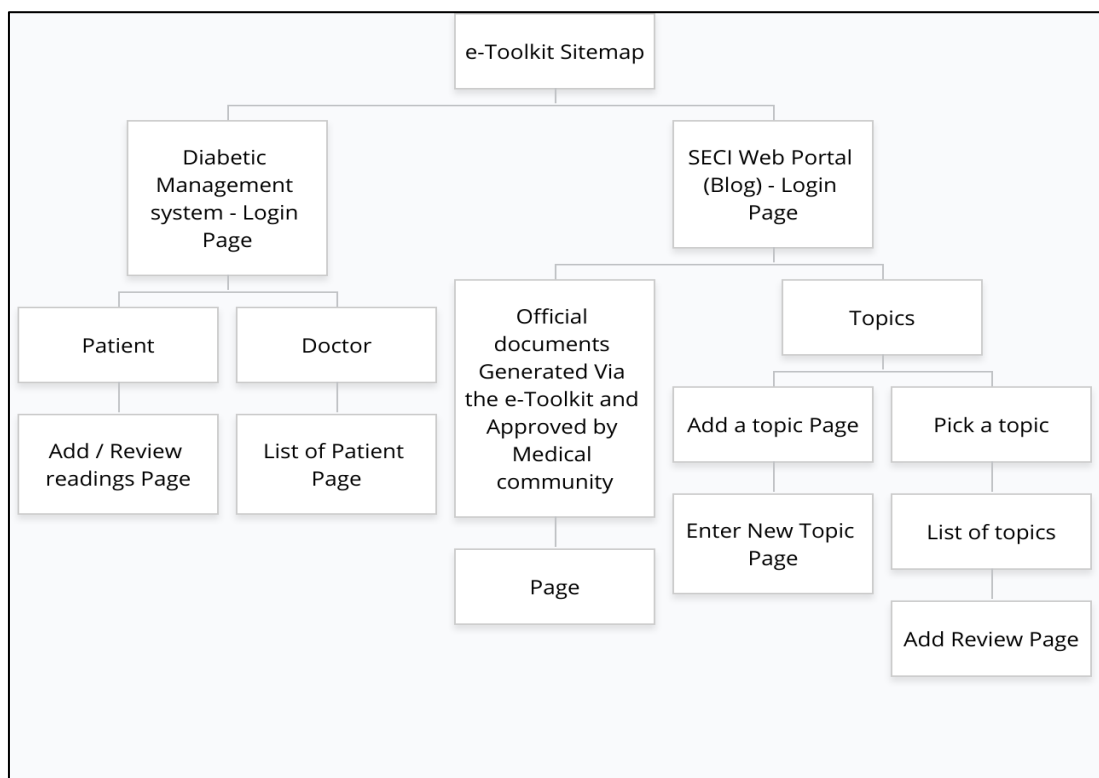
Figure 84 – The proposed e-Toolkit sitemap

6.3.11 E-Toolkit Security

The e-Toolkit has many security aspects i.e. All Passwords are encrypted using the SHA-256 hashing function as discussed in section 6.3.9.3 and illustrated in figure 75, the use of other security measures is put in place such as SSL Certificates which is a digital computer file or small lines of code that can help verify and authenticate some details of the e-toolkit. SSL can also be used to encrypts the exchange of information between the website and the server, the information cannot be accessed or read by anyone but the intended recipient (symantec.com, 2019).

6.3.12 Functionality

This part of the e-Toolkit will serve as the communication hub for patients and medical health care professional to share data and communicate.

6.3.13 Landing Page (Diabetes Management System – Layer 1)

Figure 85 illustrate the initial landing page of the login page
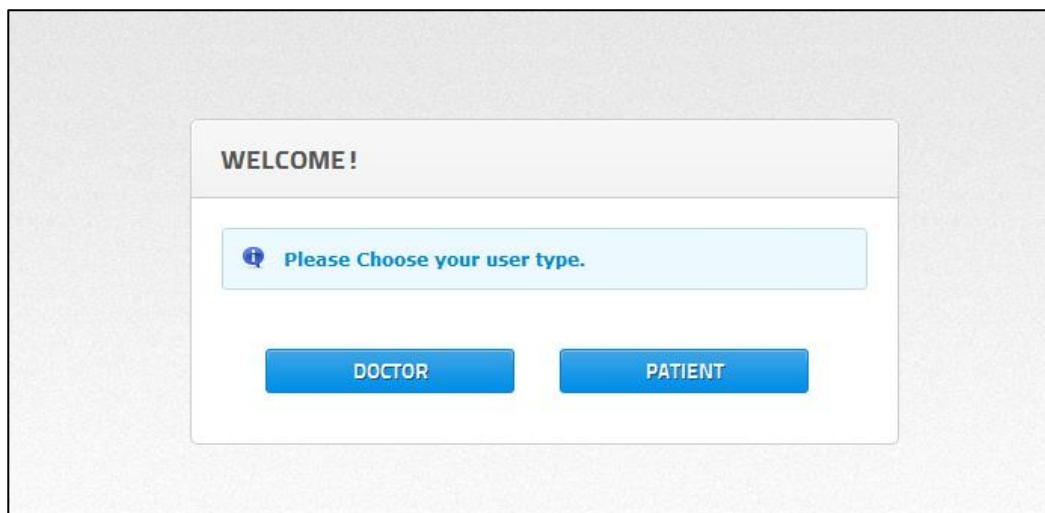
Figure 85 – Login Page

6.3.14 User verification

Anybody needing access to the site must first register. The system will automatically check if the user already exists. If the users forget their username or password, they could click on forgot username or password, an email to change their password will be sent as illustrated on figure 86 below.
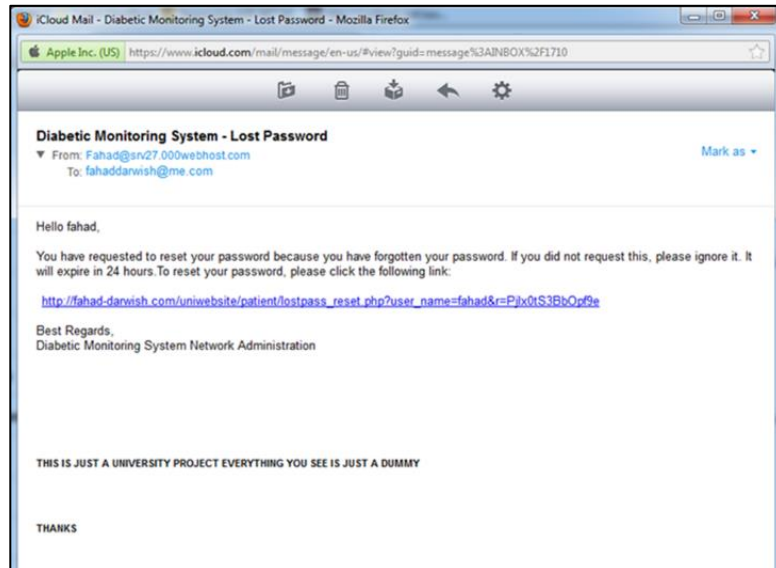
Figure 86 – Password Recovery

6.3.15 Diabetic Data (Doctors View)

This is the landing page that the medical health professionals will see.
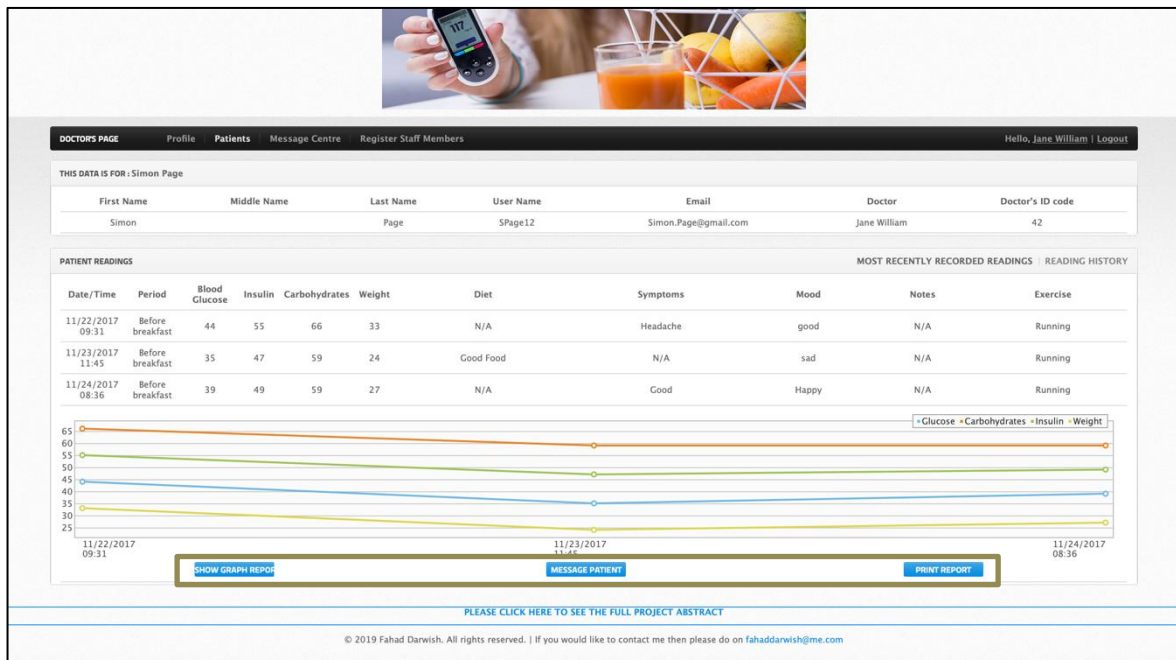
Figure 87 – Landing Page for Medical Health care Professionals

6.3.16 Registration Page (Doctors View)

This is where the health care professionals' details are recorded. They will have the option to assign a nurse to have access to specific patient details.
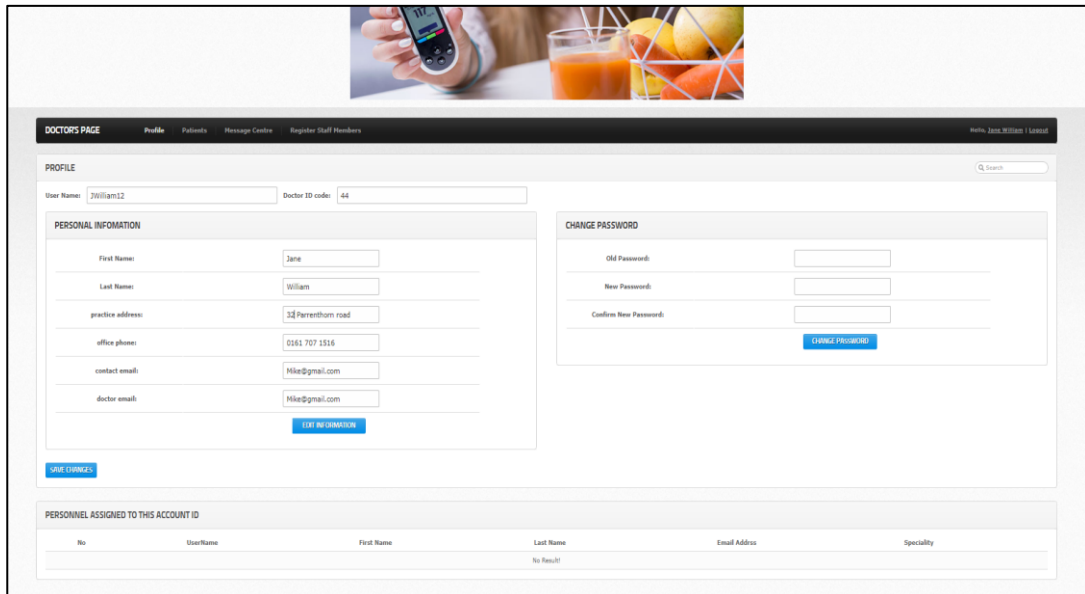
Figure 88 – Registration page for the Health Care Professionals.

6.3.17 Diabetic Login

Figure 89 illustrate the login page that diabetic patient will see after clicking on the 'patient' page, patients will get the chance to 'Sign Up' if they do not have username or password but they will need their health care provider to provide them with an ID in order to register and use the e-Toolkit.
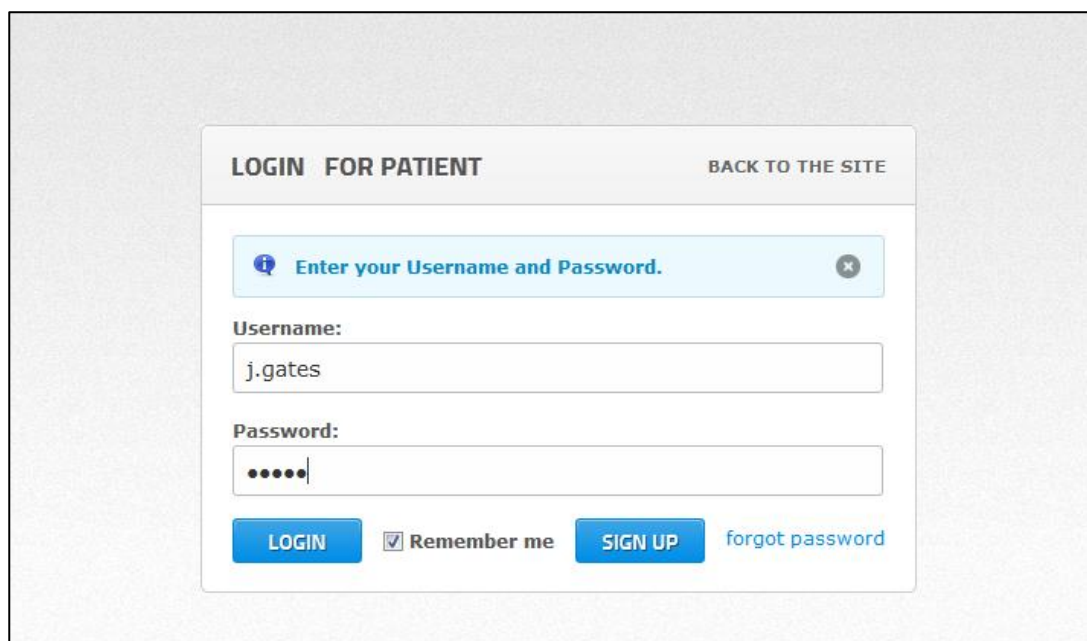
Figure 89 – Login Page for Diabetic Patient

6.3.18 Diabetic Patient Registration Page

Figure 90 Illustrate the registration Page in which diabetic will see when they register.

Figure 90 – Diabetic Patient Registration Page

6.3.19 Diabetic Profile

Figure 91 illustrate the page in which diabetic will see when they click on their profile, Diabetic will get to change their assigned doctor by changing the 'Doctor ID' they can also change their Password or any other details.

Figure 91 – Diabetic Profile page

6.3.20 Diabetic recording page

This is the page that diabetic will see upon clicking on the Daily Reading button at the top,

Diabetic will get to record in their weights, Blood glucose and more.
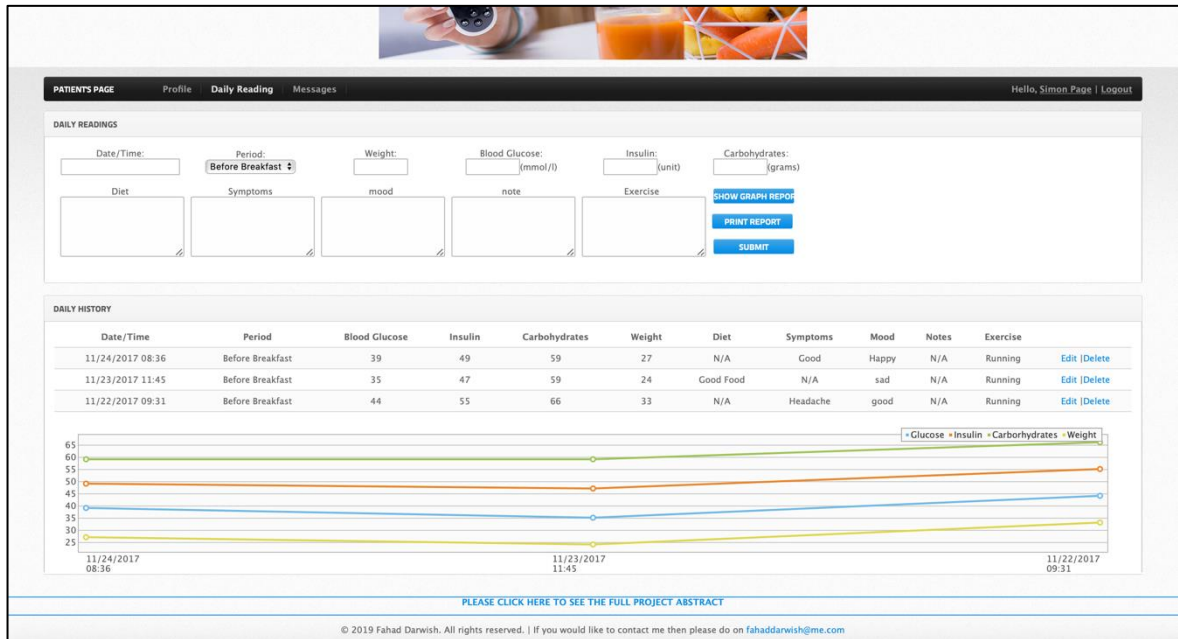
Figure 92 – Diabetic Recording Page

6.3.21 Diabetic Messaging page

Diabetic can send, receive message through the provided messaging system within the e-

