



**UNIVERSITY OF NAIROBI**

**Designing an M-learning system for Community Education and information on HIV  
and AIDS in Kenya**

**A Project Presented to the School of Computing and Informatics  
University of Nairobi**

**In Partial Fulfilment  
Of the Requirements for the Degree Masters of Science in Computer Science**

**By**

**Oduor Fredrick Oluoch  
Registration Number: P58/61703/2010**

**August 2012**

**Supervisor: Dr. Robert Oboko**

## DECLARATION

This thesis is my original work and has not been presented for a degree in any other University.  
This thesis has been submitted for examination with my/our approval as university supervisor(s).

Student' Name

Date

Signature

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Supervisor's Name

Date

Signature

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

## **DEDICATION**

To my parents,  
For the hard work, they did raising and schooling me.

## **ACKNOWLEDGMENT**

I would like to acknowledge all teachers that have imparted academic and interpersonal knowledge, skills and experience that has made it possible for me to attain this level of education and Dr. Robert Oboko for his insight and motivation that made it possible for me to complete this project.

## **ABSTRACT**

The aim of the study is to determine how m-learning can be used to create an interactive learning environment that can help improve awareness on HIV in Kenya by improving the contextualisation of information and providing a seamless platform for discussions on issues relating to HIV AIDS. The specific objectives that are addressed include to identify the pedagogical and technical challenges that are likely to be faced in using the m-learning model and the e-platform for discussing HIV AIDS related issues; to determine if the m-learning model influences gain of knowledge on HIV AIDS; to personalise the information presented to individual model users in terms of their learning interests and to develop an m-learning model that allows for the aggregation and contextualisation of information on HIV AIDS to the specific needs of the information seeker. The study involved developing a prototype m-learning system, use of the system by selected 50 truck drivers and reporting on selected aspects of the system. The data collection involved administration of questionnaire and acquisition of system data on the use of the prototype. The findings reveal technical and pedagogical challenges such as internet access and use of technical words. Additionally, the findings show that m-learning model influences knowledge gain, allows for personalisation of information and integration of information from different sources. These findings have multiple implications on policy, practice and research. Importantly, the findings show the potential offered by m-learning in community education on HIV AIDS and highlight the need for improved dialogue and support for learner autonomy within m-learning.

## **LIST OF ABBREVIATIONS**

Pife: Performance in first exams

Pise: Performance in second exam

FCS: Feed categories selected

FCI: Feed categories Ignored

FCA: Feed categories adopted

FCC: Feed categories checked

Rel\_comm: Relevance of recommendations engine

Rel\_reccactual: Actual relevance calculated from users' interaction with system

HIV: Human immunodeficiency syndrome

AIDS: Acquired immune deficiency syndrome

VCT: Voluntary counselling and testing

## TABLE OF CONTENTS

<b>DECLARATION .....</b>	<b>I</b>
<b>DEDICATION .....</b>	<b>II</b>
<b>ACKNOWLEDGMENT .....</b>	<b>III</b>
<b>ABSTRACT .....</b>	<b>IV</b>
<b>LIST OF ABBREVIATIONS .....</b>	<b>V</b>
<b>LIST OF FIGURES.....</b>	<b>VIII</b>
<b>LIST OF TABLES.....</b>	<b>IX</b>
<b>CHAPTER ONE: INTRODUCTION.....</b>	<b>1</b>
1.1. Background .....	1
1.2. Problem Statement .....	2
1.3. Research aim .....	2
1.4. Research Objectives .....	2
1.5. Rationale .....	2
1.6. Significance .....	3
1.7. Keywords.....	3
<b>CHAPTER TWO: LITERATURE REVIEW .....</b>	<b>4</b>
2.1. Theoretical Framework .....	4
2.2. HIV and AIDS Management in Kenya .....	5
2.3 M-Learning .....	6
2.4. Support for Pedagogy .....	6
2.5. M-Health Applications.....	7
2.6. Subject-Based Mining Algorithms.....	8
2.6.1. Clustering.....	9
2.6.2. Artificial Neural Networks.....	9
2.6.3. Vector-Based Algorithms.....	9
2.7. Personalisation .....	10
2.8. Web Syndication.....	11
2.9. Proposed Solution .....	13
<b>CHAPTER THREE: METHODOLOGY.....</b>	<b>15</b>
3.1. Introduction.....	15
3.2. Prototype Development Methodology .....	15
3.3. Research Design .....	15
3.4. Research Approach .....	15
3.5. Data Collection .....	16
3.6. Data Analysis.....	16
3.7. Validity and Reliability.....	16
3.8. Ethical Considerations .....	16
<b>CHAPTER FOUR: SYSTEM DESIGN AND IMPLEMENTATION .....</b>	<b>18</b>
4.1. Introduction.....	18
4.2. Data Set.....	18
4.3. System Design .....	18
4.4. System Modules.....	18
4.5. Prototype Testing.....	19
4.6. Prototype Deployment .....	19
4.7. Limitations of the Design.....	19

<b>CHAPTER FIVE: RESULTS AND DISCUSSION.....</b>	<b>20</b>
5.1. Introduction.....	20
5.2. Results.....	20
5.3. Findings .....	24
5.4. Discussion.....	27
5.4.1. Objectives.....	27
5.4.2. Framework .....	28
5.4.3. Methodology .....	29
<b>CHAPTER SIX: CONCLUSION AND RECOMMENDATIONS.....</b>	<b>30</b>
6.1. Introduction.....	30
6.2. Conclusion .....	30
6.3. Recommendations.....	30
6.3.1. Recommendations for Policy and Practice.....	30
6.3.2. Recommendations for Research.....	31
<b>REFERENCES .....</b>	<b>32</b>
<b>APPENDIX .....</b>	<b>35</b>
Appendix A: Sample Questionnaire .....	35
Appendix B: Quiz .....	37
Appendix C: User Interface Screenshot.....	39
Appendix D: User Feed Preferences .....	40



## LIST OF FIGURES

Figure 1: Moore's Transaction Distance Theory (Source: Giossos, Koutsouba, Lionarakis and Skavantzios, 2012)	4
.....	4
Figure 2: Proposed Solution .....	13
Figure 3: Distribution of Challenges Faced by participants .....	24

## LIST OF TABLES

Table 1: Summary of Key variables .....	20
Table 2: Integration Frequency Table.....	21
Table 3: Usability Frequency Table.....	21
Table 4: Relevance Frequency Table.....	22
Table 5: Comments Frequency Table.....	23
Table 6: Challenges Faced By Users .....	23
Table 7: Paired correlations pife and pise.....	24
Table 8: Paired Sample t-test pife and pise.....	25
Table 9: One-Sample t-test test-value 5.....	25
Table 10: One sample t-test integration, usability, relevance, rel-recc test-value 3.....	26
Table 11: Paired sample t-test FCI-rel-reccactual .....	26

## CHAPTER ONE: INTRODUCTION

### 1.1. Background

The HIV epidemic in Kenya is categorised as generalised. This means that the epidemic affects all sectors of the population. It is however important to note that the prevalence of HIV tends to differ across geographical locations, ages and gender. 80% of new infections in 2008 were transmitted in heterosexual sex whilst 20% in casual relations (Avert, 2011). Some groups are affected more than others are. High prevalence rates have been reported in sex workers, injecting drug users, gay men, truck drivers and cross-border mobile populations. Some of these groups' for instance gay men are marginalised within the Kenyan society. Homosexuality is both socially unacceptable and illegal in Kenya. This marginalisation compounds the difficulties that already exist in reaching HIV prevention, treatment and care. Additionally, marginalisation has been blamed for the low exploration of the extent to which HIV affects these groups.

HIV in Kenya disproportionately affects women. In 2008/9, HIV prevalence in Kenyan women was twice as high as for men. Women prevalence was 8% whereas men recorded a prevalence level of 4.3% (Avert, 2011). The disparities are greater in women aged between 15 and 24 who are four times more likely to be infected than men of their age. One of the factors attributed to this disparity is high rates of exposure to violent sexual exposure by the Kenyan woman. Statistics from 2003 show that almost half of Kenyan women have experienced sexual violence and one quarter of women aged between 12 and 24 lose their virginity by force. People infected through violent sexual contact are less likely to seek help and conventional information avenues due to fear of stigmatisation. Alternative, private and personalised information avenues are required to deal with the plight of this special group.

Even though the HIV prevalence in urban areas (8.4%) is higher than in rural areas (6.7%), more people living with HIV reside in rural Kenya. This is because 75% of the total Kenyan population live in rural areas (USAID, 2011). Due to historical factors such as poor resource allocation, rural areas are often characterised by poor access to infrastructure and social amenities. This compounds the challenges faced by social workers in managing and educating the large number of people living with HIV in rural areas.

Rapid social and technological change is a distinguishing feature of the society in the twenty first century. Advancements in technology have led to innovations and developments that allow for fast communication and information processing that are currently supporting new social patterns. In the modern society, communities are no longer based on only the geographical proximity as technology allows for global communities based on shared interests, work patterns and opportunities (Bridgland and Blanchard, 2005). Mobile communication and information technologies are vital to enabling the new social structure. Most researchers are of the view that the society is experiencing the first generation of truly portable information and communication technology (Abbad and Albarghouthi, 2011). Their assertion is supported by the advent of small portable mobile devices that provide telephone, internet, data storage and data management functionalities. In addition, the advent of smartphones that support mobile telephony, diaries, emails, word processing, spread sheets, data transfer and data storage has further improved the potential of mobile technology.

Communication and data transfer policies created by the m-technologies have the potential of reducing the dependence on fixed location for work and study. M-technologies can potentially revolutionise the way humans work and learn. However, there are challenges that are associated with mobile delivery. In general, people expect learning to be just in time, enough and tailor made for their learning needs. Understanding the

pedagogical challenges, actual use of mobile by the targeted audience and the nature of the learning problem are all vital in actualising the potential offered by m-learning.

### **1.2. Problem Statement**

HIV AIDS is a major killer in Kenya. In spite of efforts by different parties including the government, NGOs and institutions of learning aimed at dealing with the scourge it has persisted. Currently, there are different sources of information on HIV AIDS such as mass media, VCT centres, and internet and government booklets. The sources lack in aggregating and contextualising information to specific user needs. For instance, mass media often report on developments that have had adverse effects on the management of HIV whereas governmental and non-governmental organisations booklet often report on changes in the demographics and spread of the virus. Thus, though there are multiple sources of information on HIV AIDS, lack of specificity to the context of information seeker reduces the relevance of much of the information. Another issue is that most of the available information avenues do not offer support for discussions and sharing experiences. For people living with HIV AIDS, it is not enough to access information as sharing plays an important role in improving their psychological wellbeing. The existing avenues for instance social media do not offer the level of privacy needed to avoid stigma which is a major challenge in information dissemination for people suffering with AIDS. Thus, contextualisation of information to the specific needs of an information seeker and provision of a platform where people suffering from HIV AIDS (either directly or indirectly) can share their views without fear of stigmatisation are the two main problems that were addressed.

### **1.3. Research aim**

To determine how m-learning can be used to create an interactive learning environment that can help improve awareness on HIV in Kenya by improving the contextualisation of information and providing a seamless platform for discussions on issues relating to HIV AIDS. From the research aim, it is evident that the study is limited to Kenya and the use of m-learning to disseminate information on HIV.