Toolkit as illustrated in figure 93 below.
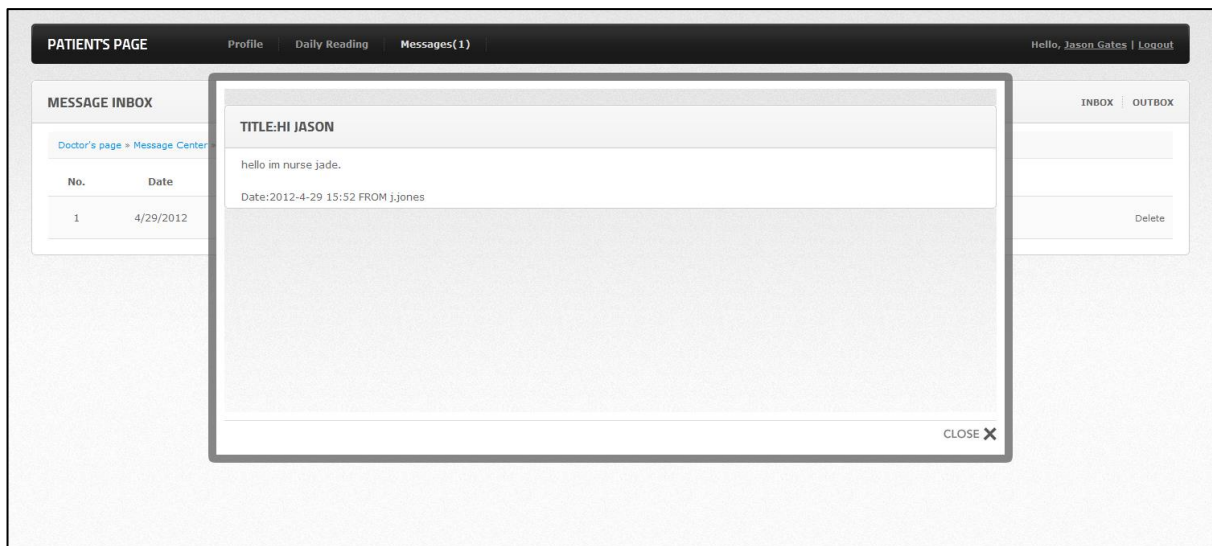
Figure 93 – Diabetic Messaging System Page

6.3.22 E-Toolkit Database Diagram
Figure 94 illustrate the Database Diagram for part of the e-Toolkit (first layer).
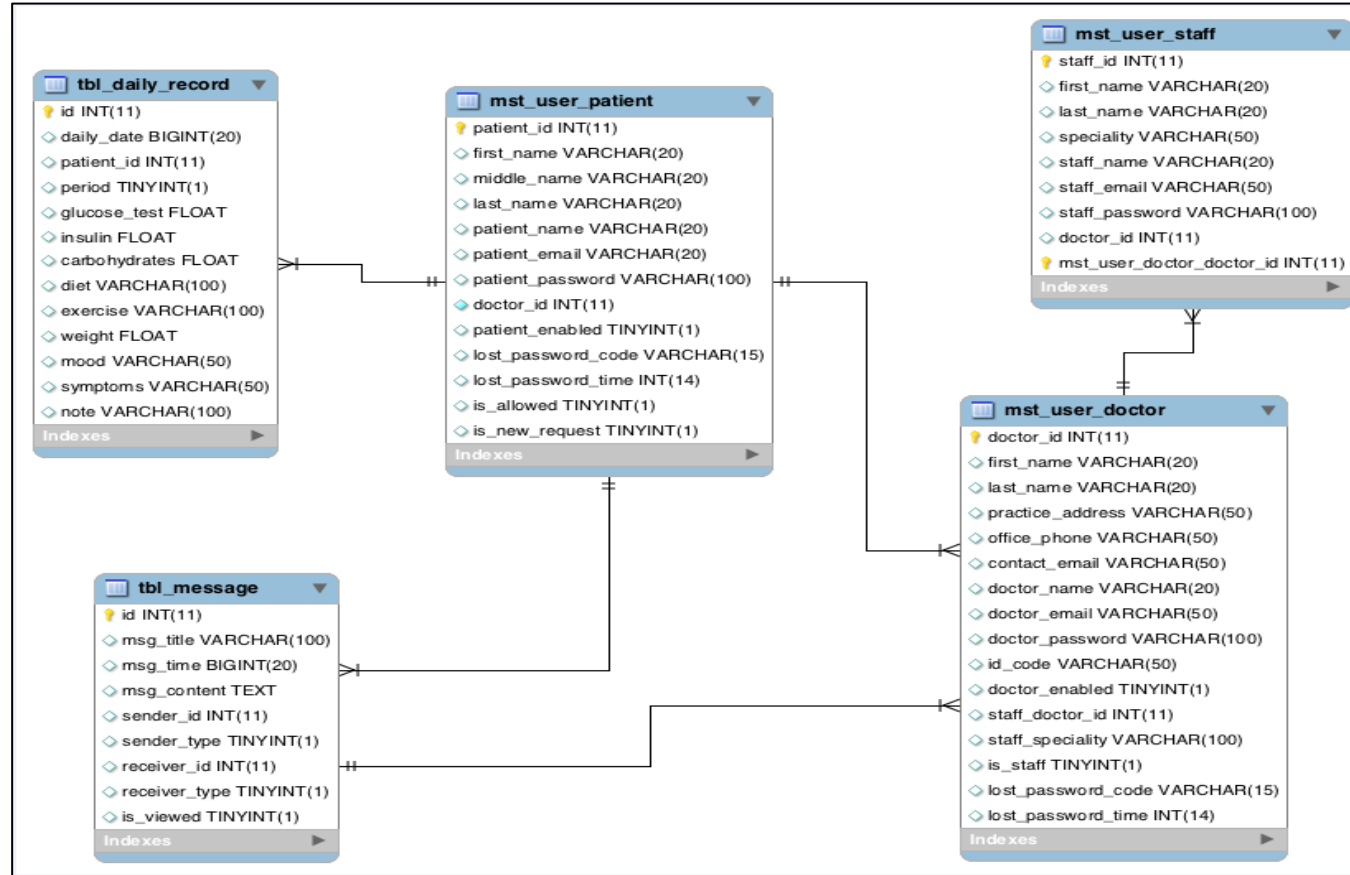
Figure 94 – MySQL Database Diagram

6.3.23 Data Collection for reviews on the e-Toolkit (SECI Portal - Layer 2)

As shown on Figure 90, patients could add reviews on a particular topic and the algorithm used will generate rules as shown on figure 95- 98.
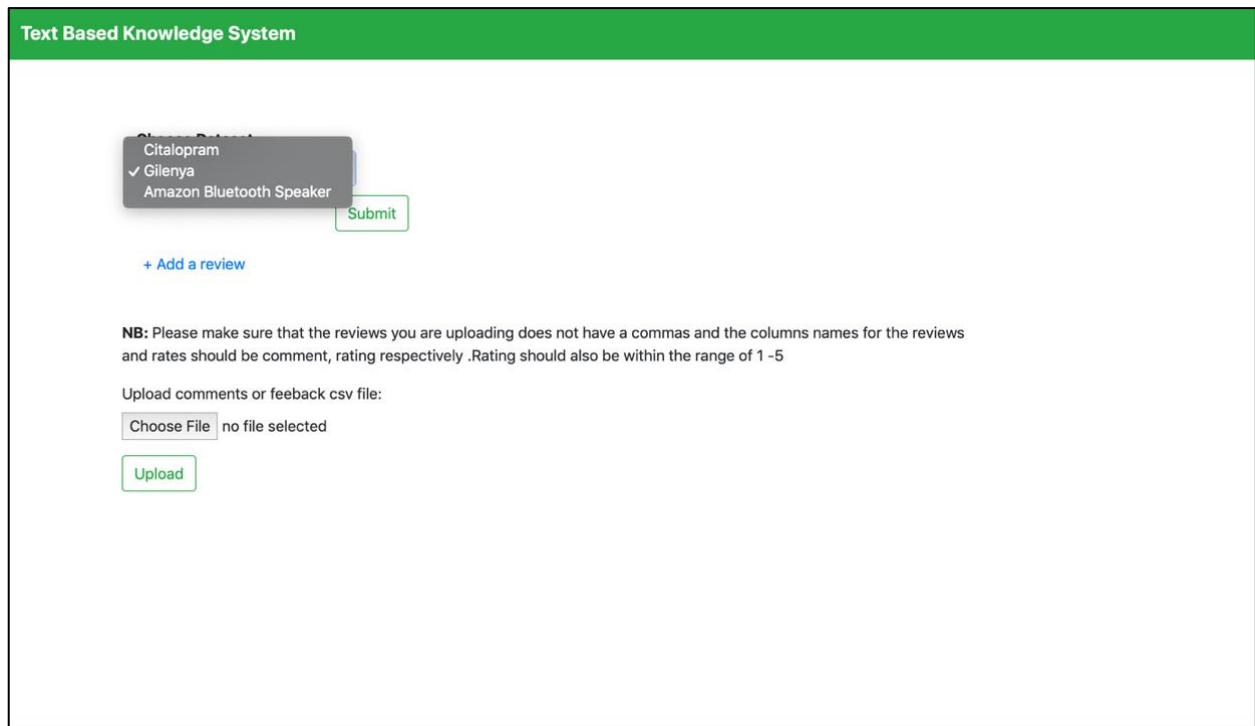
Figure 95– Patient could either select a topic of create a new topic.
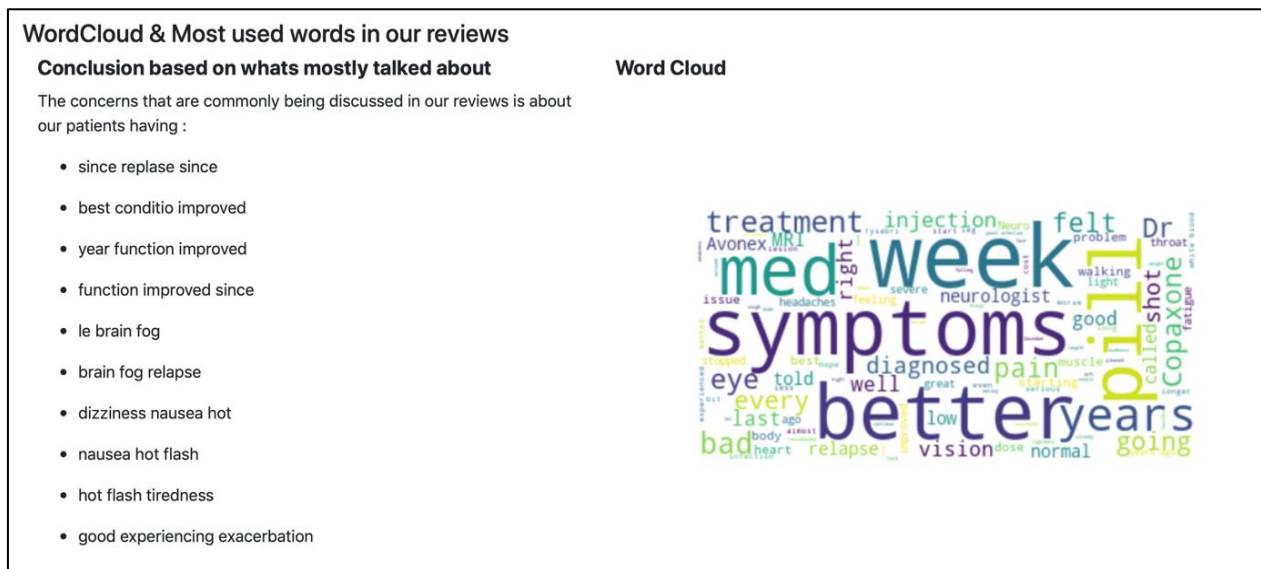
Figure 96 – The e-Toolkit displaying results of Gilenya dataset.
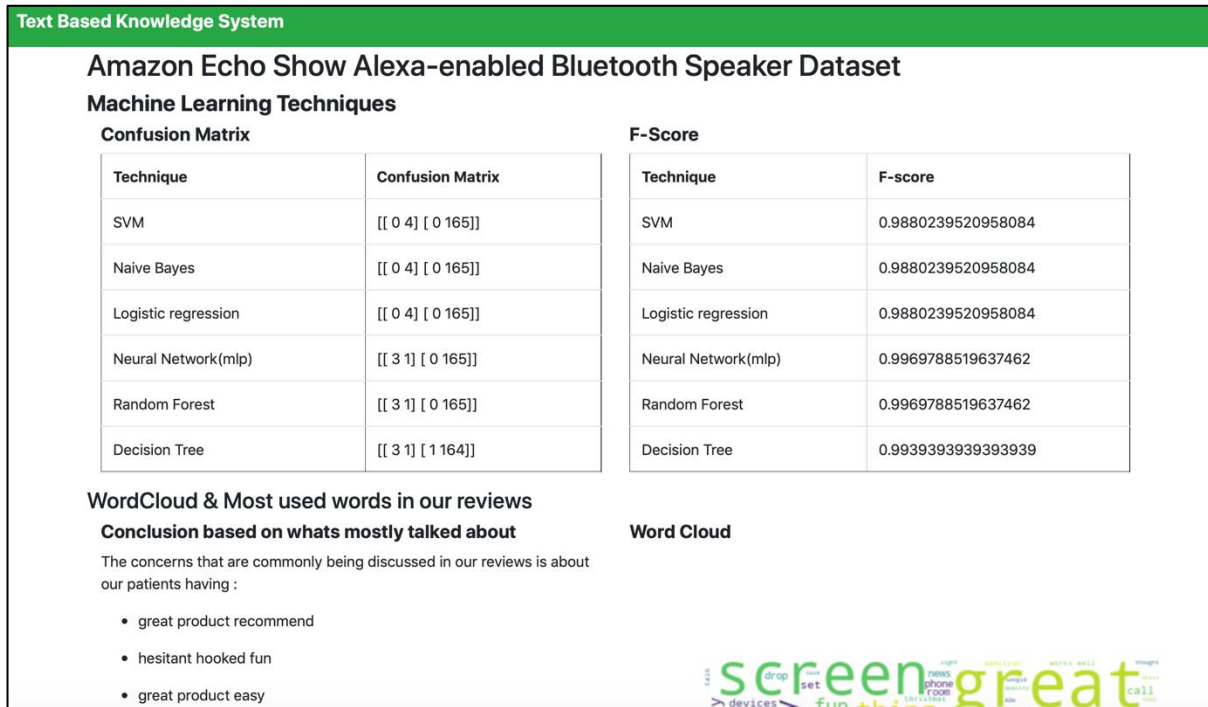
Figure 97 - Comparison of techniques on a specific dataset.
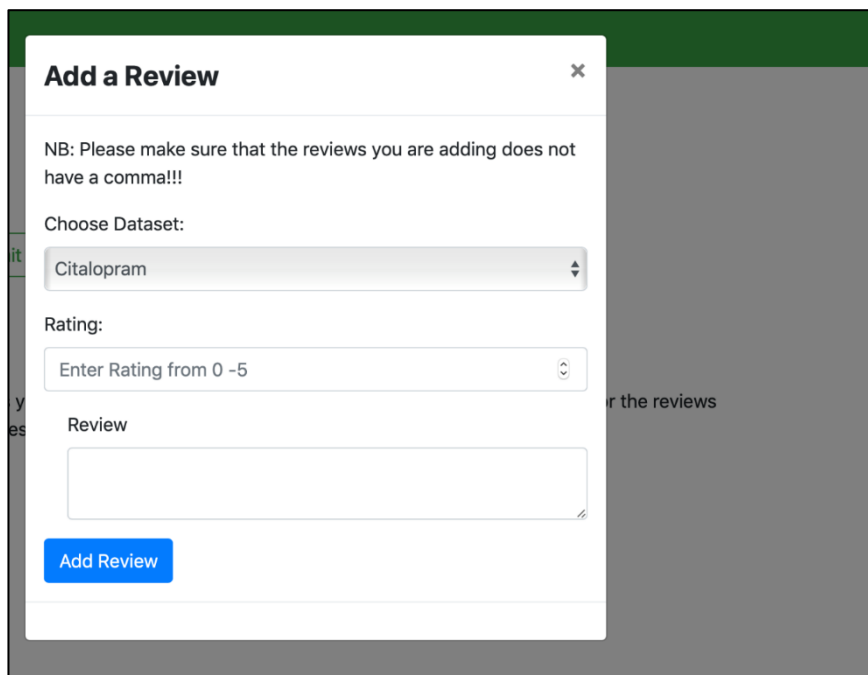
Figure 98– Adding reviews to be used on our system.

6.3.24 Goals achieved

Our results show that random forest and neural network techniques give a better result than other techniques as illustrated in the figures above. Moreover, the dataset is also an important factor in these cases. When considering the Bluetooth speaker, Amazon's dataset gives better results with all classifiers as it has a large number of reviews. This shows the more patients that use the e-Toolkit the more accurate the results will be which will in turn help diabetic patients.

6.4   Debugging

Component testing was performed to make sure that the system is working properly. The diagram below illustrates the process. It was obtained from the British Computer Society, (British Computer Society, 2018)
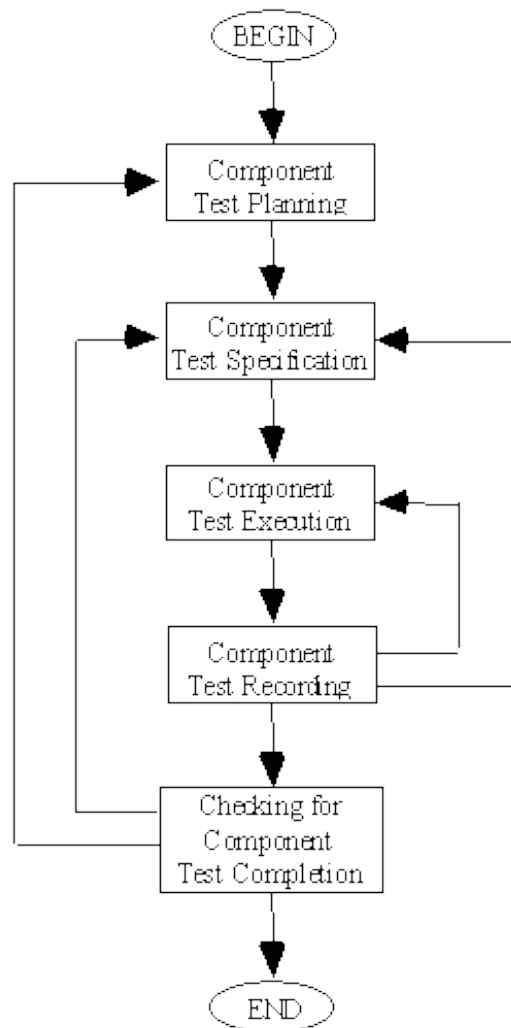
Figure 99 – Component Testing Diagram – Source: (British Computer Society, 2018).

6.5   The framework validation

One of the main objectives in this thesis is to validate the developed framework. As stated by (Christopher J. Roy, 2010) "To validate means to test the accuracy a model represents in the real world from the perspective of the intended users of the model". Validation was conducted via an online workshop and consisted of two parts. The first part focussed on our integrated framework and the second part was to evaluation of the e-toolkit which aimed to overcome the communication gap discussed in Chapter 1 & 2.

A workshop was conducted which was held using www.zoom.us, which is a website that delivers remote conferencing services using cloud computing including video conferencing, online chat, meetings and mobile teamworking (Maldow, 2019).

The workshop was carried out with eight participants who partook in both framework validation sessions. The eight participants were divided into four different backgrounds: two were medical doctors, two were pharmacists, two were diabetic patients and two were IT professionals. The online workshop took approximately an hour to complete and was split into the sections outlined below.

1. A presentation was given discussing the aims and objectives of this research and explaining the framework. The framework and the e-Toolkit were presented to participants with a description of all the components and capabilities (See Figure 100).
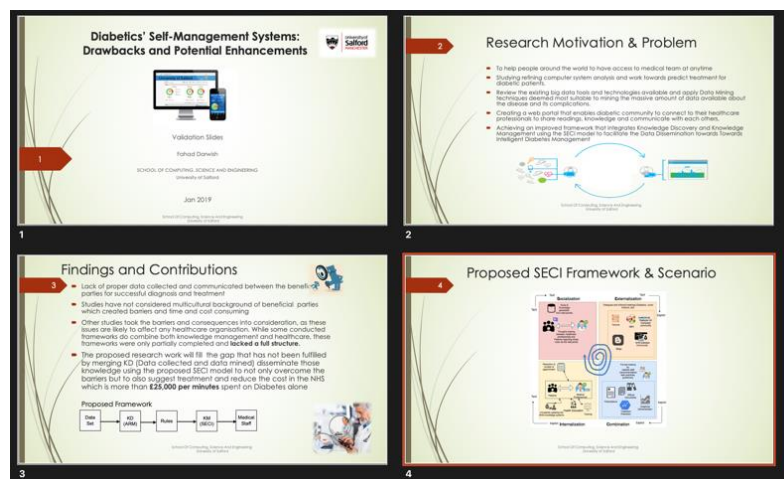
Figure 100 – Validation presentation PowerPoint slides

2. Findings such as medication side effects and other datasets were discussed, to verify the proposed SECI model.

   - Participants were directly involved in the production of a document detailing common side effects that were highlighted from the e-toolkit and findings.

3. A demonstration of the e-Toolkit was presented to determine if it had achieved its proposed requirements or not.

4. Discussion session with participants

5. Many validation forms were given to participants. (for sample forms see Appendix J & I).

To validate the e-toolkit from a technical point of view a questionnaire was carried out. This was completed the patient and medical staff as seen in appendix I.

6.5.1 e-Toolkit Validation

All participants stated they were able to access the site on both phones/tablets. However, the majority found quick and easy access on their phones, with one participant (IT professional) stating they felt it had the potential to be a great digital logbook for all to access.

Comments received for elements that didn't work as well on the e-Toolkit have been subsequently fixed. According to the pharmacists, they really liked the features where the patient is able to openly discuss topics and where concerns would be flagged up automatically. Specifically, in regard to medicine issues as that could help both patients and pharmacists by offering a faster way to find answers. The diabetic patients believed that the e-Toolkit was a good platform for all patients with diabetes to come forth as a community to support one another in managing their own diabetes. Another well liked feature was the freedom that comes with using the e-toolkit.

In particular the fact that they could now travel for a few months without having to worry about meeting their doctors as this e-Toolkit will automatically inform and update their health provider. Patients and IT flagged the issue of data protection and who might have access to data submitted to the e-toolkit. The IT team asked about the security measures being undertaken to ensure that everything is protected.
The researcher reassured the participants by highlighting the fact that everything is stored in a secured server and that passwords are encrypted from both ends (server and client ends).

It was also made clear at the beginning of the study who would have access to this data and consent was taken prior to participation.

6.5.2   Validation of Layer 2(SECI Model)

The validation of the knowledge discovery capability with our integrated framework is achieved through asking the relevant participants, specifically medical professionals and the diabetic patients, questions with regards to the e-Toolkit. Firstly, both doctors and one diabetic patient agreed that the e-Toolkit had met its purpose. When asked about the usefulness of the e-Toolkit, both diabetics agreed that it was of great tool for them to utilise, whereas one doctor disagreed on the usefulness due to the time it might take for doctors to train on the e-Toolkit. It was explained that the e-toolkit can be accessed by any authorised member of the multi-disciplinary healthcare team such as nurses and pharmacists working in the practice. The e-toolkit is designed to be user-friendly and easy to read, so training time can be kept to a minimum.

6.6   Chapter Summary

This chapter has explained in detail the development of the e-Toolkit from start to finish, including an illustration of the final e-Toolkit in the form of screenshots. The validation of the framework and the e-Toolkit is also discussed in this chapter. These two parts were part of the aims and objectives (5 & 6) which are:

- To build an e-Toolkit to support diabetes education and self-management in the UK.
- To validate the developed framework and the e-Toolkit.

Having discussed the implementation, testing and validation of the framework and the e-Toolkit the next chapter will conclude the research and findings as well as a review the research contribution and discuss opportunities for future work.

CHAPTER SEVEN: CONCLUSIONS AND FUTURE WORK

7.1   Introduction

This chapter presents the conclusions and future work derived from this thesis.   The conclusions will consist of reviewing the work accomplished and the objectives achieved throughout this thesis. The first part of this chapter is the conclusions, followed by the Future work. The conclusions will consist of introducing and summarising the current research, followed by reviews of the aims and objects. The second part of the present chapter highlights the research contributions followed by the limitations faced whilst conducting this study. The final part of the current chapter highlights opportunities for future research work and discusses any possible limitations.

7.2   Conclusion

Knowledge management within the healthcare sector in the UK faces a number of elements that can hinder the application of the two fields. Such drawbacks result in insufficient practice of diabetes self-management and education to support the fast-growing number of diabetic patients, as well as healthcare professionals. The current research has developed a framework which integrates KM and KD to support the diabetics.

This integrated framework provides a forum to include not only human knowledge and experience but also knowledge discovered from data. The developed SECI model has been applied through the proposed e-Toolkit to support diabetes self- management. Additionally, the framework is ambitious in its approach as it promotes the merits of KM in the healthcare sector by including a number of these tools familiar to patients. this will help to address the barriers discussed in chapter 1.

The framework, eventually, aims to encourage the practice of self-management and education worldwide in terms of the prevalence of diabetes mellitus. The first chapter has indicated several objectives which has achieved in this research thesis. Several objectives were accomplished in previous chapters. For example, the creation of a universally accessible e-Toolkit has been successfully researched by comparing other applications in the healthcare sector. This was described in Chapter 2 and illustrated in Chapter 6 and identified where the needs are not being met by getting responses from both the healthcare professionals and the diabetics patients as clearly shown in Chapter 2.

In creating this e-Toolkit, the main goal was to create something that is easy to use and very simple for diabetics to record their readings. In the survey, the patients expressed their feelings about the tools that they used previously, citing a common theme that their current system was not user friendly or intuitive, leading to the diabetic community being deterred from using such tools.

Additionally, the medical health care and diabetic patients respondents have highlighted that user friendly is a must for the proposed e-Toolkit and easy to navigate, which this research has taken into account and achieved.

The e-Toolkit is very cleaned designed, straight-forward and simple, with a clean layout where everything is there and the user does not need to search for a feature or a tool.
The minimalistic approach was intentional and has served its goal well. The e-Toolkit has achieved all its basic requirements (the must haves) and most of its secondary goals (the should haves). There is however, much room for refining the design and expanding the features.

The survey shown in appendix H was also given to evaluators to ensure that all parts of the e-Toolkit were accessed and inspected. The overall outcome was that the e-Toolkit has accomplished what it set out to accomplish.

7.2.1   Review of Research Aims and Objectives

Objective	Method of investigation	Main outcome	Chapter
1. **A) To conduct a literature review of KM and its model as well as identifying issues within diabetes in the UK**	Literature review from published academic documents, industry reports thesis	Gaps within the KM in the UK were discovered and were fulfilled by the proposed framework	Chapter 2, 3 & 5
1. **B) To conduct a literature review of knowledge discovery**	Literature review from published academic documents, industry reports thesis	Gaps within the knowledge Discovery were discovered and fulfilled by the proposed framework	Chapter 2, 3 & 5
1. **C) To conduct a literature review of SECI models and its applications**	Literature review from published academic documents, industry reports thesis	How the SECI model can be used to support the proposed model.	Chapter 2 &3
1. **D) To conduct a literature review of Diabetes Mellitus and its complications and prevalence globally and in United Kingdom.**	Survey conducted with 60 Participant who are patients and 40 Doctors	Paper published (Please see appendix C)	Chapter 2
2. **To carry out a survey in order to establish limitation of the existing diabetic monitoring systems.**	Finding lead to creating an e-Toolkit that could help diabetic and medical health care	Paper published (Please see appendix C)	Chapter 4

3. **To Develop a framework to integrate KM and KD via the developed SECI Model to support the e-healthcare of diabetes mellitus patients and the NHS professionals in the UK.**	Integration of KM and KD.	Integrated framework of using the developed SECI model and KD via the use of Data Mining techniques to help diabetes better manage themselves.	Chapter 4, 5 & 6
4. **To carry out a Data Mining analysis of reviews data to find the most common issues within certain medication to validate our framework idea, as well as applying this into non-medical dataset such as Amazon**	The use of DM techniques, Python programming and various software were used to achieve this.	Sentiment Analysis techniques and other data mining tools used to potentially find side effects of medications and give suggestions of treatment options for the healthcare professionals according to the data which the patients have input.	Chapter 4, 5 & 6
5. **To build an e-Toolkit to support diabetes education and self-management in the UK.**	To plan, design and test the web portal through reliable methods.	e-Toolkit was planned and developed using multiple web development lifecycle such as Diffily 2006.	Chapter 5 & 6

6. To validate the developed framework and the e-Toolkit.	An online workshop using zoom.us was conducted with 8 participants 2 doctors 2 pharmacist 2 diabetic patients and 2 IT professionals To validate the framework and its components as well as the application of the SECI model on the e-Toolkit	Positive feedback was the overall results of the validity of the framework. Evlaution of the e-Toolkit was also ahcieved via multiple methods such as acceptance testing	Chapters 6,7 & 8

Table 9 – Overall research objectives and outcome

As well as the points mentioned in Table 9 the following e-Toolkit main aims that were also achieved:

- Creation and making of universally accessible e-Toolkit that will help the diabetic community in the UK and globally. The e-Toolkit will have many purposes offering an area which is simple to use so that diabetic patients are able to document risk factor data (such as glucose levels etc). Diabetics patients will be able to access this e-Toolkit from anywhere in the world using their smart phones, tablet, PC or any other devices which have internet access.

- The ability to communicate with the patient's appointed medical professional via a direct messaging system, with measures in place for regular monitoring. Therefore, the e-Toolkit is able to be incorporated further into the patient's day-to-day routine in real time.

- The ability to display data graphs and charts; high risk patients alerts to medical professionals and the ability to share and store part of the patient's medical data (e.g. test results)

- The capability to have a blog feature, where diabetic patients can discuss and review medications that medical professionals can access. Reports will also be made of both the successful and ineffective types of medication prescribed.

7.2.2   Review of Research Contributions

The current research has produced a number of academic contributions. These contributions are presented as follows:

- A novel framework that integrates KM and KD within the developed SECI model via the e-Toolkit to help diabetics in the UK to better communicate with medical professionals. The novel framework integrates KM and KD. KD applies sentiment analysis techniques and uses other data mining tools to potentially find side effects of medications and provide suggestions of treatment options for the healthcare professionals according to the data the patients have input. It also aims to provide patients around the globe who have diabetes and doctors with the relevant solutions in relation to overcoming these complications as well as better communication between those parties.

- The integration of KM and KD is implemented in a new domain. That is, the proposed framework aims to overcome difficulties and problems of diabetics around the world as well as data mining the issues that they are facing from the data they input to discover new knowledge and to potentially help reduce the number of patients with diabetes.

- The novelty of the implementation of the SECI model to disseminate useful knowledge related to diabetes self-management and education. Despite the usefulness of the SECI model in disseminating useful knowledge in different domains, the model lacks in its implementation to support people with diabetes to self-manage and educate them about their disease. The current study applied the developed SECI model to discover and disseminate useful diabetes self-management and education recommendations and showed a promising impact in empowering diabetic patients.

- The generation of an open-source e-Toolkit for future diabetes or other domains of research. The current research indicated a deficiency of open-source e-Toolkits. It is planned to create an open source dataset gathered through diabetic visitors to the web portal via an optional participant on the knowledge gathered and will be published in the portal periodically.

- The development of an e-Toolkit to support the community of diabetes mellitus patients using the SECI model. Despite the high prevalence of diabetes mellitus within the NHS, there are still deficiencies regarding the use of web technologies to increase the awareness of the diabetes community, Furthermore, there isn't a single

available website or an e-Toolkit kit that proposes the idea that this thesis proposes where by one location a patient is able to record all their vital readings and instantly share those with their health care providers as well as educating patients using an effective KM approach. Internet technology is a valuable tool for providing educational content. The application of the SECI model in the web environment provides an effective solution to disseminate useful knowledge to support people with diabetes mellitus, especially for those who live in remote areas.

7.3   Opportunities for Future Research

The discussion in the current research has identified five areas for future research. These are illustrated as follows:

- The proposed framework was applied just to diabetes mellitus. One future research opportunity would be the application of the framework to overcome other diseases such as cardiovascular disease or respiratory disorders or other fields. Potential domains include education, military, industry and engineering, where ICT has been utilised for a period of time. Moreover, the application of the proposed framework in one of these domains offers some promising feasible opportunities as the challenges facing KM and the SECI model can often be the same.

- As mentioned in chapter 1 and in the abstract, there are many other diseases where this framework could be applied. They include cardiovascular diseases that require consistent self-management. The application of the framework to extract knowledge encountered by cardiovascular disease patients and to disseminate guidelines and best practice can provide a useful initiative to tackle these diseases.

- The current research focused on sentiment analysis of knowledge discovery from the e-Toolkit to extract knowledge related to information input by the patient. The utilisation of other KD approaches such as text mining, web mining and link mining offer an opportunity for future research. The application of these KD approaches can also produce a valuable contribution to the healthcare domain and/or be integrated with KM.

- The current system allows patients to access the e-toolkit on mobiles and tablet devices as well as laptops and PCs. For future improvements, a topic of ease of use to the older generation could be added, this is for those who might not have the best understanding of technology and systems. The medical health care team can show a hands-on introductory tutorial to help them with the use of the system.

8   REFERENCES

8.1   Reference List

A. Sallaberry, N. P. (2011). Sequential patterns mining and gene sequence visualization to discover novelty from microarray data. *Journal of Biomedical Informatics*, 760-774.

A.D.A.M. 2. (2011). Type 1 diabetes. In Z. Eckman (Ed.), *A.D.A.M. Medical Encyclopedia.* Elizabeth, New Jersey, USA: A.D.A.M. Medical Encyclopedia.

A.D.A.M. 3. (2018). Type 2 diabetes. In *A.D.A.M. Medical Encyclopedia.* Elizabeth, Ney Jersey, USA.

A.D.A.M. 4. (2017). Gestational diabetes. In *A.D.A.M. Medical Encyclopedia.*

A.D.A.M. Medical Encyclopedia. (2011). Diabetes. In E. E. Zieve (Ed.), *A.D.A.M. Medical Encyclopedia.* Elizabeth, new Jersey, USA: A.D.A.M.