### **1.4. Research Objectives**

The following are the research objectives:

- a) To identify the pedagogical and technical challenges that are likely to be faced in using the m-learning model and the e-platform for discussing HIV AIDS related issues.
- b) To determine if the m-learning model influences gain of knowledge on HIV AIDS.
- c) To personalise the information presented to individual model users in terms of their learning interests.
- d) To develop an m-learning model that allows for the aggregation and contextualisation of information on HIV AIDS to the specific needs of the information seeker.

### **1.5. Rationale**

Many Kenyans own mobile phones in Kenya. There has been an increase in the number of mobile users across Kenya. Mobile phone ownership transcends age, gender and location (urban or rural areas) (AudienceScapes, 2011). Therefore, mobile platform can be used to disperse information across Kenya. Companies such as Safaricom and Airtel having realised the potential offered by mobile telephony introduced m-banking which has so far been a success in Kenya (Safaricom, 2011b). M-learning can also use the extensive platform offered by mobile telephony to disseminate information that can help prevent and manage the scourge that is HIV. The success of m-banking in Kenya is the main motivation behind the study.

## 1.6. Significance

The study developed a model that allowed for the development of m-learning systems that are interactive and can be sustained using minimal resources. As a result, the study is important in the fight against HIV in Kenya and may help minimise rates of new infections and ensure that infected people have access to relevant information. Furthermore, the study looked into the pedagogical issues and the tools that can be used to deal with the challenges commonly faced in counselling, discussing and learning issues related to HIV in Kenya. The study findings highlight the potential offered by m-learning in social and community education. This awareness is vital in driving research and funding that is critical to the development of m-learning and its application to social problem.

## 1.7. Keywords

The following are some of the keywords used in the study:

- a) **HIV-Human immunodeficiency virus:** It is the virus responsible for causing
- b) **AIDS:** Acquired immune deficiency syndrome: This condition results in the loss of human disease fighting mechanisms.
- c) **M-learning:** The use of mobile devices in acquiring information
- d) **Ubiquitous:** A universal ever-present resource or tool
- e) **VCT:** Voluntary counselling and testing centres for HIV AIDS
- f) **Web syndication:** A process via which website material is availed to multiple sites
- g) **Subject based mining:** Use of an initiating individual or datum that is considered, to determine other data related to the initiating datum based on other information.

## CHAPTER TWO: LITERATURE REVIEW

### 2.1. Theoretical Framework

Transactional distance theory is an educational theory that defines the critical concepts of distance learning. It presents a definition of distance education which implies the separation of teachers and learners (Moore, 2007). Since its first appearance in publications (Moore, 1972, 1973), this theory has influenced numerous researchers and practices. Many scholars praise it as a classical and all-encompassing theory of distance learning (Gokool-Ramdoos, 2008; Saba, 2005) and view it as a major contribution to the field of distance education.

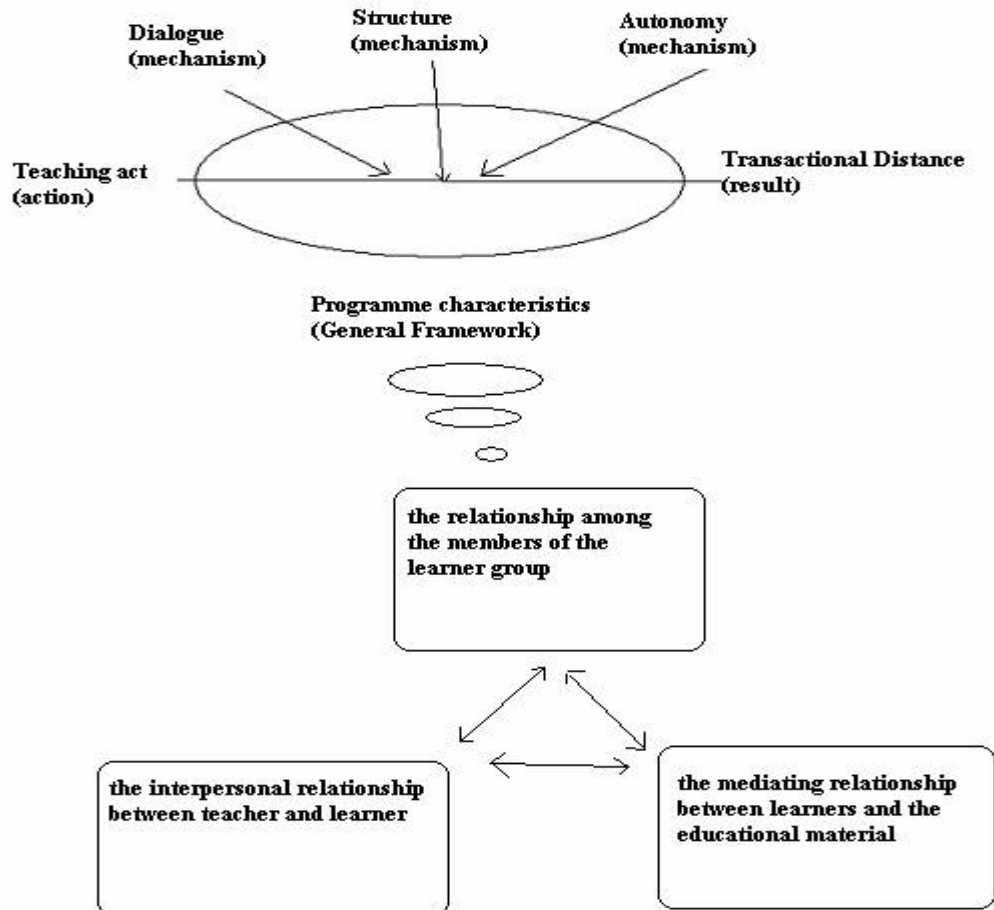


Figure 1: Moore's Transaction Distance Theory (Source: Giossos, Koutsouba, Lionarakis and Skavantzios, 2012)

Transactional distance theory is defined by the fact that distance is considered not only as geographic separation but also (and more importantly) as a pedagogical concept (Moore, 1997). As a result, the theory enables the inclusion of both types of education, that is, “a program in which the sole or principal form of communication is through technology” and where “technology-mediated communication is ancillary to the classroom” (Moore 2007, p. 91). This is especially important for mobile learning because mobile devices sometimes enter the school setting (Tatar, Roschelle, Vabey and Pennuel, 2003) as an ancillary element but mostly they extend beyond the classroom to non-traditional, informal, and non-institutional settings. The inclusive nature of transactional distance theory and its applicability and flexibility illustrates its important contribution to the framework for mobile learning.

This theory was derived from the concept of “trans-action,” which is considered by many scholars to be the most evolved level of inquiry, compared to self-action and inter-action (Dewey and Bentley, 1946), and “the interplay among the environment, the individuals and the patterns of behaviours in a situation” (Boyd and Apps, 1980, p. 5). Thus, transactional distance is defined as the “interplay of teachers and learners in environments that have the special characteristics of their being spatially separate from one another” (Moore 2007, p. 91). In short, transactional distance is the extent of psychological separation between the learner and the instructor (Shearer, 2007).

The transactional distance is controlled and managed by three interrelated factors: (1) the program’s structure; (2) the dialogue that the teacher and learners exchange; and (3) the learners’ autonomy. Moore (2007) explained that these three factors were derived from the analysis of (1) curricula of the distance learning program; (2) communication between teachers and learners; and (3) the role of learners in deciding what, how, and how much to learn (Figure 1). However, the most appealing component of Moore’s transactional distance theory is the inverse relationship between structure and dialogue. That is, as structure increases, transactional distance increases. However, as dialogue increases, transactional distance decreases. This hypothesis has been verified in several studies (Saba, 1988; Saba and Shearer, 1994). The theory becomes more complex by adding the third variable, learner autonomy, because it is unclear whether this represents the learner’s personal autonomy or the autonomy associated with learning materials. However, the theory explains that as transactional distance increases, so does learner autonomy.

## **2.2. HIV and AIDS Management in Kenya**

The Kenya National HIV and AIDS strategic plan was developed with the principle aim of reducing the number of new infections via the use of evidence based approaches to prevention. The strategic plan identified and focussed on six main outcomes that were deemed critical to the attainment of the set goals. The first goal was to reduce risky behaviour in the general population, infected people, at risk groups and the vulnerable population. The other area of focus was increasing and sustaining the proportion of eligible people living with HIV on care and treatment. The strategic plan also identified the need to ensure that health systems delivered comprehensive HIV services. The need for communities and people living with HIV to respond with HIV within their local contexts was also identified as a major issue in the fight to reduce the number of new infections (USAID, 2011; Avert, 2011). The last two outcomes areas include mainstreaming HIV in sector specific policies and strategies and ensuring that stakeholders are aligned and held accountable for results in the fight against HIV.

Improving the availability of information on management of HIV can support the attainment of four out of the six set outcomes. M-learning if effectively used can bridge the gap in accessibility and target the vulnerable and at risk groups to maximise the overall effect of healthcare education and counselling. Generally, HIV prevention has focussed on prevention due to the realisation that there is a significant increase in the number of discordant couples. Transmission between discordant couples is at an all-time high and is the leading causative factor of new infection. Couple based testing and encouraging partner disclosure and condom use are some of the strategies that have been identified as potentially useful in reducing HIV transmission. Improving the counselling and information avenues is useful and vital to supporting and sustaining some of the strategies that have been identified.

### **2.3 M-Learning**

M-learning has been made possible and is propagated by availability of mobile and wireless devices. Mobile communication is no longer limited to companies and individuals that can afford huge investment outlays or specialised software. Individuals can easily access inexpensive mobile telephony. The cost of mobile internet access is also reducing. In Kenya, leading mobile telephony providers such as Safaricom offer daily internet data plans (Safaricom, 2011a). Mobile technologies have enabled new ways of communication typical among young people among whom mobile communication is part of their normal daily interaction. Mobile technologies have created an opportunity for the delivery of learning via devices such as PDAs, mobile phones, PC tablets and laptops. This type of delivery (over a mobile interface) is known as m-learning. M-learning can be looked at as a subset of e-learning (which is web-based delivery of content and learning management). However, the emerging potential of mobile technologies show that m-learning can directly link to the 'just enough, just in time and just for me' model of flexible learning (Marquet, 2010). This implies that m-learning can attain levels of flexibility that allows for customisation of content and pedagogy to meet the learning needs and the context of the learner.

M-learning offers educators unique educational affordances that are vital in dealing with certain environmental and social factors that may impede the learning process. Portability, social interactivity, context sensitivity, connectivity and individuality are some of the benefits and key strengths of m-learning. Organisations have turned to mobile devices in their learning initiatives due to the realisation that they no longer needed large infrastructure and support cost. Small organisations can deliver mobile learning by structuring learning around web-based content that can easily be accessed via web-based mobile devices. Studies reveal that simultaneous personal development can be attained for separate organisations (contexts) via m-learning as long as the strategies are designed and targeted for mixed audiences (Suki, and Suki, 2011a). The portability and mobility afforded by mobile learning devices allows for learning within the context being learned. A study on the use of m-learning to train new sales employees reveals that m-learning eliminates the challenges associated with abstracted studies, and improves the quality and usefulness of training (Keskin and Metcalf, 2011).