AAFP, F. L. (2010, Oct). *Diabetes: Monitoring your blood sugar level.* Retrieved 2012, from FamilyDoctor.org: http://familydoctor.org/familydoctor/en/diseases-conditions/diabetes/treatment/monitoring-your-blood-sugar-level.html

accu-chek.co.uk. (2018, 10 02). *Accu-Chek Smart Pix device reader product support*. Retrieved from Accu-Chek Smart Pix: https://www.accu-chek.co.uk/help/data-management/smartpix

accu-chek.co.uk. (2018, 02 02). *Accu-Chek Smart Pix Management System*. Retrieved 2012, from accu-chek.co.uk: https://www.accu-chek.co.uk/gb/products/datadevices/smartpix.html

accu-chek.com.ph. (2018). *About Roche Diabetes Care*. Retrieved 01 02, 2019, from https://www.accu-chek.com.ph/about-roche-diabetes-care

Aidera. (n.d.). *GENERAL TERMS AND CONDITIONS ("GTC")* . Retrieved 2012, from Diasend.com: http://www.diasend.com/docs/gtc.pdf

AI-Magazine, F. U.-S. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine V7 (3)*, 37 – 54.

Alberti, A. A. (1999). *Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications.* Geneva: World Health Organisation.

allkm.com. (2018, 03 02). *Knowledge Creating process, SECI Model*. Retrieved from http://www.allkm.com: http://www.allkm.com/km-basics/knowledge-process.php

AL–MADI, M. A. (2017). *A NOVEL KNOWLEDGE MANAGEMENT FRAMEWORK FOR MANAGING INFORMATION OVERLOAD IN THE DIABETES CLINICS OF THE HOSPITALS IN JORDAN.* Staffordshire: University of Staffordshire.

Almuayqil, S. A. (2017). Integrated Framework of Knowledge Management and Knowledge Discovery to Support E-Health for Saudi Diabetic Patients. *International Journal of*

Advanced Life Sciences (http://eprints.staffs.ac.uk/3937/ - https://file.scirp.org/Html/2-8701433_78665.htm), 111-121.

amazon.co.uk. (2018, 12 12). *Amazon - Echo Show, Smart speaker and 7-inch screen with Alexa*. Retrieved from amazon.co.uk: https://www.amazon.co.uk/Introducing-Echo-Show-White-Now-Alexa-Can-Show-You-Things/dp/B01BLTQ4XU

American Academy of Family Physicians, F. L. (2018, Feb). *Diabetes: Monitoring your blood sugar level.* Retrieved 2012, from FamilyDoctor.org: http://familydoctor.org/familydoctor/en/diseases-conditions/diabetes/treatment/monitoring-your-blood-sugar-level.html

American Diabetes Association. (2012). *Checking for Ketones*. Retrieved 2012, from American Diabetes Association: http://www.diabetes.org/living-with-diabetes/treatment-and-care/blood-glucose-control/checking-for-ketones.html

American Diabetes Association. (2012). *What is Gestational Diabetes?* Retrieved 2012, from American Diabetes Association: http://www.diabetes.org/diabetes-basics/gestational/what-is-gestational-diabetes.html

American Diabetes Association. (2018, 04 02). *American Diabetes Association*. Retrieved from http://www.diabetes.org: http://www.diabetes.org/diabetes-basics/

anaconda.org. (2019, 01 03). *spyder*. Retrieved from anaconda.org: https://anaconda.org/anaconda/spyder

Anderson, A. B.-M. (1992). Empowerment: An Idea Whose Time Has Come in Diabetes Education. *The Diabetes Educator*, 37-41.

Angela M. Pazzaglia, E. T. (2016). *Survey methods for educators: Analysis and reporting of survey data .* Education Development Center, Inc. .

Angela M. Pazzaglia, E. T. (2016). *Survey methods for educators: Analysis and reporting of survey data (part 3 of 3).* Washington, DC: IES and the U.S. Department of Education.

Antano Solar John. (2009). *MODX WEB DEVELOPMENT .* Brimingham : Packt Publishing Ltd.

Apple. (2012). *iPhone*. Retrieved 2012, from apple.com: http://www.apple.com/iphone/specs.html

Aumasson, J.-P. (2018). *Serious Cryptography: A Practical Introduction to Modern Encryption.* San Francisco - CA: William Pollock.

Bachle, M. K. (2007). Ruby on Rails. *Software, IEEE*, 105 - 108.

Bakharia, A. (2007). *Ruby on Rails Power!: The Comprehensive Guide.* Technology PTR.

Basta, A. (2014). Hands on Projects. In *Computer Security and Penestration Testing* (p. 208). Stamford - USA: Cengage Learning.

BBC. (2012, April 24). *Diabetes: Most NHS Costs Wasteful, says Diabetic Medicine*. Retrieved April 24, 2012, from BBC: http://www.bbc.co.uk/news/health-17829012

BBC. (2017, April 24). *Diabetes: Most NHS Costs Wasteful, says Diabetic Medicine*. Retrieved April 24, 2012, from BBC: http://www.bbc.co.uk/news/health-17829012

BCS. (2001). *Standard for Software Componenet Testing.* Shrivenham, Swindon, Wilts: British Computer Society Specialist Interest Group in Software Testing.

Bloomberg.com. (2018, 05 01). *Diasend AB: Private comapny Information*. Retrieved from Bloomberg.com: https://www.bloomberg.com/research/stocks/private/snapshot.asp?privcapid=34419030

Blumenthal, B. C. (2018). *Division of Nephrology: Diabetic Mellitus.* Retrieved 2012, from Medical College of Wisconsin: https://www.mcw.edu/Nephrology/Patient-Information.htm

Boulton, V. R.-T. (2005). The global burden of diabetic foot disease. *The Lancet*, 1719-1724.

Bramble., J. R. (2003). Uncontrolled Chronic Disease: Patient Non-Compliance or Clinical Mismanagement? *Disease Management*.

Brian Hambling, P. v. (2013). *User Acceptance Testing: A Step-by-step Guide.* BCS.

British Computer Society. (2018, 02 02). *The Test Managment Guide*. Retrieved from http://www.ruleworks.co.uk: http://www.ruleworks.co.uk/testguide/BS7925-2.htm

Britton. (2009). Choosing a Programming Language. *Skyscrapr*.

Brixius, N. (2018, 03 03). *The Logit and Sigmoid Functions*. Retrieved from Nathan Brixius: https://nathanbrixius.wordpress.com/2016/06/04/functions-i-have-known-logit-and-sigmoid/

Brown M. (2016). *Web Hosting on Aws: A Step by Step Guide - Meg Brown.* CreateSpace Independent Publishing Platform.

Brownlee, J. (2017, October 18). *machinelearningmastery.com*. Retrieved from How to Clean Text for Machine Learning with Python: https://machinelearningmastery.com/clean-text-machine-learning-python/

Bryman. (2012). Social Research Methods. In A. Bryman, *Social Research Methods* (pp. 50-62). Oxford: Oxford University Press; 4th edition.

Burd. (2007). *Ruby on Rails for Dummies.* Indianapolis Indiana USA: Wiley Publishing.

Burd, B. (2007). *Ruby on Rails For Dummies.*

Canadian Diabetes Association. (2008). Canadian Journal of Diabetes. *Canadian Journal of Diabetes*.

CBE, D. a. (2019, 03 03). *GPs to 'prescribe' apps for patients*. Retrieved Jan 11, 2018, from www.gov.uk: https://www.gov.uk/government/news/gps-to-prescribe-apps-for-patients

Chandra Wiranto, J. C. (2019, 03 15). *Knowledge Management - SECI MODEL*. Retrieved from youtube.com: https://www.youtube.com/watch?v=X63l7RwOgtE

Check J., Schutt R. K. (2012). *Research methods in education.* Thousand Oaks, CA: Sage Publications.

Chen. (2012, March 15). *Choosing a Database Platform*. Retrieved 2012, from Database Journal: http://www.databasejournal.com/sqletc/article.php/3869736/Choosing-a-Database-Platform.htm

Chen, Y. L. (2012). Key Factors Affecting Healthcare Professionals to Adopt Knowledge Management: The Case of Infection Control Departments of Taiwanese Hospitals. Expert Systems with Applications. *Elsevier.* , 450-457.

Chikh, A. (2011). A Knowledge Management Framework in Software Requirements Engineering Based on the SECI Model. *Journal of Software Engineering and Applications*, 718-728.

Chin et al., 2. M. (2001). Barriers to providing diabetes care in community health centers. *Diabetes Care*, 268-274.

Christopher J. Roy, W. L. (2010). *Verification and Validation in Scientific Computing.* Cambridge: Cambridge University Press .

CIA World FactBook. (2018, 05 05). *The world FactBook - Central Intelligence Agency*. Retrieved from www.cia.gov/: https://www.cia.gov/library/publications/the-world-factbook/fields/2018.html

Cihakova. (2001, 9 10). *Type 1 Diabetes Mellitus.* Retrieved 2012, from Johns Hopkins Medical Institutions: http://autoimmune.pathology.jhmi.edu/diseases.cfm?systemID=3&DiseaseID=23

Coleman, V. (1998, April). A quality software process for rapid application development. *Quality Software Journal, 7*, 107-122.

Creswell, J. W. . (2013). Qualitative Inquiry & Research Design: Choosing among Five Approaches (3rd ed.).

Currie, G. ,. (2008). THE LIMITS OF KNOWLEDGE MANAGEMENT FOR UK PUBLIC SERVICES MODERNIZATION: THE CASE OF PATIENT SAFETY AND SERVICE QUALITY. *Public Administration*, 363-385.

Dalkir, K. (2011). *Knowledge Management in Theory and Practice, Second Editionpub.* Cambridge, Mass.: MIT Press.

Daniel T. Larose, C. D. (2015). *Data Mining and Predictive Analytics.* Canada: John Wiley & Sons, Inc.

Darwish, F. (2012). *Diabetes Management Website.* Huddersfield - United Kingdom: University of Huddersfield.

Davidov, B. C. (2009). Diabetic retinopathy and health0related quality of life. *Graef's Archieve for Clinical and Experimental Ophthalmology*, 267-272.

Davydova, O. (2019, 01 02). *Text Preprocessing in Python: Steps, Tools, and Examples*. Retrieved from Text Preprocessing in Python: Steps, Tools, and Examples: https://medium.com/@datamonsters/text-preprocessing-in-python-steps-tools-and-examples-bf025f872908

Dawson, C. (2009). The stage-wise and classical waterfall models. In Dawson, *Projects in Computing and Information Systems* (2nd ed., p. 119). Prentice Hall.

Diabetes drugs cost NHS nearly £1bn a year. (2016, August 03). Retrieved October 10, 2017, from www.theguardian.com: https://www.theguardian.com/society/2016/aug/03/diabetes-drugs-cost-nhs-1bn-a-year-presciption

Diabetes UK. (2012). *Diabetes iPhone Tracker*. Retrieved 2012, from Diabetes UK: http://www.diabetes.org.uk/How_we_help/Diabetes-iPhone-Tracker-app/

Diabetes UK. (2018, 01 20). *Diabetes UK*. Retrieved from https://www.diabetes.org.uk: https://www.diabetes.org.uk/about_us/news_landing_page/nhs-spending-on-diabetes-to-reach-169-billion-by-2035

Diabetes UK. (2018, 02 13). *What is diabetes?* Retrieved 02 13, 2012, from diabetes.org.uk: http://www.diabetes.org.uk/Guide-to-diabetes/Introduction-to-diabetes/What_is_diabetes/

diabetes.co.uk. (2018, 02 03). *Type 1 Diabetes*. Retrieved from Diabetes.co.uk: https://www.diabetes.co.uk/type1-diabetes.html

Diabetes.co.uk. (2018b, 04 22). *Diasend HOME*. Retrieved from www.diabetes.co.uk: https://www.diabetes.co.uk/diabetic-products/diasend-home.html

diabetes.org. (2010, 12 02). *Diabetes in the UK 2010.* Retrieved 02 10, 2019, from https://www.diabetes.org.uk/resources-s3/2017-11/diabetes_in_the_uk_2010.pdf

diabetes.org. (2018, 05 11). *Your Health Care Team*. Retrieved from American Diabetes Association: http://www.diabetes.org/living-with-diabetes/treatment-and-care/whos-on-your-health-care-team/your-health-care-team.html

diabetes.org.uk. (2018). *Methods and criteria for diagnosing diabetes*. Retrieved 01 04, 2019, from https://www.diabetes.org.uk/professionals/position-statements-reports/diagnosis-ongoing-management-monitoring/new_diagnostic_criteria_for_diabetes

diabetes.org.uk. (2019, 03 08). *Diabetes and checking your blood sugars*. Retrieved from diabetes.org.uk: https://www.diabetes.org.uk/guide-to-diabetes/managing-your-diabetes/testing

diabetessupportsite.com. (2018, 04 05). *How To Test Blood Sugar Levels*. Retrieved from
        https://diabetessupportsite.com: https://diabetessupportsite.com/how-to-test-blood-
        sugar-levels

Diabetes-UK. (2012, 02 22). *GPs to prescribe smartphone apps*. Retrieved 01 19, 2018, from
        www.diabetes.org.uk/:
        https://www.diabetes.org.uk/About_us/News_Landing_Page/GPs-to-prescribe-
        smartphone-apps

Diabetes-UK. (2018, 04 29). *Cost of Diabetes*. Retrieved from www.diabetes.co.uk:
        https://www.diabetes.co.uk/cost-of-diabetes.html

Diagnostics, R. (2010). *Accu-Chek Smart Pix Device Reader User's Manual.* Mannheim, Germany:
        Roche Diagnostics GmbH.

Diasend. (2017). *Diasend.com*. Retrieved 2012, from Diasend.com:
        http://www.diasend.com/tutorial_flash/dot_english/tutorial_english.html

Diasend.com. (2018, 02 02). *About Us*. Retrieved 2012, from Diasend: https://diasend.com//en

Diffily, S. (2006). *The Website Manager's Handbook.* Lulu.com.

documentation.online.net. (2018, 04 20). *PHP CONFIGURATION*. Retrieved from
        documentation.online.net: https://documentation.online.net/en/web/web-hosting/web-
        management/configure-php

Dolan. (2009, March 18). *Interview: Lifescan on iPhone 3.0.* Retrieved 2012, from
        mobihealthnews.com: http://mobihealthnews.com/973/interview-lifescan-on-iphone-30/

Dolan, B. (2011, November 23). *Glooko picks up where Lifescan left off*. Retrieved 2018, from
        mobi health news: http://mobihealthnews.com/14861/glooko-picks-up-where-lifescan-
        left-off/

Dr Barry Boehm. (1988). *A Spiral Model of Software Development.* IEEE.

Dr Sandra Cabot MD, Margaret Jasinska ND. (2007). Tests to diagnose and monitor Diabetes. In M.
        J. Dr Sandra Cabot MD, *Diabetes Type 2: You Can Reverse it Naturally.* USA: SCB
        International Inc.

Dr. Sandra Cabot MD, Margaret Jasinska ND. (2011). Tests to dianose and monitor Diabetespage.
        In M. J. Dr. Sandra Cabot MD, *Diabetes Type 2: You Can Reverse it Naturally.* USA: SCB
        International Inc.

DSDM, C. . (2018, 02 21). *MoSCoW-prioritisation*. Retrieved from www.agilebusiness.org:
        https://www.agilebusiness.org/content/moscow-prioritisation

elitedatascience.com. (2018, 12 10). *Train Test Split*. Retrieved from elitedatascience.com:
        https://elitedatascience.com/wp-content/uploads/2017/06/Train-Test-Split-Diagram.jpg

Elizabeth Blair, M. P. (2013). *Joslin's Diabetes Deskbook, Updated 2nd Ed., Excerpt #33: Patient Education, Part 1.* Boston, MA (USA): Joslin Diabetes Center ; Wolters Kluwer Health/Lippincott Williams & Wilkins.

Exodus3000. (n.d.). *MD5 Encryption Tool*. Retrieved 2012, from Exodus3000.c0m: http://www.exodus3000.com/tools/md5/

Fahad Darwish, C. S. (2019). Diabetics' Self-Management Systems: Drawbacks and Potential Enhancements. *In Proceedings of the 2019 2nd International Conference on Geoinformatics and Data Analysis (ICGDA 2019)* (pp. 76-82). New York, NY, USA: ACM (DOI: https://doi.org/10.1145/3318236.3318247).

Fayyad, G. P.-S. (1996). *Fast Discovery of Association Rules. In Advances in Knowledge Discovery and Data Mining.* AAAI Press.

fda.gov. (2019, 03 10). *Citalopram (marketed as Celexa) Information*. Retrieved from fda.gov: https://www.fda.gov/Drugs/DrugSafety/ucm053339.htm

Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, 6-16.

Ferlie, R. A. (2006). *Networks, Organizational Learning and Knowledge Management: NHS Cancer Networks.* NHS Cancer Networks, Public Money & Management,.

Fink, A. (2013). *How to conduct surveys: A step-by-step guide (5th ed.).* Thousand Oaks,.

Fisher, T. (2013). *Ruby on Rails Bible.* Wiley.

Flick, U. . (2011). *Introducing research methodology: A beginner's guide to doing a research project. .* London: Sage.

Giuseppe Ciaburro, B. V. (2017). *Neural Networks with R.* Birmingham: Packt Publishing Ltd.

glooko.com. (2018, 04 04). *Glooko - Payer*. Retrieved from glooko.com: https://www.glooko.com/payer/

Goodwill, J. (2000). *Pure JSP: Java Server Pages.* Sams.

Gottesdiener. (1995, August). RAD Realities: Beyond the Hype to How RAD Really Works. *Application Development Trends*, pp. 28-38.

Gurusamy, V. &. (2014). Preprocessing Techniques for Text Mining. *RTRICS, At Podi*.

Haddi, E. (2015). *SENTIMENT ANALYSIS: TEXT PRE-PROCESSING, READER VIEWS AND CROSS DOMAINS.* London: BRUNEL UNIVERSITY LONDON.

Haddi, Emma. (2015, 03 05). *SENTIMENT ANALYSIS: TEXT PRE-PROCESSING, READER VIEWS AND CROSS DOMAINS.* Retrieved from Brunel.ac.uk: https://bura.brunel.ac.uk/bitstream/2438/11196/1/FulltextThesis.pdf

Hamel, L. H. (2011). *Knowledge Discovery with Support Vector Machines.* New Jersey: John Wiley & Sons Inc.

Hatton, S. (2007). Early prioritisation of goals. In Advances in conceptual modeling . In S. Hatton, *Advances in conceptual modeling* (pp. 235-244).

Heather Graz, V. K. (2018). Diabetes-related amputations create considerable public health burden in the UK. *sciencedirect*, 3.

Herath H, H. R. (2017). Gestational diabetes mellitus and risk of type 2 diabetes 10 years after the index pregnancy in Sri Lankan women-A community based retrospective cohort study. *PLoS One*.

Hilbe, J. M. (2009). *Logistic Regression Models.* Tempe - USA: CRC Press LLC, 2017.

Hollingsworth, L. A. (2013, 6 9). Reversal of type 2 diabetes: normalisation of beta cell function in association with decreased pancreas and liver triacylglycerol. *Diabetologia, 54*(10), 2506-2514.

Hussain, F. &. (2004). A Knowledge Management Framework to Operationalize Experiential Knowledge: Mapping Tacit Medical Knowledge with Explicit Practice Guidelines. *Proceedings of the National Conference on Emerging Technologies.*, 93–99.

IDF. (2019, March 07). *Type 2 diabetes.* Retrieved from idf.org: https://idf.org/52-about-diabetes.html

idf.org. (2017). *Diabetes facts & figures*. Retrieved 03 02, 2019, from https://www.idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html

Ikujiro Nonaka and Hirotaka Takeuchi. (1995). *The Knowledge-Creating Company - How Japanese Companies Creat the Dynamics of Innovation* . New York: Oxford University Press, Inc.

Independent. (2018, August 16). *Diabetes could 'bankrupt the NHS' after 60% rise in cases in a decade, charity warns*. Retrieved from www.independent.co.uk: https://www.independent.co.uk/life-style/health-and-families/health-news/diabetes-could-bankrupt-the-nhs-after-60-rise-in-number-of-cases-charity-warns-10458353.html

Indiana University. (2018, 04 03). *what is JSP?* Retrieved from https://kb.iu.edu: https://kb.iu.edu/d/anqh

Indiana University. (2018, 04 03). *What is JSP?* Retrieved from https://kb.iu.edu: https://kb.iu.edu/d/anqh

Ioannis Kavakiotis, O. T. (2017). Machine Learning and Data Mining Methods in Diabetes Research. *Computational and Structural Biotechnology Journal, Volume 15*, 104-116.

Ioannis Kavakiotis, O. T. (2017). Machine Learning and Data Mining Methods in Diabetes Research,. *Computational and Structural Biotechnology Journal*, 104-116.

Isadore Newman, C. R. (1998). Qualitative-quantitative Research Methodology: Exploring the Interactive Continuum. *SIU Press*, 218.

Jakob, N. (1999, 3 October). *10 Good Deeds in Web Design*. Retrieved from nngroup.com: https://www.nngroup.com/articles/ten-good-deeds-in-web-design/

Jen Stirrup, R. O. (2017). *Advanced Analytics with R and Tableau.* Birmingham: Packt Punlishing LTD.

Jianhua Yang, A. G.-D. (2009). *Information Systems: Modeling, Development, and Integration: .* Germany: Springer Science & Business Media.

Jiawei Han, M. K. (2012). knowledge discovery from data. *In The Morgan Kaufmann Series in Data Management Systems,Data Mining (Third Edition)*, 1-38.

Jixia Cao, Z. Y. (2010). Utilizing SECI Model for knowledge management in library. *International Conference on Educational and Information Technology*, pp. V3-504-V3-506.

JNJDI. (2012). *Johnson and Johnson Diabetes Institute*. Retrieved 2012, from Johnson and Johnson Diabetes Institute: http://www.jjdi.com/index.html

Jobs. (2010, April). *Thoughts on Flash*. Retrieved 2011, from apple.com: http://www.apple.com/hotnews/thoughts-on-flash/

Jon, W. (2018, 04). *Blood Sugar Chart.* Retrieved 2012, from Vertex42: http://www.vertex42.com/ExcelTemplates/blood-sugar-chart.html

Joslin. (2012). *Ketone Tesing: What you need to know*. Retrieved 2012, from Joslin Diabetes Center: http://www.joslin.org/info/ketone_testing_what_you_need_to_know.html

Jurafsky, D. a. (2014). *Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.* India: Dorling Kindersley Pvt, Ltd.

Kapur. (2011). Unite for Diabetes. *WDF Newsletter*. Gentofte Denmark: World Diabetes Foundation.

Kenneth D. Lawrence, S. K. (2008). *Data Mining Methods and Applications.* Taylor & Francis Group.

Kothari, A. H. (2011). Lessons from the Business Sector for Successful Knowledge Management in Health Care: A Systematic Review. *BMC health services research. *, 173.

Kuhlman, D. (2012, April 22). *A Python Book: Beginning Python, Advanced Python, and Python Exercises*. Retrieved from https://web.archive.org/web/20120623165941/http://cutter.rexx.com/~dkuhlman/python_book_01.html

Kurgan, L.A. and Musilek, P. (2006). A Survey of Knowledge Discovery and Data Mining Process Models. *The Knowledge Engineering Review,*, 21, 1-24.

Lamb. (2008). DIASEND: one device to read them all and in the clinic guide them. *Endocrine Abstracts, 17*, 48.

Lancet. (2011, July 9). The diabetes pandemic. *The Lancet, 378*(9786), 99.

Larme & Pugh, A. L. (2001). Evidence-based guidelines meet the real world: The case of diabetes care. *Diabetes Care*, 1728-1733.

Larme, P. (1998). Attitudes of primary care providers toward diabetes: barriers to guideline implementation. *Diabetes Care, 21*(9), 1391-1396.

Leedy, P. a. (2001). Practical Research: Planning and Design. *7th Edition, Merrill Prentice Hall and SAGE Publications*.

Lesley Weaving, R. B. (2017, August 17). *The diabetic foot.* Retrieved from www.pharmaceutical-journal.com: https://www.pharmaceutical-journal.com/learning/cpd-article/the-diabetic-foot/20203299.cpdarticle

Li, X., Zhu, Z., & Pan, X. (2017, 04 04). *Knowledge cultivating for intelligent decision making in small & middle businesses.* Retrieved from Procedia Computer Science. [Online]: http://linkinghub.elsevier.com/retrieve/pii/S1877050910002814.

LifeScan (OneTouch). (2018). *OneTouch Reveal Diabetes*. Retrieved 03 01, 2018, from OneTouch Reveal Diabetes: https://www.myonetouch.co.uk/OneTouchReveal#upgrade-your-mobile-app

LifeScan. (2018). *OneTouch Reveal Diabetes*. Retrieved 03 01, 2018, from OneTouch Reveal Diabetes: https://www.myonetouch.co.uk/OneTouchReveal#upgrade-your-mobile-app

LifeScan.com. (2013, 03 25). *LifeScan Announces Voluntary Recall of All OneTouch® Verio®IQ Blood Glucose Meters*. Retrieved 01 05, 2018, from www.lifescan.com: https://www.lifescan.com/about-us/news/244

LifeScanInstitute. (2018, 05 03). *lifescandiabetesinstitute*. Retrieved from https://www.lifescandiabetesinstitute.com/: https://www.lifescandiabetesinstitute.com/

Lin Zhanga, K. H. (2014). Sentiment Analysis on Reviews of Mobile Users . *The 11th International Conference on Mobile Systems and Pervasive Computing*, 458 – 465 .

Lior Rokach. (2015). *Data Mining with Decision trees.* London: World Scientific Publishing Co. Pte. Ltd.

M.A., C. K. (2007). The Web 2.0 driven SECI model based learning process. Seventh IEEE International Conference on Advanced Learning Technologies.

Ma, Q. (2009). *The Effectiveness of Requirements Prioritization.* Auckland University of Technology.

Maginnis. (2011, October 27). *Diabetes UK Tracker iPhone App Review*. Retrieved 2012, from Trusted Reviews: http://www.trustedreviews.com/diabetes-uk-tracker-iphone-app_Mobile-App_review

Maldow, D. S. (2019, 01 08). *Zoom's Full Featured UME Videoconferencing Platform Exceeds Expectations*. Retrieved from telepresenceoptions.com: http://www.telepresenceoptions.com/2013/01/zooms_full_featured_ume_videoc/

Margaret O'Malley, Kelly Kennedy, RD. (2018, 02 02). *uncontrolled diabetes* . Retrieved from www.everydayhealth.com: https://www.everydayhealth.com/type-2-diabetes/living-with/what-its-like-to-have-uncontrolled-type-2-diabetes/

Mathieu-Fritz, Alexandre & Guillot, Caroline. (2017). *Diabetes self-monitoring devices and transformations in "patient work".* (Vols. (Vol. 11, N°4)). S.A.C. (http://www.socanco.org).

Mayo Clinic. (2017, October 24). *Diabetes - Blood sugar testing: Why, when and how.* Retrieved 2017, from Mayo Clinic: http://www.mayoclinic.com/health/blood-sugar/DA00007

Mayoclinic.org. (2018, January 03). *Diabetes*. Retrieved 2012, from Mayo, Clinic: http://www.mayoclinic.com/health/type-2-diabetes/DS00585/TAB=indepth

McNeill, P. and Chapman, S. (2005). *Research Method.* London : Rouledge.

MEARS, C. (2018, 04 11). *Sitemaps – The Beginner's Guide*. Retrieved from theuxreview.co.uk: https://theuxreview.co.uk/sitemaps-the-beginners-guide/

Melonfire. (2006, September 22). *Understanding the pros and cons of the Waterfall Model of software development*. Retrieved 2012, from TechRepublic: http://www.techrepublic.com/article/understanding-the-pros-and-cons-of-the-waterfall-model-of-software-development/6118423

Merck Manual. (2007). Diabetes mellitus (DM). In Crandall (Ed.), *Merck Manual for Health Care Professionals.* New York: Simon and Schuster.

Metzger, C. R. (1993, December). Prepregnancy Weight and Antepartum Insulin Secretion Predict Glucose Tolerance Five Years After Gestational Diabetes Mellitus. *Diabetes Care, 16*(12), 1598-1605.

Michel, E. (2001, Sept- Oct). Perceived difficulty of diabetes treatment in primary care: does it differ by patient ethnicity? *Diabetes Education, 27*(5), 678-684.

Mirza, R. (2009). *Knowledge Management and Clinical Framework for Cross Country Healthcare Organizations.A Thesis Submitted in partial fulfilment of the Requirements of Blekinge Institute of Technology for the Degree of Master of Science in Computer Science. .* Sweden: Blekinge Institute of Technology.

Morr, C. El. and Subercaze, J. (2010). Knowledge Management in Healthcare. Handbook of Research on Developments in E-Health and Telemedicine, . *Technological and Social Perspectives*, 490-510.

Moukheiber, Z. (2014, 11 12). *How One Start-Up Is Using Fitness Trackers To Help Diabetes Patients*. Retrieved from forbes.com: https://www.forbes.com/sites/zinamoukheiber/2014/11/12/how-one-start-up-is-using-fitness-trackers-to-help-diabetes-patients/#6023ba513c86

Muddasarul Hoda, S. H. (2019). *Role of Phenolic Phytochemicals in Diabetes Management.* Singapore: Springer Nature Singapore Pte Ltd.

MySQL. (2018). *What is MySQL; About*. Retrieved 2011, from My SQL.com: http://dev.mysql.com/doc/refman/5.1/en/what-is-mysql.html

mySQL.com. (2018). *Limits on Table Size*. Retrieved from mysql.com: https://dev.mysql.com/doc/refman/8.0/en/table-size-limit.html

Nam Han Cho, J. K. (2017). *The Diabetes Atlas (IDF Diabetes Atlas) Eighth Edition.* International Diabetes Federation.

Natek, S. (2016). *KNOWLEDGE MANAGEMENT SYSTEMS SUPPORT SECI MODEL OF KNOWLEDGE-CREATING PROCESS.* Slovenia: International School for Social and Business Studies.

National Institutes of Health. (2019, 02 11). *Insulin Resistance & Prediabetes*. Retrieved from National Institutes of Health: https://www.niddk.nih.gov/health-information/diabetes/overview/what-is-diabetes/prediabetes-insulin-resistance

newscientist.com - Matt Reynolds. (2017, 02 13). *NHS to start prescribing health apps that help manage conditions*. Retrieved 01 23, 2018, from newscientist.com: https://www.newscientist.com/article/2121164-nhs-to-start-prescribing-health-apps-that-help-manage-conditions/

NHS. (2012, 04 05). *Clinical Commissioning Priority Update: Diabetes.* Retrieved from www.westcheshireccg.nhs.uk: https://www.westcheshireccg.nhs.uk/document_uploads/Papers-may2012/WCCCGB120551DiabetesUpdate.pdf

nhs.uk. (2017, 08 08). *Diabetes, Type 2 - NHS Choices.* Retrieved 02 13, 2012, from nhs.uk: http://www.nhs.uk/Conditions/Diabetes-type2/Pages/Introduction.aspx

NHS.UK. (2018, 02 01). *NHS - Gestational Diabetes*. Retrieved 02 01, 2018, from www.nhs.uk: https://www.nhs.uk/conditions/gestational-diabetes/

nhs.uk. (2018, 03 01). *NHS-Type 1 Diabetes*. Retrieved from www.nhs.uk: https://www.nhs.uk/conditions/type-1-diabetes/

Ni, G. W. (2010). Research on the Knowledge Management System of the Vicarious Management Corporation. In G. W. Ni, *International Conference of Information Science and Management Engineering.* (pp. 62–67). http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5573886.

Niaksu, O. (2015). CRISP Data Mining Methodology Extension for Medical Domain. *Baltic J. Modern Computing*, 92-109.

NICE.org.uk. (2018, 04 04). *National Institute for Health and Care Excellence (NICE)*. Retrieved from nice.org.uk: https://www.nice.org.uk/about

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, 14-37.

Nura Esfandiari, M. R.-M. (2014). Knowledge discovery in medicine: Current issue and future trend. *Expert Systems with Applications*, 4434-4463.

Oded Maimon, L. R. (2005). *Data Mining and Knowledge Discovery Handbook.* Boston, MA: Springer.

OECD. (2012, 04 24). *Health: The highest cost of Diabetes*. Retrieved from http://www.oecd.org: http://www.oecd.org/eu/healththehighcostofdiabetes.htm

Omar Al Awad. (2015). *The Uptake of Advanced IT with Specific Emphasis.* Manchester - United Kindom: University of Salford .

Onur R B, A. K. (2019). NURBS-Python: An open-source object-oriented NURBS modeling framework in Python. *SoftwareX*, 85-94.

opensourceforu.com. (2017, May 16). *Choose the Right Database for Your Application*. Retrieved from www.opensourceforu.com: https://opensourceforu.com/2017/05/choose-right-database-application/

oracle.com. (2018). *Oracle*. Retrieved 2012, from oracle.com: http://www.oracle.com/technetwork/java/javaee/jsp/index.html

Orzano, A. M. (2008). Family Medicine Practice Performance and Knowledge Management. *Health Care Management Review. *, 21-28.

Patil, B. J. (2010). Hybrid prediction model for Type-2 diabetic patients. *Expert Systems with Applications.*, 8102– 8108.

Paul Kurdyak, S. V. (2017). Diabetes quality of care and outcomes: Comparison of individuals with and without schizophrenia. *General Hospital Psychiatry*, 7-13.

Peggy, E.-M. (2007). Doing My Best: Poverty and Self-Care Among Individuals With Schizophrenia and Diabetes Mellitus,. *Archives of Psychiatric Nursing,*, 49-60.

Petrash, G. (1996). *Managing Knowledge Assets for Value, Knowledge-Based Leadership Conference.* Boston: Linkage, Inc.

Phillip A. Laplante. (2004). *SOFTWARE ENGINEERING for IMAGE PROCESSING SYSTEMS.* Florida - USA: Taylor & Francis.

php.net. (2018). *http://uk3.php.net/tut.php*. Retrieved 2012, from php.net:
http://uk3.php.net/tut.php

php.net. (2018, 04 20). *what is php?* Retrieved from php.net: http://php.net/manual/en/intro-
whatis.php

Qiao Ma . (2009). *The Effectiveness of Requirements Prioritization Techniques for a Medium to
Large Number of Requirements: A Systematic Literature Review.* Auckland University of
Technology.

Quinn, E. H.-S. (2014). How Can Knowledge Exchange Portals Assist in Knowledge Management
for Evidence- Informed Decision Making in Public Health? . *BMC Public Health.*, 233.

Resnicow, H. (2007, October). Helping Patients Make and Sustain Healthy Changes: A Brief
Introduction to Motivational Interviewing in Clinical Diabetes Care. *ADA Clinical Diabetes,
26*, 161-165.

Resul Das, I. T. (2009). Effective diagnosis of heart disease through neural networks ensembles.
Expert Systems with Applications, 7675-7680.

Rhoda K. Moise, B. D. (2017). *Diabetes Knowledge, Management, and Prevention Among Haitian
Immigrants in Philadelphia.* Miami: sagepub.com / Department of Public Health Sciences,
University of Miami Miller School of Medicine.