In an environment characterised by high levels of mobility, mobile devices allow for synchronous communication thereby allowing for connectivity and overcoming barriers associated with mobility. Mobility is one of the key factors responsible for new numbers of infections in Kenya. Mobile groups such as truck drivers and sex workers are some of the worst affected groups. Despite efforts to ensure that self-help groups and voluntary counselling are spread all over Kenya, the mobility characterising these group of people minimises the gains that these groups can attain from the developments.

### **2.4. Support for Pedagogy**

M-learning has the potential of supporting new instructor-learner relationships. The digital age has led to lasting reversal of the generation gaps in that children are increasingly becoming the technology experts within their homes. As a result, educators and training experts that are steeped in traditional delivery approaches face difficulties when faced with digitally literate learners (Yeonjeong, 2011). Learners no longer limit their learning needs to receiving and memorising the wisdom of their teachers and instructors, rather, students are demanding training that meet specific needs. Traditional teaching techniques and attitudes of contemporary youths are no longer in accord largely due to increased availability of information brought about by the digital



information age (Suki and Suki, 2011a; Suki and Suki, 2011b). M-learning is a platform that if effectively used can bridge this gap.

M-learning creates learning opportunities that differ from those afforded by e-learning and paper based learning. Some of the principle considerations when designing m-learning delivery include urgency of learning needs, the need for knowledge acquisition, mobility of the learning setting, interactivity of the learning setting, interactivity of the learning process, situatedness of the instructional activities and integration of the instructional content (Suki and Suki, 2011b). Younger learners are more comfortable with the notion of using mobile phones for learning. Though some older learners use their mobile phones to arrange meetings and work on assignments, they are generally less comfortable with using mobile phones for learning.

The challenges of creating a learning environment that delivers content via mobile phones have not been adequately addressed by teachers many of whom have recently migrated to the digital world. There are concerns about educators' ability to understand and respond to the digital teaching opportunities due to the aging educators' population and lack of comfort with digital ICTs. Pedagogical differences also exist with most educators aligned to philosophies that focus on teaching and memorisation whereas digital natives seek information. Educators have long established a culture of individualism and secretiveness; as a result, they feel challenged by the need to collaborate with technicians, instructional designers and programmer's web-based education (Yeonjeong, 2011).

Many researchers introduced new approaches for adapting content in a mobile environment. One approach suggested adapting the delivery of small chunks of complementary learning content to mobile devices using Bayesian belief networks. The authors categorize these variables into four main categories: device limitations, connectivity, learner's profile, and content types. A tree organization for the learning assumed an approach where each object in the tree have certain goal (level) and specific features which are compared with the learners' contexts and preference. The sequence of objects is dynamically changed according to the learner contexts and preference. This model enables the learner to participate in customizing his own course. However, this approach is based on the assumption that learners have considerable knowledge in the area of study. This is however not the case for people seeking information on HIV AIDS. Most subject based mining algorithms do not make such assumptions since they tend to be unsupervised and focus on clustering.

Current research efforts in the field of mobile learning have been in many cases guided by a learner-centred approach. Context and content adaptivity are crucial components in mobile learning environments. One important challenge is how to design and implement technological tools and methods to support them. In order to tackle this challenge, learners' context should be defined. Allowing learners to identify the content that they are interested in is one approach that can be used to improve content adaptivity. Another approach involves the use of statistical and data mining techniques to determine aspects that a learner is likely to be interested in.

## **2.5. M-Health Applications**

M-Health applications fall into five broad categories: remote data collection, remote monitoring, communication and training for healthcare workers, diagnostic treatment support, and education and awareness. A key early application, EpiHandy, provided form-based data entry capability on handheld computers (PDAs like PalmPilot) (Danis et al, 2010). It solved some key problems with traditional paper-based solutions, including error reduction methods such as in-the-field data correctness checks and the elimination of handwritten forms, which often take significant time to reach the appropriate ministries and then must be input

into computers (an error-prone process). However, this application was not used for dissemination of information.

One of the many applications developed on top of RapidSMS is ChildCount (Danis et al, 2010). The application was aimed at improving the speed with which community health workers in Kenya were trained to monitor malnutrition and malaria in children under the age of five. Health care workers communicated results through structured SMS messages (messages composed of a series of keywords followed by data values) following a two to three-hour training session that covered the use of the mobile phones and the correct formatting of the SMS messages. Even with training about 10% of messages were rejected by the system because of improper formatting and additional training was required whenever new fields (for the collection of new measurements) were introduced (Danis et al, 2010). This is major undoing of most SMS based m-learning systems.

Another use of structured SMS messages by trained workers was SMS BloodBank, in which nurses were trained to send SMS messages to request replenishment of blood products from a central location (Danis et al, 2010). The main obstacle in that deployment was the cost of the SMS message to the nurses. This very efficient mechanism was in danger of failing until the developers reimbursed nurses for the sending of messages. ELMR is an SMS-based application for updating and querying lightweight mobile health records (Danis et al, 2010). Intended for use not only by health care workers but also by patients (though perhaps with more limited functionality), the intent is to allow users to interact with a database table through SMS.

Millennium Villages in Kenya, Uganda and Rwanda are using mLearning modules on their mobile phones (GSMA, 2010). Reproductive health and care for new-borns are two of the subjects M-Health applications can download from a central database to their mobile phone. Mobile networks provide the privacy, quality of service and interoperability required to deliver this learning application. A good example in Kenya is the Text to Change (TTC) application (GSMA, 2010). TTC uses mobile phone technology to make life saving knowledge easily available to the public and especially to community and family level caregivers. TTC is specialised in interactive and incentive based SMS programs addressing a wide range of health issues such as HIV/AIDS, Malaria and Reproductive Health. By 2009, TTC had reached over 200,000 people with low-tech mobile phone based health messages (GSMA, 2010). The approach used in the TTC application can be summarized as an incentive-based mobile phone quiz: questions are sent to the participants' mobile phones, participants reply, and those participants with the highest accuracy and participation rates are entered into drawings for free "airtime" (pre-paid mobile phone minutes) and other prizes (Mobileactive.org, 2011). The quiz format is intended to provide a fun way for participants to engage with educational content. However, a critical examination of the m-learning approach adopted shows that it lacks in contextualisation. Users of the system have no control over the nature of information that they are sent. This makes the approach less effective for informed public and those that seek information on specific aspects relating to HIV AIDS.

## **2.6. Subject-Based Mining Algorithms**

To solve some problems digital dictionaries are required. In cases where the subjects to be mined are texts, a dictionary is required. A stop dictionary removes those words from the data to be analysed, in which little or no predictive power is expected, as for example, often in articles like "the" or "some" of the case. The problems of polysemy (the ambiguity of words) and synonymy (the equal importance of different words) are solved by the use of digital dictionaries (Sobh, 2010). Dictionaries, which are often domain specific thesaurus,

minimise the synonym problem with increase in large automatically generated corpora. Depending on the analysis type, it may be possible that phrases and words by part-of-speech tagging can be classified linguistically, but this is often not necessary.

### **2.6.1. Clustering**

The common cluster analysis methods (k-means and hierarchical clustering) use self-organizing techniques on fuzzy logic. Very often in subject based mining k-means clusters are formed. These clusters belonging to algorithm seeks the sum of the Euclidean distances over all clusters to minimize within. The main problem is the number of to-find cluster to determine a parameter. Prior knowledge is necessary for parameter setting. Clustering algorithms are often very efficient though there is the risk that only local optimum is found.

In the equally popular hierarchical cluster analysis, documents a hierarchical in cluster tree of their similarity to groups (Miner et al, 2012). This method is clearly more complicated than that for k-means cluster. There are two major approaches used in hierarchical cluster analysis. The first is where the amount of documents in successive steps of shares whereas the second is by document first as a separate cluster perceives each and most similar clusters in the sequence gradually aggregates (Miner et al, 2012). The practice results in but usually only the latter approach to meaningful results. In addition to the runtime problems, another weakness is that for good results on the expected cluster structure required background knowledge. As with all other methods of clustering, the human analysts ultimately decide whether the identified cluster structures of meaning reflect.

### **2.6.2. Artificial Neural Networks**

The 1982 Teuvo Kohonen approach developed the first self-organizing maps another widely used approach to clustering in subject based mining (Sahami, 2009). This is (usually in two-dimensional) artificial neural networks created. They have an input layer, in which each text to be classified document vector represents is multidimensional and as the one neuron assigned to be the centre than, and an output layer in which the neurons sequence of the chosen distance measure enabled the under (Miner et al, 2012). The use of artificial neuron networks complicates the processes and increases the processing demands.

### **2.6.3. Vector-Based Algorithms**

A large number of text-mining methods are vector-based. The resulting high-dimensional vector space is the result of a significantly lower dimensional imaged. Plays since 1990, increasing the Latent Semantic Analysis (LSA) has a significant role in the traditionally singular value decomposition relies. Probabilistic Latent Semantic Analysis (PLSA) is a more statistically formalized approach based on the latent class analysis is based and to estimate the latent class probabilities of the EM algorithm used. Algorithms that are based are LSA on, however, very computationally intensive. A typical desktop computer can hardly have few hundred thousand documents analysed. LSA achieve more of the slightly worse, but less computationally demanding results compared to covariance based subspace methods (Sahami, 2009). The analysis of relationships between documents through reduced matrices makes it possible to identify documents that refer to the same situation, though its wording is different (Sahami, 2009). Evaluation of relationships between terms in this matrix makes it possible to establish relationships between terms associative, and often correspond to the semantic relations in ontology can be represented.

## 2.7. Personalisation

Initially Web services were mainly intended to engage in dynamic business-to-business (B2B) interactions with services deployed on behalf of other enterprises or business entities. Broad interest in standardization/customization efforts was aimed at reducing the necessary user interaction. However, with the advancement of Web service technology the complexity of possible tasks and the availability of services anytime anywhere, e.g. through powerful mobile client devices, will strongly increase. Thus networked services will not only become increasingly sophisticated, but also move into the area of business-to-consumer (B2C) interactions bringing back the user as active participant during interactions. The 'user in the loop' will enhance the service offerings and increase the competing variety that can be sensibly provisioned, but due to this excess information will also demand a great deal of personalization. The development of Internet portals like Yahoo.com already demonstrates the usefulness of a user-centred approach for e.g. social-networking and related areas. In addition, the necessity to provide a personalization/customization mechanisms or a support for browsing available services has become apparent to deal with the number of different offerings that may be more or less relevant for each individual user. For instance, MyYahoo.com allows individual users to define their own portal entry page and determine which personal content to display or which services to activate on start-up. With Web services becoming an integral component of the future Web, a user-centred approach to personal Web services is easily conceivable, e.g. through Web service technology in the backend of Internet portals.

The problem of how to interactively deal with Web services of course has impact on the areas of human computer interaction (HCI) and user interface design (UID) (e.g. for mobile devices with limited capabilities). Interaction with a truly user-centred service can thus be divided into three major phases: the discovery of services that are basically able to perform a task based on a user's service request, the querying of services for subsequent selection based on user preferences for deliverable objects and the final execution of a service, after a decision for one of the available objects.