Richard S. Surwit. (2013). Caffine & Diabetes. In R. S. Surwit, *The Mind-Body Diabetes Revolution:
A Proven New Program for Better Blood Sugar Control.* New York: Publisher Simon and
Schuster.

Robinson, J. S. (2017). *Text Mining with R.* United States of America: O'Reilly Media Inc.

Roche.co.uk. (2018, 05 01). *Roche Diabetes Care*. Retrieved from www.roche.co.uk:
https://www.roche.co.uk/home/company-portrait/diabetes-care.html

rstudio.com. (2018, 08 05). *https://www.rstudio.com/*. Retrieved from https://www.rstudio.com/:
https://www.rstudio.com/

rubyonrails.org. (2011). *Ruby on Rails*. Retrieved 2012, from rubyonrails.org:
http://rubyonrails.org/

Rusdah, R. &. (2013). Review on Data Mining Methods for Tuberculosis Diagnosis. . *Information
Systems International Conference (ISICO)*.

S C.M. Croxson, A. J. (2009). Diabetes Mellitus. In A. J. Croxson, *In Brocklehurst's Textbook of
Geriatric Medicine and Gerontology (SEVENTH EDITION)* (pp. 760-774). Philadelphia:
Philadelphia, PA: Saunders/Elsevier.

Salus, P. H. (1998). *Handbook of programming languages.* Indianapolis, IN : Macmillan Technical
Pub.

Sarah Itani, F. L. (2019). Specifics of medical data mining for diagnosis aid: A survey. *Expert Systems with Applications*, 300-314.

Sarah Itani, F. L. (2019). Specifics of medical data mining for diagnosis aid: A survey,. *Expert Systems with Applications, Volume 118,*, 300-314,.

Saunders et. (2009). Research Methods for Business Students. In *Research Methods for Business Students* (p. 130). Hall: Prentice.

Saunders, Thornhill and Lewis. (2015). Research Methods for Business Students. In *Research Methods for Business Students* (p. 130). Hall: Pearson.

Schlichting. (2008, September 5). *What is SQL Server*. Retrieved 2011, from databasejournal.com: http://www.databasejournal.com/features/mssql/article.php/3769211/What-is-SQL-Server.htm

scikit-learn.org. (2018, 02 05). *Feature extraction*. Retrieved from scikit-learn.org: https://scikit-learn.org/stable/modules/feature_extraction.html

Shayban, D. A. (2017). PDB48 - Using Diabetes Self-Management Questionnaire (DSMQ) to Assess Diabetes Self-Care Activities for Diabetes Patients in King Fahad University Hospital - Saudi Arabia. *Value in Health,*, A482-A483.

SiDiary. (2018). *Online shop*. Retrieved 04 30, 2018, from SiDiaryonline: https://shop.sinovo.net/index2.php?language=en&idMenu=3

Silhavy, R. (2018). Stop Words Removal. In R. Silhavy, *Artificial Intelligence and Algorithms in Intelligent Systems* (pp. 32-50). Zlin - Czech Republic: Springer International Publishing.

Singhal, B. (2018). *Beginning Blockchain.* Germany: Apress Media LLC: Welmoed Sphr .

Snieder R. & Larner, K. (2009). *The Art of Being a Scientist: A Guide for Graduate Students and their Mentors.* Cambridge: Cambridge University.

Stephens, R. (2015). MoSCoW. In R. Stephens, *Beginning Software Engineering* (p. 57). Indianapolis: John Wiley & Sons Inc.

Suffer, D. (2009). *Designing gestural interfaces.* Beijing: O'Reilly.

Sut, N. &. (2011). Comparison of regression tree data mining methods for prediction of mortality in head injury. *Expert Systems with Applications*, 15534–15539.

Sut, N. S. (2011). Comparison of regression tree data mining methods for prediction of mortality in head injury. *Expert Systems with Applications*, 15534–15539.

symantec.com. (2019, 01 05). *Beginner's Guide to SSL Certificates*. Retrieved from https://www.symantec.com: https://www.symantec.com/content/en/uk/enterprise/white_papers/b-beginners-guide-to-ssl-certificates_WP.pdf

Takeuchi Hirotaka, Shibata Tsutomu. (2006). *Japan, Moving Toward a More Advanced Knowledge Economy* (Vol. 2). Washington, DC: The International Bank for Reconstruction and Development/The World Bank .

Tamer Uçar, A. K. (2011). Predicting existence of Mycobacterium tuberculosis on patients using data mini approaches. *Procedia Computer Science*, 1404-1411.

Tang, H. (2009). Tang,H.,Tan,S.,andCheng,X.Asurveyonsentimentdetectionofreviews. *Expert Systems with Applications*, 36(7):10760–10773.

Tenderich. (2009, March 17). *Lifescan's New Diabetes iPhone App*. Retrieved 2012, from Diabetes Mine: http://www.diabetesmine.com/2009/03/lifescans-new-diabetes-iphone-app.html

The Guardian. (2016, August 03). *Diabetes drugs cost NHS nearly £1bn a year*. Retrieved October 10, 2017, from theguardian.com: https://www.theguardian.com/society/2016/aug/03/diabetes-drugs-cost-nhs-1bn-a-year-presciption

The Lancet . (2011, July 9). The diabetes pandemic. *The Lancet - Elsevier, 378*(9786), 99.

The Lancet. (2011, July 9). The diabetes pandemic. *The Lancet, 378*(9786), 99.

The New Encyclopaedia Brittanica. (1998). Diabetes. In *The New Encyclopaedia Brittanica* (Vol. 4 Micropedia , p. 60). Chigago, Illinois, USA: Encyclopaedia Brittanica.

Tim Converse, J. P. (2004). Hashing using MD5. In J. P. Tim Converse, *Php5 and MySQL Bible* (p. 435). USA: John Wiley & Sons Inc.

Tudor, D. &. (2006). Using an agile approach in a large traditional traditional organisation. Proceedings of AGILE 2006 Conference (AGILE'06).

Turner, M. N. (1998, March 14). Risk factors for coronary artery disease in non-insulin dependent diabetes mellitus: United Kingdom prospective diabetes study (UKPDS: 23). *British Medical Journal*, 316:823.

United Nation. (2019, 03 02). *Population 2030*. Retrieved from un.org: https://www.un.org/en/development/desa/population/publications/pdf/trends/Population2030.pdf

University Of Salford Website. (2019, 03 19). *Ethics and Research Governance*. Retrieved from salford.ac.uk: https://www.salford.ac.uk/ethics

Unwin, W. G. (2011). *The Diabetes Atlas Fifth Edition.* Brussels: International Diabetes Federation.

Veglis, A. (2005). PHP and SQL made simple . *Distributed Systems Online, IEEE*.

Vohra, D. (2007). *Ruby on Rails for PHP and Java Developers.* Springer.

W. Jason Gilmore. (2008). *Beginning PHP and MySQL.* Berkeley, CA: W. Jason Gilmore.

w3techs.com. (2019, 01 02). *Usage statistics and market share of PHP for websites*. Retrieved from w3techs.com: https://w3techs.com/technologies/details/pl-php/all/all

Walker, J. R. (1964). Carbohydrate consumption and diabetes. *Proceedings of the Nutritional Society, 23*, 143-149.

Wasson, C. S. (2006). THE EVOLUTIONARY DEVELOPMENT MODEL. In C. S. Wasson, *System Analysis, Design, and Development Concepts, Principles, and Practices* (p. 295). Hoboken, New Jersey.: John Wiley & Sons, Inc., .

waterfall-model. (2012). *All about the Waterfall Method*. Retrieved 2012, from Waterfall-model.com: http://www.waterfall-model.com/

Wayne Goddard, S. M. (2004). *Research Methodology.* Juta and Company Ltd.

webhostingmedia.net. (2018, 02 02). *PHP vs Python – What Are the Differences & Similarities?* Retrieved from webhostingmedia.net: https://webhostingmedia.net/php-vs-python/

webmd.com. (2018, 12 12). *Gilenya*. Retrieved from webmd.com: https://www.webmd.com/drugs/2/drug-154676/gilenya-oral/details

Weinger, Y. P. (2003, march). The Confidence in Diabetes Self-Care Scale. *Diabetes Care, 26*(3), 716-718.

Weinhandl, K. I. (2015). Information Overload in Healthcare: Too Much of a Good Thing? . *The Journal of Evidence and Quality in Health Care.*, 285–290.

Whitehead, P. (2001). *PHP: Your Visual Blueprint for Creating Open Source, Server-Side Content. .* Visual.

William D. McArdle, B. (. (2009). Exercise Physiology : Nutrition, Energy, and Human Performance. In B. (. William D. McArdle, *Exercise Physiology : Nutrition, Energy, and Human Performance* (pp. 5-435). Philadelphia, United States: Lippincott Williams and Wilkins.

William D. McArdle, F. I. (2010). Exercise Physiology : Nutrition, Energy, and Human Performance. In B. (. William D. McArdle, *Exercise Physiology : Nutrition, Energy, and Human Performance* (pp. 5-435). Philadelphia, United States: Lippincott Williams and Wilkins.

Williams, C. (2007). Research Methods. *Journal of Business & Economic Research*, 65-72.

Wittwer. (2003). *Blood Sugar Chart.* Retrieved 2012, from Vertex42: http://www.vertex42.com/ExcelTemplates/blood-sugar-chart.html

World Health Organization. (2018, 04 03). *Diabetes Programme*. Retrieved from www.who.int: http://www.who.int/diabetes/en/

Wyster. (2008, September 28). *Using Rapid Application Development for your software project*. Retrieved 2012, from The Wysteria: http://wysterdesir.com/2008/09/28/using-rapid-application-development-for-your-software-project/

Yangyang Li a, ∗. J. (2009). Hydraulic conductance and vulnerability to cavitation in corn (Zea mays L.). *Environmental and Experimental Botany*, 1-10.

Yin Kia Chiam, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics Volume 36*, 82-93.

Yue-Long OU, L.-n. Z. (2010). Researching on MD5's characteristics based on software reversing,. *The Journal of China Universities of Posts and Telecommunications*, 127-130. Retrieved 2012, from Exodus3000.c0m: http://www.exodus3000.com/tools/md5/

Zheng, Y. L. (2017). Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nature Reviews Endocrinology*, 88.

9   Appendices

9.1   Appendix A: Email confirmation of SPARC 2018 Acceptance

From: sparc <Rgc-Sparc@salford.ac.uk>
Sent: 09 May 2018 16:52
To: Darwish, FAHAD (PG)
Cc: Mo Saraee
Subject: SPARC 2018

Dear Fahad,

We're writing to inform you that your abstract has been **accepted with minor amendments** to SPARC taking place 4-5th July 2018 at Media City UK. This emails provides details of the next steps and further information about SPARC.
This year all abstracts were matched to the themes of the University's Research Beacons, written for a lay audience and submitted as a competitive process. All abstract submissions have been peer reviewed and we've copied details of the feedback given below for your information. Please also find attached further guidance on abstract writing that you may find useful and we will be running a SPARC re-write writing retreat on:
https://myadvantage.salford.ac.uk/students/events/detail/549495

"The title needs to be removed from the beginning of the main body of the abstract – this will affect the formatting of the booklet"

To re-submit your abstract, please copy and paste it into this form no later than 18th May 2018: https://salford.onlinesurveys.ac.uk/abstract-re-submission-form-sparc-2018

Due to the number of abstracts submitted and accepted we can only offer participants their first choice. We have therefore allocated you an **oral presentation** and the full programme will be circulated to all presenters shortly.

Oral Presenters will be asked to provide their presentations by email in advance by 4pm on 4th June 2018. Presentations are to last no more than 5 minutes, and if using Powerpoint – contain no more than 4 slides. Questions will be reserved until the end, once all presenters have presented and all slides will be loaded onto the lecterns in advance.

Please note: parts of the conference will be recorded and the posters will be made available online after it has taken place. Please consult your with supervisor to consider the copyright implications should you wish to publish the work you are presenting, and notify us if your presentation is not to be made publicly available.

Should you have any queries, please do not hesitate to contact us,

Kindest Regards,
Tracy and Davina (on behalf of the SPARC Committee)

University of Salford MANCHESTER

SPARC 2018
Research & Knowledge Exchange
Address ie: 827 Maxwell Building, University of Salford, Manchester M5 4WT
T: +44(0) 0161 295 3671
SPARC@salford.ac.uk / www.pg.salford.ac.uk / @SalfordPGRs

9.1.1   Appendix A (Continued) : SPARC copy of Abstract

Intelligent Diabetes Management for Better Management and Reduced Costs to Health Providers

Fahad Darwish

School of Computing, Science & Engineering
Email: f.darwish@edu.salford.ac.uk

Supervisor
Dr Mohamad (Mo) Saraee - m.saraee@salford.ac.uk

Abstract

Diabetes is a growing worldwide pandemic that, by the year 2030 is expected to claim three people every ten minutes. By the year 2035, 17% of the NHS annual budget is expected to be spent on diabetes treatment, a figure of £16.8 billion, costs from this disease's complications could be responsible for bankrupting the British National Health System.

It is of great social importance that these costs and complications be reduced as much as possible by preventing new cases and minimising existing cases. This goal will be accomplished through patient education and self-management.

A major component of self-management involves regular blood testing and detailed record keeping. Detailed comparative analysis, extensive diabetic and medical surveying were used to determine how to best create universally available diabetes management website that will help the medical community and diabetic community come together to reduce the harmful effects of this disease.

The web application that has been created successfully connect the doctor to patient whilst providing the patient with a single means of recording every important vital health factor and simultaneously permitting the doctor real-time monitoring and input.

Our approach aims to review the existing big data tools and technologies available and apply data mining techniques on the massive data. This research will also study refining computer system analysis and mining to predict treatment for diabetic patients.

Keywords

Diabetes, NHS, patient education, self-management

9.2   Appendix B: Abstract on IEEE 11th International Conference on the Developments in eSystems Engineering (DeSE 2018)

Diabetes Self-Management System: Review of existing systems and potential enhancements	Fahad Darwish, Dr Mohamad Saraee

Diabetes is a global pandemic with growing devastating human, social and economic impacts. By 2025 it is estimated that in the UK five million people will be diagnosed with diabetes and by 2030 diabetes will claim three lives every ten minutes. According to the most recent reports, 10% of the British National Health Service (NHS) budget for England and Wales is spent on diabetes, equating to £1.5m an hour. With diabetes increasing at its current rate, this spending is unsustainable. The most shocking finding of this report is that almost four-fifth of NHS diabetes spending is on preventable complications. Self-Monitoring and Self-Management systems play a significant role in preventing and reducing diabetes complications. Minimising costs, and improving patient outcomes could be achieved through self-management and empowering the diabetic patients with knowledge management (KM). This research provides a review of existing monitoring and management systems which have been created by blood glucose or insulin pump manufacturers. Initial findings suggest that profit rather than patient outcomes are the main motivator behind the existing systems. To this end, this research proposes that the prevalence of personal computers and mobile smart phones be leveraged to develop a web portal for diabetes self-management. This overcomes the limitation of the existing system, which is lack of ability to communicate with health professionals and none of the existing options are free to the patient. The portal would facilitate a closer relationship between the medical and diabetic communities. This strengthened relationship would allow patient and doctors to work together to find effectively ways in which the damaging effects of this disease can be reduced. The proposed portal will help patients to record each imperative health factor. This would simultaneously allow the doctor access to real-time monitoring and give them the ability to communicate directly with the patients. Additionally, this paper investigates the potential advantages of integrating knowledge management approaches to disseminate and share knowledge among related parties.

Source: " http://dese.org.uk/wp-content/uploads/ProgrammeFull_DeSE2018_V7-1-1.pdf "

9.3 Appendix C – Final Paper of ACM--2019 2nd International Conference on Geoinformatics and Data Analysis (ICGDA 2019)--Ei Compendex and Scopus

Diabetics' Self-Management Systems: Drawbacks and Potential Enhancements

Fahad Darwish
School of Computing, Science
and Engineering
University of Salford
Manchester, M5 4WT, UK
f.darwish@edu.salford.ac.uk

Charith Silva
School of Computing, Science
and Engineering
University of Salford
Manchester, M5 4WT, UK
h.c.e.silva@edu.salford.ac.uk

Mo Saraee
School of Computing, Science
and Engineering
University of Salford
Manchester, M5 4WT, UK
m.saraee@salford.ac.uk

ABSTRACT

Diabetes is a pandemic that is growing globally, and by the year 2030 it is expected to effect three people every 10 minutes. In the UK, it is estimated that by 2025, 5 million people will have diabetes. Diabetes is currently costing the British National Health Service (NHS) over £1.5m an hour. This equates to 10% of the NHS budget for England and Wales or over £25,000 being spent on diabetes every minute. It is important to minimise these costs by employing new techniques to curtail existing cases and limit new cases. This goal will be accomplished through patient education and self-management. Self-management and self-monitoring play a significant role in restraining diabetes complications. A major component of self-management involves regular blood testing and detailed record keeping through comparative analysis, extensive surveying of diabetic's patients and medical professionals were carried out.

The outcome of this research was used to determine how to best create a global and easily accessible diabetes management e-toolkit, that will help the diabetic community alongside the medical community to reduce the harmful effects of this disease. The e-toolkit, that has been formed successfully, can connect the patient to their doctor. It will provide the patient with a single means of recording every important vital health factor and simultaneously allowing the doctor to access real-time monitoring.

The e-toolkit proposed in this paper would facilitate in bringing both the medical and diabetic group of people closer together resulting in a strengthened relationship. Patients will be able to record each imperative health factor whilst having the ability to communicate with their doctor and in turn, effectively managing their diabetes to their utmost potential.