The amount of research studies that combines Web Usage Mining and Semantic Web is quite limited. Most of the work shows that integrating the semantics involved in structural links with the Web Usage Mining process can improve the discovered patterns. On the other hand, much work was also done in the integration of the content features of site with Web usage mining process (Eirinaki, Lampos, Paulakis and Vazirgiannis, 2004). Studies have elaborated different ways of how the fields of Semantic Web and Web Mining can cooperate (Berendt, Hotho and Stumme, 2002). The first part of the work is on extracting semantics from Web page. The second part is on the improvement of Web Usage Mining by using semantics structures in the form of ontology. Subsequently the authors sketched out the benefits of combining Semantic Web and Web Mining. They further elaborated a process for learning ontologies by mining the Web. They emphasized that by constructing a pattern space over ontology, navigation primitives and Web Mining methods one can find patterns which follows a semantic approach.

The studies presented in Dai and Mobasher (2002), extracted domain level objects from user sessions and created a user profile for each user by aggregating these objects according to their weights and a merge function. It is assumed that there already exists a domain level ontology for the Website and merge functions have already been defined on every attribute of objects. Another related work given in Stumme, Berendt and Hotho (2004), investigated the notion of semantics and ontologies in detail and discussed the possible ways in

which the Semantic Web can improve the results of Web Usage Mining by exploiting the Semantic structures such as ontologies and how Web Usage Mining can help to build up the Semantic Web. The paper discusses how Web Usage Mining can be utilized to learn ontological structures and classify the instances. The studies presented in integrated Content Mining into Web Usage Mining, for effective personalization (Mobasher, Dai, Luo, Sun and Zhu, 2000; Mobasher, Dai, Luo, Sun and Zhu, 2009).

In a recent work a Web Usage Mining Framework for mining evolving User Profiles of dynamic Web sites by exploiting the external ontology, used for mapping and relating dynamic Web pages was proposed (Mobasher, Dai, Luo, Sun and Zhu, 2009). SEWeP integrates the Web usage logs with the semantics of Web site's content to improve the personalization (Eirinaki, Vazirgiannis and Varlamis, 2003). The innovative feature of the architecture was C-logs, an extended form of Web usage log that encapsulates the site semantics. However, the framework was limited only to concept hierarchy. A framework for personalization combining usage information and domain knowledge based on ideas from bioinformatics and information retrieval has been proposed (Bose, Beemanapalli, Srivastava and Sahar, 2006).

It is apparent that there are different approaches that can be used in the personalisation of web content. Ontology based approaches; text mining and the use of agents are the most common approaches. Of these, ontology based approaches offer great potential. However, the development of the ontologies is a complex task that requires significant commitment by the developers. Agents on the other hand either use semantic web approaches or data mining techniques. Use of subject based text mining techniques thus offers a balance between ease of implementation and performance.

## **2.8. Web Syndication**

Web syndication is a process via which website material is availed to multiple sites. Syndication is often used to provide other sites users with summaries or updates on recently added contents and forums. For the receiving website, syndication is an effective way of adding to the depth, accuracy and relevance of the information. Content syndication can also facilitate link building since links embedded within syndicated content can be optimised around anchor items that point at optimised links (Lee, Miller and Newnham, 2008). Subscribing to feeds also has the advantage of minimising the risk of private information sharing, as is the case in subscribing for newsletters. This improves the suitability of web syndication to collection and aggregation of information on sensitive health matters.

Web site syndication has revolutionised the consumption of web-based information especially for dynamic content. Web syndication enabled websites often generate feeds that summarise the content of their sites. By using aggregators, users can poll feeds for updates which are delivered to their mobile devices (Lee, Miller and Newnham, 2008). Additionally, users can specify criteria, and design the feeds such that multiple sources are used to suit their specific interests and needs. Web syndication allows for effective search and use of content since users do not have to manually go through different sites looking for relevant content. Additionally, web syndication allows for the automatic monitoring of sites. Ubiquitous access to information has been greatly improved by the increased usage of web syndication (Lee, Miller and Newnham, 2008). The web 2.0 web syndication has the potential to change higher education and information transfer (Lee, Miller and Newnham, 2008). Some of the known benefits associated with the use of web syndication include ease of access and simplicity of the technology used. Another benefit is that users can easily combine and share information across various sources (Lee, Miller and Newnham, 2008).

There are two commonly used web syndication formats namely really Simple Syndication (RSS) and Atom. RSS is widely used than atom. However, atom offers greater flexibility than RSS. It is worth noting that whereas RSS is copy written, users can copy and even modify the atom codes. This offers it greater extensibility and therefore applicability to several situations. Another benefit associated with the use of atom over RSS is that it had a carefully designed payload container (Hammersley, 2005). This guarantees that recipients cannot arbitrarily and inappropriately process content that they should not. This is a vital feature in improving the validity and reliability of information being transmitted to the subscribers. Whereas RSS has only the <description> element, atom has separate <summary> and <content> elements. This allows greater accessibility of content in cases where the content is non-local or non-textual. RSS and atom also differ in their implementation of auto discovery features. RSS often relies on unregistered Multipurpose Internet Mail Extensions (MIME) type, which is a common source of difficulty for non-technical users. On the other hand, atom uses a standardised auto discovery that includes a self-pointer. This allows users to auto subscribe with only the content of the feed and use web standard dispatching techniques (Hammersley, 2005). Another major difference between atom and RSS that is highly relevant to the study is in extraction and aggregation. RSS recognises documents of <rss> form (Wittenbrink, 2005). Atom on the other hand allows standalone atom entry documents that can be transferred using any network protocol. Additionally, atom supports aggregated feeds thereby allowing feeds to point back to their sources when they are included in other feeds.

## 2.9. Proposed Solution

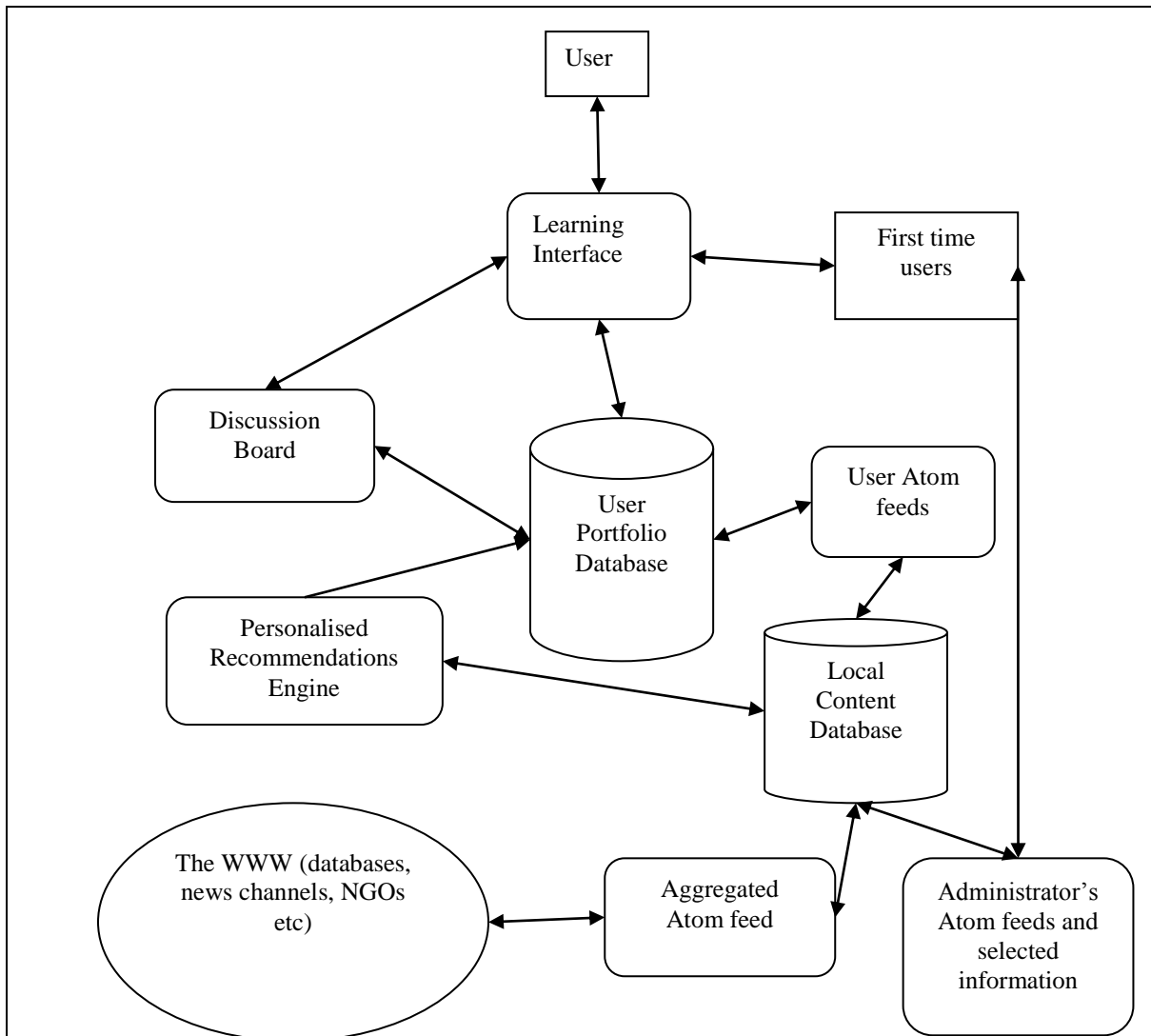


Figure 2: Proposed Solution

The solution allows the user to login into the learning interface that offers access to the discussion board and user portfolio databases. The user portfolio database is populated not only by the atom feeds from the user but also information from the personalised recommendations engines (Fig. 2). The personalised recommendations engine implemented a subject based mining protocol that allows users to access information commonly accessed by users who use similar search words. In general, the personal recommendations engine is aimed at using data mining to provide users with information commonly accessed by other users who access similar information. Aggregated atom feeds from all users are used to gather information from selected databases and information sources. These feeds populate the local content database from which data conforming to specific users' feeds are transferred to those users' portfolio database. The creation of feeds is hidden from the users who create the feeds by simply typing in aspects that they need to be updated. The approach used

ensures that the content is incrementally developed and users can access information that they need (direct feed updates) and relevant information (personalised recommendations engine). The approach also allows for context awareness by presenting new or updated context (feeds) to an interested user and making the context persistent (personalised recommendation engine) for the user to retrieve later.



## CHAPTER THREE: METHODOLOGY

### 3.1. Introduction

This section presents the prototype development methodology, research design, the research approach, data collection, data analysis, validity and reliability, ethical considerations and lastly the limitation of the adopted research approach.