CCS Concepts

• Information systems → Data analytics
• Applied computing → Health care information systems

Keywords

Diabetes, British National Health Service, E-toolkit, Doctor, Patient

1. INTRODUCTION

Diabetes is a long-lasting health disorder which distresses people of all ages, race and economic classes. If not treated, the effects may be fatal, involving long-term harm to vital organs. Usually, diabetes is a lasting health disorder, medicinal specialists concur that it requires close supervision, observation and intense self-organisation (American Diabetes Association, 2018).

The global healthcare expenditure on people with diabetes was estimated to be USD 850 billion (£651 billion) in 2017 (N. H. Cho, 2018). As these numbers race out of control, the medical community struggles to keep up with the demand for care and societies around the world attempt to adjust to this 'catastrophic' pandemic (webmd.com, 2019). On the 8th November 2018, according to the National Diabetes Audit Report, diabetes prescriptions now cost NHS £1bn (National Diabetes Audit Report, 2017).

Diabetic patients play a big role in controlling the disease, with self-management by ensuring that regular measurements of their blood glucose levels are recorded after waking up, before a meal and after a meal, before and after an exercise, before going to bed, and at other occasions, if and when required (Alan J. Sinclair and Simon C.M. Croxson, 2009). Consequently, doctors are frustrated in their treatment of diabetic patients and find it difficult to adjust medications and other factors when patients are unable to express how various influences affect their condition (Larme AC1, 1998). When doctors don't get the information, they require in a timely manner, this places diabetic patients at a much higher risk of developing complications.

To this end, this research has utilised the ever-increasing use of personal computers and mobile smart phones to develop a newer and better e-toolkit for diabetics to monitor and record their health risk factors. (Darwish F, 2012). The existing diabetes monitoring tools and applications were created by blood glucose or insulin pump manufactures. Therefore, these applications are almost always created from a commercial vantage, and not necessarily with either the patient's or the medical community's best interests in mind hence they are often lacking in their effectiveness and applicability. Thus, this research has proposed an e-toolkit which overcomes the limitations of the existing applications.

2. LITERATURE REVIEW

This section discusses in detail the existing diabetes monitoring systems in the market. It highlights their limitations from the perspectives of diabetic patients and the concerned medical staff.

2.1 Diabetes UK App

This application enables diabetics to track something beyond blood glucose and insulin readings. It can record calories, carbohydrates, body weight and ketones. The recorded information can then be transferred into graphs to illustrate the data in a visual format (Diabetes UK,

2018). However, the ability to send patient data to medical professionals is not available and therefore this restricts the communication of readings between the patient and their doctor (NHS Croydon Clincical Commisioning Grp, 2018). They would need to print the data which adds further steps and equipment resulting in an impact on the environment as paper wastage is involved. Since the lack of funding from the charity Diabetes UK, the app is no longer used by many patients and therefore has been discontinued on the IOS software.

2.2 Accu-Chek
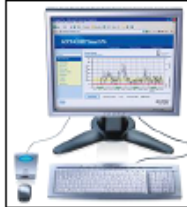
Accu-Chek Smart Pix system is a product by Roche Diabetes Care based in the United Kingdom, the most famous, global producer of diabetic equipment (accu-chek.co.uk, 2019). The system is efficient and has an interface that records blood glucose and insulin levels directly from the insulin pump made only by Accu-Chek. The data is then generated on to the user's PC, which is then filtered automatically and accordingly, reducing any human error.

The drawbacks of this system are the need for other equipment like a laptop or PC and a specific software to upload the data to be converted into relevant information. This is restrictive as not all patients, such as the elderly or low-income class, have accessibility to this equipment and they may also struggle to set the system up. Also, the system can be quite costly depending on where you are residing, as in the UK it is priced at £49.99. (amazon.co.uk, 2019). However, in the US the same system is priced at a much higher price of $148. (glucosemeters4u.com, 2019).

Finally, the system may be enough for a simple glucose reading to inform the patient to help them keep track however, the system has no capability to inform or transfer this data to medical professionals through a portal or even by email (accu-chek.co.uk, 2019). Therefore, the only option is to print the data and take it on appointments, which again is limits the efficiency and accuracy.

2.3 One Touch

OneTouch (formerly known as LifeScan) is a mobile/smart phone application made by a division of JNJ, a well-known pharmaceutical company in America (myonetouch.co.uk, 2019). The app is readily available on the IOS platform and can link directly with blood glucose meters through Bluetooth or an iPhone cable linking to a laptop/PC. However, this is not directed to the entire population as this is only available to IOS users and is not accessible by other popular operating systems such as Android. A major downside to this app was a recall in 2013 caused by readings of high blood glucose levels of 1024mg/dl or above causing the device to shut off. Therefore, this app was not reliable and effective for diabetics. This caused a lack in trust in the company, monetary losses and major time delays.

2.4 Glooko Diasend

The Diasend System depends on a device, called a 'transmitter', which the patient must physically connect to their PC. This transmitter is intended to automatically detect and import readings from various glucose meters and insulin pumps to the server. The information can then be expanded into a chart format in order to

provide visual illustrations of the collected data. The information that is uploaded on to the system can be accessed from anywhere. However, in order to transfer the data to the server, a large transmitter is needed for this transmission, which is rather bulky and heavy. This constrains usability as the requirements for a laptop/ PC and a transmitter is restrictive. Moreover, the data can only be transferred through cables as there is no Bluetooth network which again impedes the usage of the system. Another problem is that this system can only be used by a patient who has been assigned through a referral system from the healthcare professional office. If both the patient and medical professional are enrolled, then the data can be viewed by through the server portal. Overall, this system is costly and requires a subscription which is not clear in terms of whether this is monthly or annually.

2.5 Conclusion of Diabetic Monitoring Devices

Although innovation is continually developing, and there is a continuing exertion being made to make applications and sites of use for the diabetic group. The proof introduced here, demonstrates that it is hard to discover a system that genuinely satisfies the diabetic's requirements. One of the requirements being a system which truly fulfils the diabetic's need for a sole place where all vital data can easily and quickly be recorded. Besides, it is essentially difficult to find any sort of system which gives the medicinal group ongoing access to such information in real time. This goal of this project is to provide patients with the capacity to login and share all their vital data with their medical team, in an organised and effective way.

Equally critical, is the medical professional's ability to give feedback to the patient and mediate in case of a medical complication or emergency. These features will be this web portal main functions.

Below is a table showing the main differences between the above systems and the proposed system:

Systems	Special Hardware Required	Subscription required	Connects to Doctor
Our Proposed e-Tool Kit	No	No	Yes
Diasend	Yes	Yes	No
Diabetes Tracker UK	No	Yes	No
Lifescan	Yes	Yes	No
Accu-Chek	Yes	Yes	No

Table 1– Systems Comparison

3. PROPOSED SOLUTION

The goal of the research study is to create a worldwide available e-toolkit that will overcome the current issues that the existing diabetes monitoring system has and will enable the medical professionals and patients to communicate.

The e-toolkit is expected to be user friendly and accessible to patients allowing them to record data such as individual risk factors, which are important in monitoring their condition. The patient will have a portal that can record consistent and real time readings, which can be accessed from home, work or from any

other location desired using a mobile phone. Patients will no longer have to experience the inconvenience of carrying a journal and writing notes. The portal will be accessed by the patient's doctor allowing the vital patient/doctor interaction.

Doctors and patients can interact using the direct messaging system. As the system requires a log in, procedures will be implemented to allow regular monitoring with the aim of the portal becoming integrated to the patients' daily routine. Additional features can potentially involve: the display of graphs, charts, notifications to high-risk patients as well as enabling data from lab results to be recorded and shared. The e-toolkit will suggest treatments to the doctor according to the readings, which the patients have already input.

3.1 Survey Design, Implementation and Feedback Analysis

This section covers the structure of the conducted survey, the diabetic patients and medical professional feedback as well as the analysis of the collected data. We started by giving a brief introduction about the approach which has been followed to implement the conducted survey. Furthermore, we explained the ethics standards that were agreed upon among all parties, being the diabetic patients, medical professionals and researchers.

3.2 Survey Research Process

The survey information that was gathered will be presented in this order: first, a blank copy of the issued survey (survey development), then a presentation of the results (survey administration) and finally an analysis of those results. This process is explained in Angela M book titled "Survey methods for educators: Analysis and reporting of survey data".

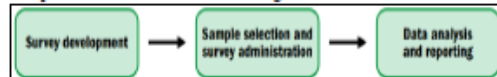The process is demonstrated in the figure shown below.

Figure 1 – Survey Process - Source: (Angela M. Pazzaglia, 2016)

3.2.1 Medical Professionals Survey Structure

There were two surveys for this research study, one for the medical healthcare professionals and the other for diabetic patients.

The medical healthcare profession survey had particular questions such as "What do you consider to be the benefits of this proposed system".

Figure 2 – Example of Medical Healthcare Professionals Survey

Also, there was questions with open answers to allow suggestions, for example, "Would you use this application to Share information and comments with other health care professionals, if such a feature was available?".

This allows the research study to have a variety of answers to the questions, improving the research qualitative data.
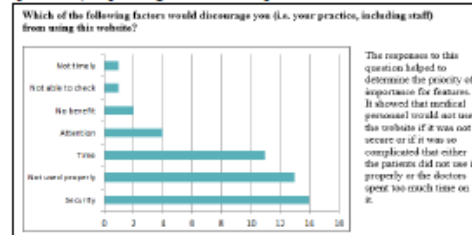
Figure 3 – Example of a question and a corresponding illustrated as a bar chart

3.2.2 Medical Professionals Survey Results Analysis

The results of the medical healthcare profession survey have been helpful for this research study, as 98% of the healthcare professionals said they would use the proposed e-toolkit and shared enthusiasm about the prospect. Some responders did give suggestions with the questions that had open answers. For example, with question 7 "If a feature were included that allowed your office, or the patients themselves, to input medical test results (for monitoring / trend tracking purposes), would you use/ support such a feature?". The overall recommendations to this were that it would be supported if it were restricted to input, from medical personnel only and would be helpful to import the patient's medical lab results (A1C, cholesterol, etc) so that it can be recorded in one convenient location.
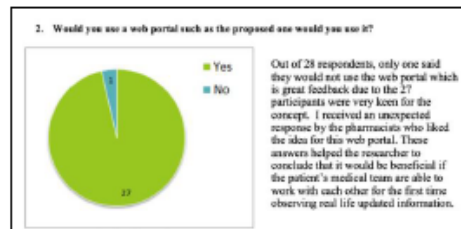
Figure 4 - Example of a question and a corresponding graph showing the results

3.2.3 Diabetic Patients Survey Structure

The diabetic patient survey questions were aimed at their personal health, such as "How long have you had diabetes?" and "Do you find it a hassle to keep track of your readings?". Also, questions aimed about current applications "What features do you dislike about your current system/monitoring e-Toolkit?" and open questions about possible suggestions/recommendations about a system they would use.

Figure 5 - Example of the diabetic patient survey

3.2.4 Diabetic Patients Survey Results Analysis

The results of the diabetic patient survey showed that 37.5% of the patients had diabetes for more than 20 years. 15 patients (with another 13 that answered "maybe") said they would use the proposed system which links the medical relationship between patients and professions. This finding is a boost of confidence that diabetics would be interested in this proposed e-toolkit.

In the data, we found out that a staggering 65% of the diabetic patients confirmed that it is a hassle to keep track of their own readings. This response emphasised on a point that there are two classes of diabetics; those who take the condition seriously and stay on top of their health with exercise and diet. Then there are those who disregard most medical advice, continue to have a lifestyle as they did before their diagnosis, showing not much care for their health. The second type of diabetics are the ones who will most likely not want to participate in such a survey or use the suggested application.

From these survey results, there are some side-by-side comparisons from the same questions to show the similarities or differences in the results. The question below is the most important analysis of the question, "which data should this e-toolkit monitor?".

This comparison shows doctors and patients agreed which health factors should be recorded. Both parties also decided the importance and priority of this documentation. Each group indicated the same fields for their top six: weight, blood glucose, insulin intake, diet, carbohydrate consumption and symptoms.

Figure 6 – Bar chart showing answers for what the website should monitor

3.2.5 Survey conclusion and analysis

Occasionally, whilst doing a research study like this one proposed, some aspects may be missed like the impact on the people themselves throughout their lives, physically and emotionally. Whilst collecting these findings many astonishing ideas were discovered from the diabetic community. Throughout the depths of the numerous subjects in this research it is repeated and accentuated that diabetes is an impactful disease affecting millions of people's everyday wellbeing and functioning of people globally. However, creating a web system to try support a huge community is neither easy nor simple. One of the most noteworthy results of this study was the awareness of the lack of communication between the medical diabetic personnel and the diabetic patients. This is mostly reported in certain areas like North America more than in the United Kingdom.

This viewpoint has been helpful leading to questions in need to be answered and could be involved in this project.

A completely distinct research study can be created and developed looking at different features of this problem. Even though it could prove to be a fascinating theme to explore, with a definite impact on the possible triumph of this web portal, it has shown to be too widespread which could leave this study steering a different direction to its original purposes.

During this project there were many findings discovered from the excruciating disease are as following:

Numerous of the diabetics from the consensus suggested and was onboard with the statement of diabetics can be split into two different groups:

Group 1	Group 2
This group takes into account of the diagnosis and try their hardest to tackle the disease by: Paying attention to their symptoms, maintain records, test frequently. This makes them feel they can control their health and it can change at the fate of their diet and exercise. As a group they particularly expressed some feelings of doubt and distrust to their healthcare professions including doctors. They have stated that the doctors have a lack of communication to the patients and are not sympathetic towards their treatment. A theme has been repeated and exaggerated in saying the doctors are not helpful and did not impact the diabetic patient's health in tackling the illness instead they felt like the doctors made the situation worse. Patients from group 1 were fearful of civil liberties being revoked, including driving, as the doctor can do this if they wish. Therefore, this group did not agree with the unlimited and unrestricted access of the web portal with their records and information.	This group seems to be the people that the medical professionals are most concerned about, they are more likely the people to benefit from this system, yet it actually turns out would be the least group to actually use it. These are the patients that appear to not to deal with their disease. They are patients who use their medication or insulin intake all the time as they do not want to change their diet or daily lifestyle, diabetic in this group will probably be hospitalised or to experience major consequences for not looking after themselves, such as limb amputation.

Table 2 – The two different groups of diabetics from the survey

The two different groups of diabetics are not representative of the population as there may be patients that are a cross over of the two groups, one that may exercise regularly but have a poor diet. However as there is not much data to explain fully the anomalies that are present. In the survey data the distribution ratio of group one to group two of the table shown above were 7:13.

The diabetic patients this proposed e-tool kit is targeted are those who do not actively track and record their vital readings, those who are less worried about their health, who do not understand fully their disease and how to control/limit negative impacts. These characteristics are similar to those in group two shown in Table 2.

Despite the worries by diabetics (from the survey results) in the discussions of the e-toolkit potential tracking and alert function, both patients and doctors voted in favour of such a feature. The alerts would need to include an on/off feature, as doctors did not want to have to receive alerts for all their patients. The consensus indicated that this would be a very useful feature, but only for certain, high-risk patients. Additionally, the parameters that would initiate such an alert would have to be adjustable. Both groups insisted that this was essential, as ranges vary depending on the individual and on the individual's specific set of circumstances at a given time. For instance, a "normal" blood

glucose range for one person may be drastically different from their own range two years ago or that of another person.

For the medical community, the biggest concern that would potentially prevent use of this site would be if it were not secure. Medical professionals would be professionally and ethically barred from using a site that is easily hack-able or does not take appropriate security precautions. Their second highest concern was that the site would be too complicated for patients to use correctly and because of this the data may be corrupted. The third highest stated concern was that the site would be too complex for doctors to use easily thus being too time consuming and impractical.

Important disclaimer: as was pointed out by some of the diabetic survey respondents, the people who were able to be contacted and who were willing to participate in this research were primarily (with the exception of four hospitalised patients) the type of patients who classify into group 1 (Mathieu-Fritz, Alexandre & Guillot, Caroline., 2017). These are the proactive, go-getter diabetics. The patient responses received from this group are slated towards the feelings and viewpoints of this group and do not necessarily represent a good cross-section of all diabetics. Unfortunately, this reality was pointed out afterwards and if this project were expanded further in-depth, a thorough cross-sampling of diabetic participants would most likely have to be secured with the assistance of the medical community.

A surprising two-thirds of the diabetics surveyed said that they do not find it difficult to keep track of monitoring and recording their health factors. Again, however, this information must be viewed in the context that the people responding are in the group 1. It is hoped an e-toolkit such as this will help to convert more people over from group two into group one. By making it easier and simpler for them to maintain their records and by providing them with more support in this task, from the doctor's office.

The result of the conducted survey shows that existing monitoring systems do not support knowledge sharing and dissemination between medical staff and diabetic patients. Therefore, an e-toolkit has been suggested to overcome the issues mentioned above.

4. Final Product and Functionality

This section shows the final product that was designed and implemented to meet the above requirements.

4.1 Development and Technical Background

After a comprehensive research on the tools and techniques to be used to implement the proposed E-Tool Kit, it was concluded that it is best to use PHP to develop the front end of the e-tool kit and MySQL for the Database.