### 3.2. Prototype Development Methodology

A rapid application development methodology was used in the study. Rapid application development methodologies require minimal planning and in general tend to favour rapid prototyping. In most cases, planning is interleaved with the coding of the software. RAD methodologies allow for faster changes in requirement and improve the speed with which software is written. Of the available RAD methodologies, the study employed agile software development. Under agile methodology, the requirements and the solutions change via collaborative work. Testing of software is often conducted as it is being developed. Testing often involves two dimensions: the developers and acceptance testing focussing on the clients. The choice of the agile methodology is guided by its support for adaptive software development approach. In general, adaptive methods allow for easy transformation to changes in realities. Software development teams under adaptive methods are required to change their approach with changes in the project. Since the study involved prototyping which typically involves changes in the specifications, adoption of an adaptive approach was in order. The other reason for selecting agile development methodology is the limited time span within which the application was developed. Agile methodology is more paid than other RAD methodologies. Thus, it is more likely that the prototype was developed within the limited time under this methodology.

### 3.3. Research Design

A quasi-experimental research design was used in the study. To develop a model that allows for interaction and sustainable resource use, an understanding of the real environment is required. Quasi-experimental research design unlike experimental design does not require randomisation (Andrew and Halcomb, 2009). The study involved the development of a prototype it usage and reporting on various assessment items. A quasi-experimental design is selected due to the definition of the population. Truck drivers are few and highly mobile and therefore the researcher does not have the luxury of being selective.

### 3.4. Research Approach

To validate a prototype and determine its usefulness, it is important to carry out a test drive. In Kenya, citizens above the age of eighteen years own mobile phones and subscribe to social networking sites. Mobile users who can access the internet are the intended users of the proposed m-learning system. This is the main reason why efforts aimed at understanding the efficacy of the prototype and challenges focused on selected Kenyans. The selected participants were informed of the aim of the study and the procedure to be used. The study involved 50 participants. This sample size is large enough to validate the use of statistical techniques in analysing the data.

### 3.5. Population, Sample and Sampling

The population is defined as Kenyans who own web enabled mobile phones. A sample of 50 truck drivers was selected. The emphasis on truck drivers is derived from the documented effect of mobility on HIV AIDS in Kenya. By focusing on this highly affected group, the findings from the study can be generalised with high precision to the population. The participants must know how to operate a web enabled mobile phone and

accessing the internet. Additionally, the participants selected were over the age of eighteen years. Convenience sampling is used in that the selected participants are those truck drivers that are willing and able to participate in the study. The use of convenience sampling is thus aimed at addressing some of the practical challenges associated with the sample of interest.

### **3.5. Data Collection**

The study involved the collection of both secondary and primary data. The collection of primary data was done via questionnaires. The focus of the data collection was determining the challenges associated with the use of the developed prototype and determining whether learning has taken place. The participants were required to rate the usability and usefulness of the prototype. An interval scale was used since perception of the usability and utility of a prototype are qualitative and subjective variables. Five-point Likert type questions were used. The scale was represented by very poor=1, poor=2, average=3, good=4, and very good=5. The questionnaire design was based upon the computer system usability questionnaire (Lewis, 1995) that has high psychometric scores. The participants were asked questions (a quiz) before using the prototype and after using the prototype. The quiz was based on information stored in the databases. The individual questions were sourced from a professional self-assessment tool on HIV AIDS knowledge (Davao Institute for Educational Research Development and Consultancy, 2011). The scores in this quiz were used to determine if learning has taken place.

### **3.6. Data Analysis**

Data collection involved the collection of both qualitative and quantitative data on the usability and usefulness of the developed m-learning prototype. The data collection also focused on the pedagogical and technical challenges faced in using the system. One sample t-test was used to determine if the participants' perception of the usability, usefulness and knowledge gained from the system is significantly different from set threshold levels. Paired two-sample t-test was used to determine if the participants' scores in the quiz prior to and after using the prototype are significantly different. Two-sample t-test is used because the data analysis involved the performance of the same sample prior to and after interacting with the system (Creswell, 2009). Qualitative content analysis was used to analyse the responses on the system and their implication on the prototype and further studies in the same area.

### **3.7. Validity and Reliability**

Qualitative techniques in general allow for collection of data from different sources and triangulation. Triangulation is vital in minimising errors and biases associated with specific sources and therefore influences the validity of the findings positively. The interview approach allows the researcher to collect explanations on the ratings awarded to the system by the user. Moreover, the study involved collection of data from truck driver (who are from different parts of the country and are worst affected by the failures of the current community education approaches). This is a move aimed at improving the overall reliability of the study and ensures that the nature of the real world is captured accurately.

### **3.8. Ethical Considerations**

The study did not involve any minors. The participants were adults who could make informed decisions on whether to participate in the study. Identification information of the participants was kept private and confidential. Lastly, the participants were informed of the aim of the study and the procedures that was used

beforehand. This is a measure aimed at ensuring that their participation in the study is based on informed consent.

## CHAPTER FOUR: SYSTEM DESIGN AND IMPLEMENTATION

### 4.1. Introduction

This chapter presents the processes that were used in the design and implementation of the software prototype. The software prototype was developed for the purpose of experimentation. The chapter describes the design and implementation of the prototype in terms of the programming and the testing strategies that were used.

### 4.2. Data Set

The system data set can be categorised into users and areas of interest. Different users were required to select areas in HIV AIDS learning that they are interested in. In addition, the system is expected to grow with time in that the users may suggest other areas that they are interested in and are not included in the sections they can select. Interest in an unlisted area by a pre-set proportion of the users results in its inclusion in the categories that users can choose. Each user has an identifier (userID) and each field or category has a unique identifier. Every user will therefore be associated with a number of fields or categories that they want to be updated. This approach allows for easy analysis of the relationship between the users and the fields that is important for the personalisation engine.

### 4.3. System Design

The prototype development methodology used is agile software development methodology. The agile unified process, which is a variant of the agile software development methodology, was adopted in the system design and implementation. It is worth noting that the system involves the addition of new fields after its completion. The agile unified process allows for configuration management that involves managing fields, tracking and controlling the changes in the fields over time. The Agile Unified Process (AUP) allows seven disciplines namely model, implementation, test, deployment, configuration management, project management and environment. Environment principle involves the inclusion of efforts to ensure that proper process, guidance and tools are available for the development team. Project management on the other hand involves all processes aimed at managing risks and ensuring that the project is delivered within time.

### 4.4. System Modules

Authentication to the system involves the use of a password and a username. Every user is required to sign-up into the system. The administrator of the system will also access the system using the same authentication mechanism. After accessing the system, the users can manage their feeds (areas that they need to be updated on) and see feed updates from different fields and a discussion board. The administrator on the other hand can access reporting tools and add feed categories. In addition, the administrator can add any additional information into the discussion board.

Text clustering was used in the implementation of the personalisation engine. Clustering in general involves the division of data points into homogeneous clusters. This implies that elements in the same class after classification have to be as similar as possible. In this study, the focus of the classification is non-numerical textual data. Similarity calculation and generalisation are key issues when clustering non-numerical textual data. Since each field is associated with a given identifier, this association is used to support the clustering. Probabilistic and hierarchical clustering algorithms can be used to implement the clustering. A variation of K-mean clustering algorithm was used in the study. Jacard's coefficient scores were used to measure similarity between user preferences. Jacard's coefficient is one of the measures used for assessing similarity under K-

means clustering the others being Euclidean distance and cosine similarity (Park, 2012). Every user has a Jacard's coefficient showing the similarity and divergence in user interests. This is a mapping between one user and all the other users. For every user, four users with the highest Jacard's coefficient with the selected user were determined. As a result, every user has a set of four other users who have the highest similarity index. Check for feeds accessed by the set members but not by the user. Which feed category is subscribed to by most users in this set but not by the selected user? Return the most commented on and most recent feed update in this category.

$$\text{Jacard's coefficient} = \frac{\text{Number of categories in Common}}{\text{Total number of categories in Common}}$$

Where categories refer to the feed categories selected by a user.

#### **4.5. Prototype Testing**

Unit testing is carried out during the implementation of the prototype. The AUP process allows for formal analysis and design during the implementation of a prototype. This minimises the need for an intensive system testing. Unit testing was carried out to test individual modules. The discussion board, login system, administrator and user interfaces and the personalisation engine are some of the modules that were tested at this phase. The user while playing the role of a user and an administrator carried out user acceptance testing. The developer came up with a list of desired functionalities and system behaviour. The user tested the system against the mock user requirements at each developmental phase.

#### **4.6. Prototype Deployment**

For demonstration purposes, the system will be accessed from the host computer using a mobile phone. The system is quite heavy and requires various tools most of which have yet to be availed on mobile platforms. However, accessing the prototype over the internet ensures that the servers handle the associated resource requirements. An extensive deployment of a related system would however require the use of dedicated web servers than can be accessed on different mobile platforms. For the study purpose, the prototype was hosted for a period of three days within which the participants interacted with the system.

#### **4.7. Limitations of the Design**

The major limitation of the design approach adopted is that it does not include the intended users in the design process. The developer is likely to be a biased user since he has prior knowledge of the system. This limitation is however addressed by gaining feedback on the system from the users after the deployment of the prototype. It is worth noting that based on users' comments and feedback; the administrators can add new fields that users can be updated.

## CHAPTER FIVE: RESULTS AND DISCUSSION

### 5.1. Introduction

This chapter presents the results and discussion of the findings from the research. Fifty truck drivers were involved in the study. This chapter is important in addressing the research objectives and the overall research aim.

### 5.2. Results

The study involved the use of questionnaire and collection of data pertaining to the use of the system by the participants. In addition, the participants had to answer questions derived from a universally accessible discussion board in the system. This measure was aimed at determining if the system helped the participants gain knowledge. The participants answered the same quiz after and prior to using the system. Other aspects that were investigated in the research include the system's ability to integrate information, usability, relevance of the information provided to users and relevance of the recommended categories generated by the system. Other issues that were investigated are the technical and the pedagogical challenges that were faced by the participants when using the system.

**Table 1: Summary of Key variables**

	N	Range	Minimum	Maximum	Mean		Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
Pife	50	6.00	2.00	8.00	5.4200	.21588	1.52650
Pise	50	7.00	3.00	10.00	7.1000	.23776	1.68123
Integration	50	4.00	1.00	5.00	3.6400	.14780	1.04511
Usability	50	4.00	1.00	5.00	3.4200	.15410	1.08965
Relevance	50	4.00	1.00	5.00	3.5200	.13774	.97395
FCS	50	8.00	.00	8.00	4.0800	.24385	1.72426
FCI	50	3.00	.00	3.00	1.3400	.11990	.84781
FCA	50	3.00	1.00	4.00	2.0200	.11949	.84491
FCC	50	3.00	.00	3.00	1.1200	.09748	.68928
rel_recc	50	4.00	1.00	5.00	3.6200	.14254	1.00793
rel_reccactual	50	3.50	1.50	5.00	3.0500	.12143	.85863
Comments	50	3.00	.00	3.00	.9200	.11006	.77828
Valid N (listwise)	50						

Table 1 presents a summary of the key variables. Performance in first exam (PIFE) denotes the performance of the participants in the quiz prior to using the system. The quiz was marked out of ten and the lowest participant had a two whereas the highest had eight. The standard deviation in the first exam is 1.526 whereas the mean is 5.42. The standard error associated with the mean is 0.215 (Table 1). Performance in the second exam (PISE) denotes the participants' performance in the quiz after using the system. The participant with the lowest performance in the second quiz had a 3 whereas the highest performer had a perfect score. It is important to point at the differences between the highest and lowest pife and pise scores. Importantly, the mean performance in the second quiz was 7.1 which is higher than the 5.42 recorded in the first quiz (Table 1). The

standard deviation and standard error associated with the mean in the first and second quiz are comparable. In the second quiz, the standard error associated with the mean was 0.237 whereas the standard deviation was 1.68.

Data on integration, usability and relevance of the system were collected using a questionnaire and input from the fifty participants that had interacted with the system. The users were required to rate the system on a scale of one to five, under the said metrics. Integration refers to the system’s ability to collect and present different information from diverse sources. Relevance on the other had refers to the appropriateness of the retrieved information to the users’ learning or information needs. Usability refers to user’s ability to freely use the modules presented by the system. These three metrics have a minimum of one and a maximum of five. The average scores for integration, usability and relevance are 3.64, 3.42 and 3.52 respectively (Table 1). Additionally, the standard deviations are 1.045, 1.0089 and 0.974 respectively.

**Table 2: Integration Frequency Table**

		Frequency	Per cent	Valid Per cent	Cumulative Per cent
Valid	1.00	2	4.0	4.0	4.0
	2.00	4	8.0	8.0	12.0
	3.00	15	30.0	30.0	42.0
	4.00	18	36.0	36.0	78.0
	5.00	11	22.0	22.0	100.0
	Total	50	100.0	100.0	

**Table 3: Usability Frequency Table**

		Frequency	Per cent	Valid Per cent	Cumulative Per cent
Valid	1.00	2	4.0	4.0	4.0
	2.00	8	16.0	16.0	20.0
	3.00	16	32.0	32.0	52.0
	4.00	15	30.0	30.0	82.0
	5.00	9	18.0	18.0	100.0
	Total	50	100.0	100.0	

**Table 4: Relevance Frequency Table**

		Frequency	Per cent	Valid Per cent	Cumulative Per cent
Valid	1.00	2	4.0	4.0	4.0
	2.00	3	6.0	6.0	10.0
	3.00	20	40.0	40.0	50.0
	4.00	17	34.0	34.0	84.0
	5.00	8	16.0	16.0	100.0
	Total	50	100.0	100.0	

Frequency tables (2, 3, and 4) provide further information on the distribution of the rating of the system. The rating of the system corresponding to average and good (3 and 4) are associated with the highest number of participants in all cases. Furthermore, in all the three cases, the ratings associated with very poor and poor (1 and 2) are associated with the least participants. The rating corresponding to very good (5) has the third highest frequency in all three metrics. The implication of these findings is that most participants are of the view that the system ranges between average and very good in terms of usability, relevance and integration. However, further analysis will be carried out to determine if the rating of the system in all the three metrics is significantly higher than average.

Feed category selected (FCS), Feed category Ignored (FCI) and Feed category adopted (FCA), Feed category checked (FCC) are variables relating to the use of the system by the participant. The highest number of categories initially selected by a participant was eight whereas the lowest was zero. On average, the users selected four feed categories with the standard deviation of 1.72 (Table 1). The highest number of categories ignored by a user is three whereas the lowest is zero. On average, the participants ignored 1.34 (2) categories with a standard deviation of 0.84 (1) between the users. The category adopted refers to feeds that users select from the recommended feed categories. The participants adopted between one and four feed categories. On average, the users adopted two feed categories with the standard deviation between the users being 0.844(1) (Table 1). The feed categories checked refers to the categories that the users checked from the recommended feed categories. The participants checked between zero and three feed categories. On average, the participants checked one feed category of the recommended feed categories.

Rel\_recc and rel\_reccactual refer to the relevance of the recommendations as rates by the participants (in the questionnaire) and as shown by their use of the system respectively. Rel-reccactual is calculated by averaging the categories that are adopted by the users and the categories that are checked. On average, the participants rated the relevance of the recommendations engine as 3.6 with a standard deviation of 1.008 (Table 1). On average, the participants either checked or adopted three feed category from the recommendations engine



(Table 1). The users could also make comments when using the system. On average, each participant made a comment (Table 1).

**Table 5: Comments Frequency Table**

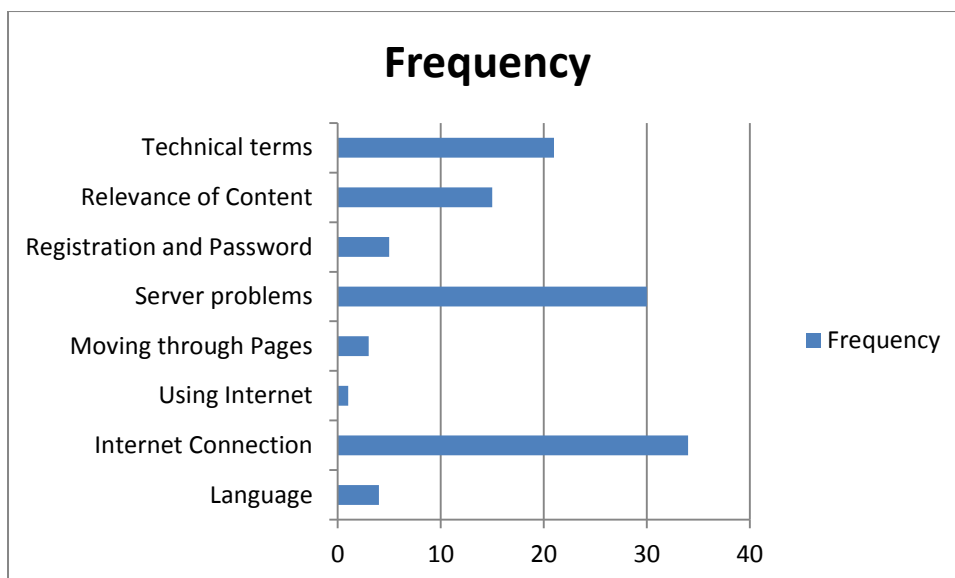
	Frequency	Per cent	Valid Per cent	Cumulative Per cent
Valid .00	16	32.0	32.0	32.0
1.00	23	46.0	46.0	78.0
2.00	10	20.0	20.0	98.0
3.00	1	2.0	2.0	100.0
Total	50	100.0	100.0	

Further analysis reveals that 32% of the participants did not comment, 46% of the participants had one comment, 20% of the participants had two comments and 2% of the participants had three comments.

**Table 6: Challenges Faced By Users**

<i>Challenges</i>	<i>Frequency</i>
Language	4
Internet Connection	34
Using Internet	1
Moving through Pages	3
Server problems	30
Registration and Password	5
Relevance of Content	15
Technical terms	21

Table 6 presents the learning and technical challenges that were faced by the users. The technical challenges that are mentioned predominantly by the users are internet connection, using internet, moving through pages, server problems and registration and passwords. The pedagogical challenges that were mentioned by the participants include language, relevance of content and technical terms.



**Figure 3: Distribution of Challenges Faced by participants**

Figure 3 shows the frequency of reporting of the technical and pedagogical challenges that were faced by the users. The problem reported by the most participants is internet connection followed by server problems. These interrelated connectivity issues arise from availability of internet, the speed of the connection and availability of the host server. Registration and password, moving through pages and using the internet are user-oriented problems that could be due to the design of the system or the user’s ability to follow instructions. The pedagogical challenge reported by most users is the use of technical terms in the feed updates. Twenty participants are of the view that the use of technical terms in the feed updates is a problem that may hamper their acquisition of knowledge. Relevance of content and language are the other pedagogical challenges reported by 15 and 4 participants respectively (Figure 3). These of feeds as the main update mechanisms is likely to present information that is not always very specific to the needs of the users. Furthermore, the updates are in English. This can be a problem for users not fluent in the language.

### 5.3. Findings

**Table 7: Paired correlations pife and pise**

	N	Correlation	Sig.
Pair 1 pise and pife	50	.421	.002

Comparison of the quiz results prior to and after interacting with the system provides important information on the role of the system in knowledge acquisition. A comparison of the scores prior to and after the study shows that they are significantly correlated with a correlation coefficient of 0.421 and a p-value of  $0.002 < 0.05$  (Table 7). This implies that the performance in the first quiz is positively correlated with the performance in the second quiz.

**Table 8: Paired Sample t-test pife and pise**

	Paired Differences					T	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 pise - pife	1.68000	1.73134	.24485	1.18796	2.17204	6.861	49	.000

Table 8 presents results of the paired two-sample t-test of pife and pise. The findings reveal a t of -6.861 showing that the mean value of pise is higher than pife. The p-value is found to be less than 0.005 and as a result, it can be concluded that there is statistically significant difference between the mean values of pife and pise (Table 8). Since pise has the highest mean, it can be concluded that the performance of the participants after interacting with the system is significantly higher than their performance prior to interacting with the system.

**Table 9: One-Sample t-test test-value 5**

	Test Value = 5					
	T	Df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
pife	1.946	49	.057	.42000	-.0138	.8538
pise	8.832	49	.000	2.10000	1.6222	2.5778

Further analysis of the performance of the participants reveals that in both cases the mean score was higher than five. The mean difference from five in both cases is positive (0.42 and 2.1) (Table 9). However, it is only pise (p-value<0.05) that has a mean value that is statistically significantly different from 5. The differences in the pife scores (p-value =0.57>0.05) are not statistically significant and therefore the differences can be attributed to random errors. The implication of these findings is that the participants' performance in the first quiz was average (5) whereas their performance after interacting with the system was above average (different and greater than 5).

**Table 10: One sample t-test integration, usability, relevance, rel-recc test-value 3**

	Test Value = 3					
	T	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
integration	4.330	49	.000	.64000	.3430	.9370
usability	2.725	49	.009	.42000	.1103	.7297
relevance	3.775	49	.000	.52000	.2432	.7968
rel_recc	4.350	49	.000	.62000	.3336	.9064

Table 10 presents findings from an analysis aimed at determining if the average rating of the system by the users is above average (3). The findings show that for all the three metrics (integration, usability and relevance) the mean differences are positive which implies that in all these cases the mean values are greater than 3. Integration and relevance have statistically significant differences from the test value with p-values<0.05 in both cases. Usability has a p-value=0.009<0.05 (Table 10). This implies that the difference from the test value is significant. The implication is that the participants rating of the system in terms of integration, usability and relevance of the information collected is above average. Rel\_recc which represents the relevance of the feed categories recommended by the personalization engine is also significantly higher than 3 with a p-value<0.05. The implication is that the participants are of the view that the recommendations made by the engine have a relevance rating greater than average.

**Table 11: Paired sample t-test FCI-rel-reccactual**

**Paired Samples Test**

	Paired Differences						t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference					
				Lower	Upper				
Pair 1 rel_reccactual - FCI	1.71000	1.28607	.18188	1.34450	2.07550	9.402	49	.000	

A comparison of the number of feed categories ignored and the relevant recommended categories (depicted by the average number of feed categories checked or adopted by the users) provides a good pointer to the efficacy of the recommendations engine. A paired t-test analysis of these variables reveals a positive t value. This implies that the FCI mean is lower than the rel\_reccactual mean value. The p-value <0.05 implies that the differences in the means of the two values are statistically significant (Table 12). The implication is that the relevant recommended categories were higher than the ignored (irrelevant) feed categories developed by the personalisation engine.