- PHP is a widely-used open source general-purpose scripting language that is especially created for web development and can be embedded into HTML. (php.net, 2018)
- MySQL (Structured Query Language) is known as "the world's most used" relational database management system (RDBMS). It is used in many high-profile products including Wikipedia, Google, Facebook, and Twitter. MySQL is open source for private use or for products created with MySQL which, themselves, will be open source. Commercial edition licenses are available for non-open source products (MySQL, 2019)

4.1.2 Conclusion

MySQL is the best choice for this project's purposes. In addition to being free of charge, it is easy to use, highly compatible with the chosen PHP language, and offers a wide selection of support options and resources. It also offers a sufficient amount of storage

space, has a high speed and a renowned, reliable, intuitive system. (Darwish F, 2012)
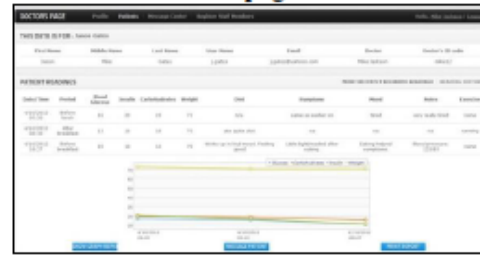
4.2 View of doctor's page on e-toolkit

Figure 7 – View of Doctors page

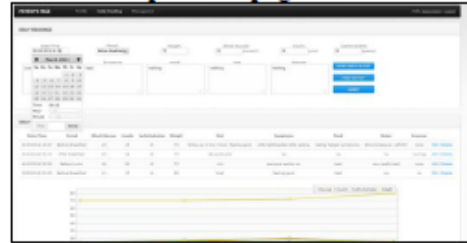### 4.3 View of patient's page on e-toolkit

Figure 8 – View of Patient page

5. PRODUCT EVALUATION AND FUTURE WORK

Whist developing this research study it became clear that the original projected objectives and features were too grand and ambitious with the limited time and resources obtainable. It is easy to get carried away with researching and developing all of the opportunities the web portal can provide and disregarding the website's fundamental main necessities and ease of use. Some of the ambitious features could have been as successful if the website was designed by an experienced team of expertise developers however this project was created by a single student and cannot obtain the same level with limited experience and time.

Upon the improvement of the e-tool kit's appeal and flexibility that can be foreseen in the future, the section below addresses some features that can be considered in potential upcoming versions of the e-tool kit.

- *Registration of multiple doctors*- If there was a feature that permitted the patients to register more than one healthcare specialist this would allow everyone having access to the same up-to-date information, allowing to input and update the records and the ability to communicate with the same channel. The healthcare specialists could include from the patient's general practitioner, dietician, dentist or the pharmacist.
- *Field importation of data* - the web portal could have fields for laboratory test results or information from the medical file inputted by the doctors. For instance, HbA1c and cholesterol readings can be projected as easily as the insulin and blood glucose readings as seen currently. It would be predominantly stored to be seen by other medical personnel.
- *Emergency alerts* - to gain the immediate attention for the doctor if a specific range has been exceeded. However, if the alerts are not serious doctors have the option to switch this feature on and off and adjust the notification criteria.

- *Alteration of chart parameters* - trends need to be organised so visualisation of the features is crucial against a timeline or in comparison of other features.
- *Field importation for Type-1* - Patients - this group needs to be provided with fields for readings, including insulin pump and basal and bolus insulin.

Currently in the web portal, there are some field headings like 'DIET' and 'NOTES', where all the data information has been stored although in the future, they could be additional field for specific factors from injuries, menstrual cycle, body temperature to meal calculation, medications and infections, and many more. This can prove to be very useful as this information can be tracked and used to predict trends. This list of factors was produced in the survey comment sections suggested by the patients.

The following survey was given to evaluators to ensure that all parts of the e-toolkit were accessed and inspected:

Figure 9 – Example of survey assessing the e-toolkit

To summarise, our analysis proves that it is hard to find a relevant system that genuinely satisfies the diabetic's requirement for a solitary place where all vital data can be recorded easily and quickly. The goal for this project is to provide patients with the capacity to login and share all their fundamental data with their medical team, in an organised and effective way.

Although technology is constantly evolving, and there is an on-going effort being made to create applications and e-Toolkits for use by the diabetic community, the evidence presented shows that it is difficult to find a system that truly fulfils the diabetic's need. Furthermore, it is virtually impossible to find any kind of system which provides the medical community with real time access to such data. This project aims to provide patients with the ability to login and share all information directly with their doctors, in a structured and efficient manner. Equally important, is the doctor's ability to provide feedback to the patient and intervene in the event of a medical complication or emergency. These features will be this e-toolkit's primary functions.

6. REFERENCES

[1] accu-chek.co.uk. (2019, 01 04). accu-chek.co.uk/. Retrieved from Accu-Chek Smart Pix device reader product specifications: https://www.accu-chek.co.uk/help/data-management/smartpix

[2] Alan J. Sinclair and Simon C.M. Croxson. (2009). Diabetes Mellitus. In A. J. Croxson, In Brocklehurst's Textbook of Geriatric Medicine and Gerontology (SEVENTH EDITION) (pp. 760-774). Philadelphia: Philadelphia, PA: Saunders/Elsevier

[3] amazon.co.uk. (2019, 01 24). ACCU CHEK Smart Pix. Retrieved from amazon.co.uk: https://www.amazon.co.uk/ROCHE-DIABETES-CARE-ITALY-3806040/dp/B001O1FQNA

[4] American Diabetes Association. (2018, 04 02). American Diabetes Association. Retrieved from http://www.diabetes.org: http://www.diabetes.org/diabetes-basics/

[5] Angela M. Pazzaglia, E. T. (2016). Survey methods for educators: Analysis and reporting of survey data . Education Development Center, Inc. .

[6] Darwish F. (2012). Diabetes Management Website. University Of Huddersfiled.

[7] Diabetes UK. (2018, 12 12). Diabetes UK Tracker app launched. Retrieved from Diabetes.org.uk: https://www.diabetes.org.uk/about_us/news_landing_page/android-version-of-diabetes-uk-tracker-app-launched

[8] glucosemeters4u.com. (2019, 01 30). Retrieved from glucosemeters4u.com: http://www.glucosemeters4u.com/accu-chek-smart-pix.htm

[9] Larme AC1, P. J. (1998). Attitudes of primary care providers toward diabetes: barriers to guideline implementation. Diabetes Care, 1391-6.

[10] Mathieu-Fritz, Alexandre & Guillot, Caroline. (2017). Diabetes self-monitoring devices and transformations in "patient work". (Vols. (Vol. 11, N°4)). S.A.C. (http://www.socanco.org).

[11] myonetouch.co.uk. (2019, 01 03). OneTouch Reveal. Retrieved from myonetouch.co.uk: https://www.myonetouch.co.uk/OneTouchReveal

[12] MySQL. (2019, 01 04). Retrieved from MySQL: https://www.mysql.com

[13] N. H. Cho, J. E. (2018). IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Research and Clinical Practice, 271–281.

[14] National Diabetes Audit Report. (2017, March 14). Retrieved from nhs.uk: https://digital.nhs.uk/data-and-information/publications/statistical/national-diabetes-audit/national-diabetes-audit-report-1-findings-and-recommendations-2016-17

[15] NHS Croydon Clincical Commisioning Grp. (2018, 04 03). Health Apps Library. Retrieved from croydonccg.nhs.uk: https://www.croydonccg.nhs.uk/get-involved/PublishingImages/Pages/Publications/Appendix%208.pdf

[16] php.net. (2018, 04 20). what is php? Retrieved from php.net: http://php.net/manual/en/intro-whatis.php

[17] Till Seuring, Olga Archangelidi, Marc Suhrcke. (2015). The Economic Costs of Type 2 Diabetes: A Global Systematic Review. . Pharmaco Economics.

[18] Van Der Ven, W. K. (2003). The confidence in diabetes self-care scale: psychometric properties of a new measure of diabetes-specific self-efficacy in Dutch and US patients with type 1 diabetes. American Diabetes Association, 713-718.

[19] webmd.com. (2019, jan 20). Diabetes Care: Managing Your Time When You Have Diabetes. Retrieved from webmd.com: https://www.webmd.com/diabetes/features/diabetes-time-management#1

[20] who.int. (2019, Jan 20). Diabetes. Retrieved from who.int: https://www.who.int/news-room/fact-sheets/detail/diabetes

9.4   Appendix D: Invitation Letter

University of Salford

University of Salford MANCHESTER

Participant Invitation Letter

Improving diabetes management between patients and healthcare professionals

Dear Sir/Madam,

I am writing to you to ask you to consider in taking part in a survey from a researcher at University of Salford. The research study aim is to create a universal web system, not restricted by location, income level or the monitor type the patient uses, to support diabetes educational and management. It will carry out numerous functions including the ability to record daily glucose readings, dietary habits, insulin intake and other information. This can then be accessed, in real time, by the patient's health care providers, to improve upon accurate and timely communication.

I would like diabetic patients in the UK to fill out a short survey telling us what feature you would like to see on such a website and how it could be designed to best suit your needs and concerns.

Participation is voluntary. You may withdraw from the study at any time and remain anonymous which is stated on a consent form which is required to sign before the study. The research has ethical approval from the University of Salford.

If you are interested in participating or require any further information about the research study, please don't hesitate to contact me.

Thank you for your help

Yours Sincerely

Fahad Darwish

PostGrad Student

f.darwish@edu.salford.ac.uk

9.5   Appendix E: Information Sheet (Page 1)

PARTICIPANT INFORMATION SHEET:
Improving diabetes management between patients and healthcare
professionals (Using Data Mining Techniques)
Lead Researcher: Fahad Darwish
Contact: f.darwish@edu.salford.ac.uk

_____

We would like to invite you to take part in this research survey which is being run by the study researcher Mr F Darwish from the University of Salford. This participant information sheet (PIS) will help you decide if you'd like to take part in the study explaining the process, the benefits of your involvement in the study to the medical field, any possible risks, your rights and confidentiality with the right to withdraw at any time and contact details.

The purpose of this study is improving the diabetes management between patients and healthcare professionals to possibly make the service in the medical field to become more efficient and up to date with today's technology. This will take you approximately 15 minutes to complete and your participation in the study is completely voluntary and can withdraw at any time.

If you agree to take part in this study, you will be asked to sign a consent form which will be provided as well.

WHAT IS THE PURPOSE OF THE STUDY?

This research study's primary focus is to create a universal web system, not restricted by location, income levels or the monitor types, to support the educational and management of diabetes, most common disease in the medical field. The study will carry out numerous functions including the ability to record daily glucose readings, dietary habits, insulin intake and other information. This can then be accessed, in real time, by the patient's health care providers, to improve upon accurate and timely communication.

WHAT WILL MY PARTICIPATION IN THE STUDY INVOLVE?

You have been chosen to participate in this survey study as you are a diabetic patient or healthcare profession, thank you for taking the time to read this information sheet and hope you can help this study. It is a survey with simple Yes/No answers to circle with the room for explanations and possible suggestions which could be considered to make the web portal accurate and up-to-date in managing this growing illness. There is only one survey that you can fill in online with small amount of time needed, with complete anonymity and confidentiality of your information. Your participation in this study is entirely voluntary and you can withdraw at any time. You are free to omit any question and have no obligations.

WHAT ARE THE POSSIBLE BENEFITS AND RISKS OF THIS STUDY?

The study alone will not directly support you at this present time but with your assistance and knowledge, it could benefit the researcher greatly to gain better insight while creating the web portal, which potentially could allow the diabetic patient's health care

Improving diabetes management between patients and healthcare professionals

Page **1** of **2**

9.6   Appendix E: Information Sheet (Page 2)

provider to access the patients updated readings with greater efficiency. This increases the healthcare professionals and diabetic patients' communication as the anomalies in the readings could send urgent alerts or notifications to the healthcare profession and they can examine the problem and possibly save the patient's life if in danger.
The study results can be developed in less time than other data collection methods like paper based, therefore creating faster and valid help for the web portal. The study can be easily accessed via online on computer or mobile devices with no need to use an email address which produces a less likelihood of invading privacy with greater anonymity. We do not expect there are any risks to you in participating in the study.

If the information on page 1 has interested you and considering taking part, please read the additional information below.

WILL MY PARTICIPATION BE CONFIDENTIAL?

Yes, we will follow the ethical and legal practices. The data will be on the encrypted USB flash drive device and fully secure and password protected with only access by the researcher. All published data will be written in a specific way to disguise your identity unless prior consent has been obtained and allowed to release the identity of that individual from the study. All information which is collected about you during the course of the research will be kept strictly confidential. The University of Salford and the main researcher will keep identifiable information about participants until a minimum of one year past the end of the study, and then its retention will be reviewed with a view to permanent disposal or long-term archiving for future research. The data on the USB flash drive can be securely erased by reformatting which destroys everything on it and then physically destroyed.

All data from this survey (including the afore-mentioned transcripts and so forth) obtained during this research project will be subject to the provisions of the Data Protection Act 2018 (http://www.legislation.gov.uk/ukpga/2018/12/contents/enacted)

WHAT ARE MY RIGHTS?

This participation of this study is voluntary with the freedom to decline in participating, or to withdraw from the research at any time. As a participant you have the right to access information about you collected as part of the study and will be contacted if any new information comes to light which could disrupt you or your confidentiality.

WHO DO I CONTACT FOR MORE INFORMATION OR IF I HAVE CONCERNS?

If you have any questions, concerns or complaints about the study at any stage, you can contact:
the lead researcher Fahad Darwish - f.darwish@edu.salford.ac.uk.
Or his Supervisor Dr Mo Saraee – m.saraee@salford.ac.uk

9.7   Appendix F: Consent form

University of
Salford
MANCHESTER

PARTICIPANT CONSENT FORM

To be completed by the participant. If the participant is under 16, to be completed by the parent / guardian / person acting *in loco parentis*.

- I have read the information sheet about this study
- I have had an opportunity to ask questions and discuss this study
- I have received satisfactory answers to all my questions
- I have received enough information about this study
- I understand that I am / the participant is free to withdraw from this study:
    - At any time (until such date as this will no longer be possible, which I have been told)
    - Without giving a reason for withdrawing
    - (If I am / the participant is, or intends to become, a student at the University of Salford) without affecting my / the participant's future with the University
- I understand that my research data may be used for a further publication in anonymous form, but I am able to opt out of this if I so wish, by ticking here. ☐
- I agree to take part in this study

Signed (participant)	Date
Name in block letters	
Signed (parent / guardian / other) (if under 16)	Date
Name in block letters	
Signature of researcher	Date
This project is supervised by:	
Researcher's contact details (including telephone number and e-mail address): Fahad Darwish - f.darwish@edu.salford.ac.uk Mobile: 0044 7590 53 7017	

Spring 2017/18

162

9.8   Appendix G: Ethics Approval Certificate

University of Salford
MANCHESTER

Research, Innovation and Academic Engagement Ethical Approval Panel

Doctoral & Research Support
Research and Knowledge Exchange,
Room 827, Maxwell Building
University of Salford
Manchester
M5 4WT

T +44(0)161 295 5278

www.salford.ac.uk/

20 August 2018

Fahad Darwish

Dear Fahad,

RE: ETHICS APPLICATION STR1718-48: Towards Intelligent Diabetes Management: A Data Mining Approach

Based on the information you provided, I am pleased to inform you that your application STR1718-48 has been approved.

If there are any changes to the project and/ or its methodology, please inform the Panel as soon as possible by contacting S&T-ResearchEthics@salford.ac.uk

Yours sincerely,

Dr Anthony Higham
Chair of the Science & Technology Research Ethics Panel

9.9 Appendix H: e-Toolkit Acceptance Testing

Example of the Acceptance Testing used

Test Number	MoSCoW requirements	Results achieved	Has it passed?
1	The e-Toolkit must allow diabetics to enter all vital health information.	The e-Toolkit has the feature whereby patients are able to input all their vital readings	Passed
2	An ID must be assigned to every physician, nurse and medical staff member.	The feature has been successfully implemented every doctor/patient has an ID assigned to them.	Passed

9.10   Appendix I: e-Toolkit Evaluation Survey

There were many e-Toolkit Evaluation forms that were issued to the patients and medical team the below is a sample of one of them, those forms helped greatly in the verification and validation of the e-Toolkit.

Please go to http://fahad-darwish.com/uniwebsite/		
Please pick one of the followings which is more appropriate to you.		

1.   Health Care Provider
    2.   Patient with Diabetes
    3. Other		

To begin, please follow the instructions		

1) Register on the e-Toolkit as a health care provider
    2) Please input your email address as an email will be sent to you for verification.
    3) Please insert the following details		

Inputted Username:	**Inputted Password:**	**ID chosen #**
4) Log in and view your profile		
5) Register a staff member **6) Does the new member registered shown at the bottom of the page?**		

9.11   Appendix J: e-Toolkit Validation Survey

There were many e-Toolkit validation forms that were issued to the patients and medical team the below is a sample of one of them, those forms helped greatly in the verification and validation of the e-Toolkit.

Please go to http://fahad-darwish.com/uniwebsite/		
Please pick one of the followings which is more appropriate to you.		
4. Health Care Provider 5. Patient with Diabetes 6. other		
Does the e-Toolkit, does it meet its purpose	Yes	No
Do you think this e-Toolkit beneficial ?	Yes	No
How did you access the e-Toolkit? Phone, Laptop, tablet or others		
Was there part of the e-Toolkit that did not work for you?		
What did you like the most about the e-Toolkit		
Any suggestion for improvement?		
More comments?		

9.12   Appendix K: Proposed SECI model (Scenario Validation)

Our Framework revealed that the most common issues on the Citalopram Dataset were:		
a) **"Severe Headache"** b) **"Caused hormonal imbalance"** c) **"Tiredness / feeling tired"**		
Are those side effects / concerns correct?	Yes	No
If yes are you happy for us to publish those on our e-Toolkit?		
Any additional comments?		

Appendix L – Official Documents generated