#### **5.4. Discussion**

This section will discuss the implication of the results and the findings in relation to the research objectives, theoretical framework and the methodology adopted in the study. This section is important in developing an understanding of the practical and theoretical implications of the results and findings.

##### **5.4.1. Objectives**

The study has four objectives that collectively play a role in attaining the main aim. The first objective is to identify the pedagogical and the technical challenges that are likely to be faced in using the m-learning model and the e-platform in discussing HIV AIDS. The findings reveal a number of technical and pedagogical challenges that were identified by the participants. Internet connection and server problems are the most frequent challenge. These technical challenges are due to connectivity problems, server maintenance and in some cases incorrect feed links. These technical challenges can be addressed by continually checking the feed links to ensure that they are accurate and improvement in internet infrastructure in the country. The latter requires input from different stakeholders. Other technical issues that were highlighted by the participants include registration, passwords, and moving through pages. Registration and password problems are a result of having similar user names and using authentication codes when registering. The system has however been designed to ensure that the users are informed of any erroneous entries when registering and has a password recovery mechanism. It is also probable that the registration and password problems are due to apprehension and lack of adequate experience in using systems that require user authentication (Weira, Douglassa, Richardsonb, and Jacka, 2010). Moving through pages is a challenges reported by only three of the participants. The problem may be due to the size of the screen in the phones that they used and inexperience in using web browsers within mobile phones.

The study also identified a number of pedagogical challenges. The first challenge is the language used. The system is implemented in English and users are required to select English typed feed categories. As a result, the feed updates are in English. Language used can be a challenge for participants that are not fluent. The use of technical terms is a related problem that has been identified by the users. When creating the feed categories the researcher endeavoured to simplify the language used and make the categories relevant to typical community education on HIV AIDS scenarios. However, the actual updates may and often do contain technical terms. This is a challenge in using the feed based system. However, the users can still use their mobile devices to get more information on the technical terms. Discussion platform offered by the system is another platform that can be used to dispel any misunderstanding arising from the use of technical terms. Importantly, the existence of technical terms highlight the potential benefits that experts in HIV AIDS management stand to gain by using such a system. Relevance of content is a pedagogical challenge that has been identified by 15 participants. Relevance of content is tied to language and the use of technical terms. The use of web syndication is also a

major factor in relevance. Web syndication searches for keywords and this may result in feed updates that are not necessarily related to what the users hoped (Hammersley, 2005). This challenge can be addressed by the manual removal of updates that are deemed irrelevant by a power user (administrator).

The second objective of the study is to determine if the m-learning model influences gain of knowledge on HIV AIDS. Pursuant to this objective, the participants were quizzed prior to and after using the system. The same quiz was presented to the participants. This measure is aimed at ensuring that the treatment is uniform. The quiz was based on information that is universally available for all users. The findings reveal that the performance of the participants in the quiz was significantly higher than the pass mark after users interacted with the system. Furthermore, the findings reveal that the scores prior to and after interacting with the system are significantly different with the latter being higher. These findings suggest that interaction with the system had some influence on the latent information retained by the participants. It is however noteworthy that this is subject to the assumption that the improvements in scores was due to interaction with the system. This is not necessarily true since the users could have gotten information from other sources after the first quiz. Another issue is that the difference in the scores may be influenced by users' interest and access to the information from which the quiz was based. Additionally, it is probable that the participants focussed on the content, however this possibility is minimised by the fact that the participants had no prior knowledge that the same quiz will be repeated. This is a form of single blinding. Thus under the research condition and assumptions, the findings show that the m-learning model influences the knowledge gains made by the learner.

The third objective is to personalise the information presented to the individual model users in terms of their interests. The model incorporated a personalised recommendations engine that recommended categories that the participants or users could find interesting. This engine implemented a similarity matching mechanism based on k-mean and Jacard's similarity coefficient matching. Users could either ignore the recommendations (if irrelevant), subscribe (adopt) the categories and check the categories in which case they would access the feed update related to the category. A comparison of the feeds ignored and those adopted or checked reveals a significant difference. The implication is that most users found the recommendations engine to be relevant to their interests. This is further supported by users' rating of the relevance of the recommendations engine. Analysis reveals that users' rating of the relevance of this engine is significantly higher than average (3).

The fourth objective is to develop an m-learning model that allows for the aggregation and contextualisation of information on HIV AIDS to specific user needs. For a system to attain this objective, it should integrate information, be usable and provide information that is relevant to users. An examination of summary statistics reveals that users could log into the system, subscribe to certain feed categories and if not select (adopt) a category from the recommendations engine. The comments statistics also reveal that some users commented on various updates within the system. Analysis of participants rating of the system in terms of the usability, integration of information and relevance of content reveals that for all these metrics the participants are of the view that the system is above average. The level of participation in the system and their rating of the system suggest that they were able to aggregate and contextualise information. With the latter being reflected on their rating of the relevance of content and relevance of recommendations made by the personalisation engine.

#### **5.4.2. Framework**

Transactional distance as per the transactional distance theory is influenced by program structure, dialogue between teachers and learner and the learner autonomy. The theory asserts that there is an inverse

relationship between structure and dialogue in that, transactional distance increases with increase in structure and decrease in dialogue. A critical examination of the design of the system shows that the components included in the system have targeted minimisation of the challenges associated with increase in transactional distance. Users are separate from the teachers (administrators) and only interact via comments and discussions. The inclusion of a personalised recommendation system is aimed at increasing learner autonomy. By allowing users to select the categories they want to be updated on, the system significantly increases learner autonomy. With increase in transactional distance, learner autonomy should increase and this is supported by these measures. An examination of the challenges reported by the participants highlight the need for increased dialogue as a mechanism for dealing with the challenges associated with increased transaction distance. Discussing technical terms and help system to deal with password and registration problems are example of dialogue based strategies that can help address some of the pedagogical and technical challenges faced in using the system. From this perspective, it is evident that the findings reinforce the transactional distance theory by highlighting the need for increase in dialogue and learner autonomy with increase in transactional distance.

#### **5.4.3. Methodology**

In relation to the methodology, there are a number of limitations highlighted in the findings. First, the conclusion that the m-learning model supports knowledge acquisition is reached subject to a number of assumptions. This is largely a result of the use of a quasi-experimental design rather than a fully experimental design. Further studies should look into the possibility of including randomised treatment, control groups and blinding to accurately determining the implication of such a system on learning. The inclusion of these measures would reduce the need for subjecting the findings to a myriad of assumptions. Another factor is that there is need for increase in exposure time. In the current study, the participants were exposed (or used the system) for a one week period. It is not clear if this period is enough for the participants to acclimatise with the system. Increasing the exposure time and or determining the time required for users to be fully aware of the functionalities of the system can help ensure they enjoy the full benefits of the system and its limitations. These methodological issues have to be addressed by future studies.

## **CHAPTER SIX: CONCLUSION AND RECOMMENDATIONS**

### **6.1. Introduction**

This chapter presents the conclusion and recommendations arising from the findings. The chapter discusses the main findings and the implications that the findings have on policy, practice and research.

### **6.2. Conclusion**

- a. The study sought to determine how m-learning could be used to create an interactive learning environment that can help improve awareness on HIV in Kenya by improving the contextualisation of information and providing a seamless platform for discussions on issues relating to HIV AIDS. The findings reveal that the use of an m-learning model that supports aggregation of information, interaction and personalisation of information to specific user needs can help support contextualisation and a seamless platform for discussion of HIV AIDS thereby supporting community education.
- b. The study focussed on fifty truck drivers in Kenya and would have to be replicated for other groups for the findings to be indicative of its potential in community education. However, due to the homogenous nature of the sample, the findings can be generalised for groups that are highly mobile who coincidentally are worst affected by HIV AIDS in Kenya.
- c. The study findings also show that the m-learning model can foster knowledge acquisition among truck drivers.
- d. However, there are a number of methodological limitations associated with the study including lack of randomisation and control groups.
- e. Other issues include limited exposure time of the sample to the treatment. These aspects can be integrated into future studies. Importantly, the study supports the need for increase in dialogue and empowering students or learners to be autonomous in learning settings where there is high transaction distance.

### **6.3. Recommendations**

The recommendations are based upon the findings developed from the study. The recommendations focus on the areas of policy, practice and research.

#### **6.3.1. Recommendations for Policy and Practice**

The study findings reveal that m-learning and m-health applications have potential in areas of community education specifically in HIV AIDS. The findings further reveal that some of the challenges associated with m-learning arising from increase in transaction distance can be addressed by increased incorporation of dialogue and empowering learners to be autonomous. These findings have multiple implications on m-learning and community education on HIV. First, policies targeting the use of m-learning in community education should be considered by the health institution and non-governmental organisations interested in improving information availability in special groups such as truck drivers. Such policies will create the platform required for continued research into and adoption of m-learning as a platform for community education on HIV AIDS. Another issue is that there is need for increased emphasis on personalisation of information in community education. Supporting initiatives aimed at improving both the availability and personalisation of information to specific user needs can help improve the relevance of community education to user interests. There is need to support other approaches to personalisation such as the use of ontologies and agents. Use of m-learning system to support knowledge acquisition by professionals in the areas of community



education and counselling should be considered. The immense potential offered by m-learning, data mining and web syndication can help ensure that professional educators and learners both gain.

### **6.3.2. Recommendations for Research**

The study findings have a number of implications in research. First, there is need to replicate the findings relating to the relevance of recommended categories, support for knowledge acquisition, integration and relevance of content in other populations and different research settings. Future studies should target the use of experimental design to minimise the effects of random errors and increase in exposure time. Longitudinal designs aimed at determining knowledge gain over long periods should be considered. These strategies will help eliminate any apprehensions arising from the approaches used in data collection. Another research area that would help facilitate the large scale deployment of such an m-learning system is focussing on the feed sources. To improve localisation of the information to Kenyan contexts there is need for increased digitalisation of information and use of web technology. In this way, m-learning system based on web syndication can offer more benefits to a learner interested in information specific to the Kenyan context. Increased research into other personalisation approaches and comparative assessment of their accuracy in an m-learning context would help improve the accuracy of personalisation in the long term. This area of research should be looked at. Lastly, there is need for studies that will help administrators develop mechanisms for selecting the best feed sources. Such studies would offer information that will be of practical benefits in populating the m-learning system.

## REFERENCES

- Abbad, M.M. and Albarghouthi, M. (2011). Evaluate Students' Perceptions of the Virtual Learning Environment at Paisley University. *International Journal of Emerging Technologies in Learning*, 6(3), 28-34
- Andrew, S., and Halcomb, E. (2009). *Mixed methods research for nursing and the health sciences*. New York, NY: John Wiley and Sons
- AudienceScapes (2011). *Mobile Communications in Kenya*. Retrieved 22 December 2011 from <<http://www.audiencescapes.org/country-profiles/kenya/media-and-communication-overview/mobile-communications/mobile-communications->>
- Avert (2011). *HIV and AIDS in Kenya*. Retrieved 20 December 2011 from <<http://www.avert.org/hiv-aids-kenya.htm>>
- Berendt, B., Hotho, A., and Stumme, G. (2002). *Towards Semantic Web Mining*. In Horrocks, I., Hendler, J. (eds.) (2002). ISWC 2002, LNCS, 2342, pp. 267-278. New York: NY: Springer
- Bridgland, A., and Blanchard, P. (2005). *Powerful, portable, personal computing: Is m-learning an opportunity in e-learning?* Retrieved 20 December 2011 from <<http://eprints.infodiv.unimelb.edu.au/archive/00000889/01/bridgland.blanchard.pdf>>
- Bose, A., Beemanapalli, K., Srivastava, J., and Sahar, S. (2006). Incorporating Concept hierarchies into Usage Mining Based Recommendations. *Proceedings of WEBKDD, 2006*, Pennsylvania.
- Boyd, R. D., and Apps, J. W. (1980). *Redefining the discipline of adult education*. San Francisco: Jessey-Bass.
- Creswell, J.W. (2009). *Research design: qualitative, quantitative, and mixed methods approaches*. (3<sup>rd</sup> ed.). Thousand Oaks, CA: Sage
- Dai, H., and Mobasher, B. (2002). Using Ontologies to discover domain- level Web Usage Profiles. *Proc. of the 2nd Semantic Web Mining Workshop at ECML/PKDD, 2002*, Helsinki, Finland.
- Danis, M.D., Ellis, J.B., Kellogg, W., van Beijma, H., Hoefman, B., Daniels, S., and Loggers, J.W. (2010). *Mobile Phones for Health Education in the Developing World: SMS as a User Interface*. Retrieved 9 March 2012 from <[http://www.mobileactive.org/files/file\\_uploads/IBM-TTC.pdf](http://www.mobileactive.org/files/file_uploads/IBM-TTC.pdf)>
- Davao Institute for Educational Research Development and Consultancy (2011). *HIV/AIDS questionnaire-test your knowledge*. Retrieved 9 March 2012 from <<http://www.slideshare.net/pamelaveroy/hivaids-questionnairetest-your-knowledge>>
- Dewey, J., and Bentley, A. F. (1946). Interaction and transaction. *The Journal of Philosophy*, 43(19), 505-517.
- Eirinaki, M., Lampos, C., Paulakis, S., and Vazirgiannis, M. (2004). Web Personalization Integrating Content Semantics and Navigational Patterns. *Proceedings of WIDM, 2004*. USA
- Eirinaki, M., Vazirgiannis, M., and Varlamis, I. (2003). SEWeP: Using Site Semantics and a Taxonomy to Enhance the Web Personalization Process. *Proceedings of the 9th SIGKDD Conf, 2003*.
- Giossos, Y., Koutsouba, M., Lionarakis, A., Skavantzios, K. (2012). *Reconsidering Moore's Transactional Distance Theory*. EuroDL. Retrieved 2 August 2012 from <<http://www.eurodl.org/?article=374>>
- Gokool-Ramdoos, S. (2008). Beyond the theoretical impasse: Extending the applications of transactional distance theory. *The International Review of Research in Open and Distance Learning*, 9(3), 1-17.
- GSMA (2010). *mLearning: A Platform for Educational Opportunities at the Base of the Pyramid*. Retrieved 10 March 2012 from <[http://www.mobileactive.org/files/file\\_uploads/mLearning\\_Report\\_Final\\_Dec2010.pdf](http://www.mobileactive.org/files/file_uploads/mLearning_Report_Final_Dec2010.pdf)>

- Hammersley, B. (2005). *Developing feeds with RSS and Atom*. New York, NY: O'Reilly Media, Inc.
- Keskin, N.O., and Metcalf, D. (2011). The Current Perspectives, Theories and Practices of Mobile Learning. *Turkish Online Journal of Educational Technology*, 10(2), 202-208
- Lee, M.J., Miller, C., and Newnham, L. (2008). RSS and content syndication in higher education: subscribing to a new model of teaching and learning. *Educational Media International*, 45(4), 311-322
- Lewis, J.R. (1995) IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction*, 7(1), 57-78.
- Lundin, J. and Magnusson, M. (2003). Collaborative learning in mobile work. *Journal of Computer Assisted Learning*, 19(3), 273-83.
- Marquet, P. (2010). Obstacles to the use of ICTs in training and consequences for the development of e-learning and m-learning. *Education, Knowledge and Economy*, 4(3), 183-192
- Miner, G., Elder, J., Hill, T., Nisbet, R., and Delen, D. (2012). *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*. New York, NY: Academic Press
- Mobasher, B., Dai, H., Luo, T., Sun, Y., and Zhu, J. (2000). Integrating Web Usage and Content Mining for More Effective Personalization. *Proc. of the International Conference on Ecommerce and Web Technologies (ECWeb2000)*, Greenwich, UK, 2000. pp. 165-176
- Mobasher, B., Dai, H., Luo, T., Sun, Y., and Zhu, J. (2009). Integrating Web Content Mining into Web Usage Mining for finding patterns and predicting users' behaviour", *International Journal of Information Science and Management*, 7(1)
- Mobileactive.org (2011). *Mobile Quizzes for HIV AIDS awareness: Zain and Text to Change*. Retrieved 10 March 2012 from < <http://mobileactive.org/case-studies/how-text-change-and-zain-used-mobile-quizzes-raise-hiv-aids-awareness>>
- Moore, M. G. (1972). Learner autonomy: The second dimension of independent learning. *Convergence*, 5(2), 76-88.
- Moore, M.G. (1973). Toward a theory of independent learning and teaching. *The Journal of Higher Education*, 44(9), 661-679.
- Moore, M.G. (1997). Theory of transactional distance. In D. Keegan (Ed.), *Theoretical principles of distance education* (pp. 22-38). NY: Rutledge Studies in Distance Education.
- Moore, M.G. (2007). The theory of transactional distance. In M. G. Moore (Ed.), *Handbook of distance education* (pp. 89-105). Mahwah, NJ: Lawrence Erlbaum Associates.
- Park, J. (2012). *Computer Science and Convergence*. New York, NY: Springer
- Ragus, M. (2004). *Mobile learning: Handheld innovations in flexible learning, case studies. New Practices in Flexible Learning*. Retrieved 20 December 2011 from <<http://pre2005.fl exiblelearning.net.au/projects/mobilelearning.htm>>
- Saba, F. (1988). Integrated telecommunications systems and instructional transaction. *The American Journal of Distance Education*, 2(3), 17-24.
- Saba, F. (2005). Is distance education losing its identity? Or what should we call or field these days? *21st Annual Conference on Distance Teaching and Learning*, University of Wisconsin-Madison.
- Saba, F., and Shearer, R.L. (1994). Verifying key theoretical concepts in a dynamic model of distance education. *The American Journal of Distance Education*, 8(1), 36-59.

- Safaricom (2011a). *Internet on Mobile*. Retrieved 22 December 2011 from <<http://www.safaricom.co.ke/index.php?id=233> >
- Safaricom (2011b). *M-PESA*. Retrieved 22 December 2011 from <<http://www.safaricom.co.ke/index.php?id=250>>
- Sahami, M. (2009). *Text mining: classification, clustering, and applications*. Boca Raton, FL: CRC Press
- Shearer, R. (2007). Instructional design and the technologies: An overview. In M. G. Moore (Ed.), *Handbook of distance education* (pp. 219-232). Mahwah, NJ: Lawrence Erlbaum Associates.
- Sobh, T. (2010). *Innovations and Advances in Computer Sciences and Engineering*. New York, NY: Springer
- Stumme, G., Berendt, B., and Hotho, A. (2004). *Usage Mining for and on the Semantic Web*. Boston, MA: AAI/MIT Press.
- Suki, N.M., and Suki, N.M. (2011a). Users' Behaviour towards Ubiquitous M-Learning. *Turkish Online Journal of Distance Education (TOJDE)*, 12(3), 118-129
- Suki, N.M., and Suki, N.M. (2011b). Using Mobile Device for Learning: From Students' Perspective. *US-China Education Review*, 8(6a), 44-53
- Tatar, D., Roschelle, J., Vabey, P., and Pennuel, W. R. (2003). Handhelds go to school: Lessons learned. *The IEEE Computer Society*, 30-37.
- USAID (2011). *USAID HIV/AIDS Health Profile for Kenya - September 2010*. Retrieved 22 December 2011 from <[http://www.usaid.gov/our\\_work/global\\_health/aids/Countries/africa/kenya.pdf](http://www.usaid.gov/our_work/global_health/aids/Countries/africa/kenya.pdf) >
- Weira, C.S., Douglasa, G., Richardsonb, T., and Jacka, M. (2010). Usable security: User preferences for authentication methods in eBanking and the effects of experience. *Interacting with Computers*, 22(3), 153-164
- Wittenbrink, H. (2005). *RSS And Atom: Understanding And Implementing Content Feeds And Syndication*. Birmingham: Packt Publishing
- Yeonjeong, P. (2011). A Pedagogical Framework for Mobile Learning: Categorizing Educational Applications of Mobile Technologies into Four Types. *International Review of Research in Open and Distance Learning*, 12(2), 78-102

**APPENDIX**

**Appendix A: Sample Questionnaire**

**Questionnaire Code** \_\_\_\_\_

After using the m-learning HIV AIDS system, state (by ticking the relevant response) whether you strongly disagree, disagree, are neutral, agree or strongly agree with the following statements:

	Strongly Disagree	Disagree	Neutral	Agree	Strongly agree
The information is relevant to what I expected					
	Strongly Disagree	Disagree	Neutral	Agree	Strongly agree
It is easy to use the system					
	Strongly Disagree	Disagree	Neutral	Agree	Strongly agree
The system allowed me to access information on different subjects					
	Strongly Disagree	Disagree	Neutral	Agree	Strongly agree
The recommended feeds were interesting					

What factors made it hard to use the system? (Please explain)

a. \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

b. \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

c. \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

d. \_\_\_\_\_  
 \_\_\_\_\_

---

---

What factors made it hard to gain information (knowledge) from the system? (Please explain)

a. 

---

---

---

---

b. 

---

---

---

---

c. 

---

---

---

---

d. 

---

---

---

---

## Appendix B: Quiz

Please answer all the questions:

1. HIV/AIDS is a contagious disease caused by?
  - a. Virus
  - b. Fungi
  - c. Bacteria
  - d. Witchcraft
  
2. HIV cannot be transmitted through the following except?
  - a. Hugging
  - b. Sharing Food
  - c. Using a public toilet seat
  - d. Sharing razor blades
  
3. Which of these statements is not true about HIV/AIDS?
  - a. HIV still has no known cure
  - b. HIV now is already pandemic (spreading worldwide)
  - c. HIV once contracted develops in months
  - d. HIV weakens and destroys the immune system
  
4. The rate of HIV transmission per sex act is highest during?
  - a. Early stage of HIV infection
  - b. Both early and advanced stages of HIV
  - c. Advanced stage of HIV infection
  - d. Middle stages of HIV infection
  
5. What is the most common manifestation of pulmonary disease in HIV patients?
  - a. Pneumonia
  - b. Bronchiectasis
  - c. Carcinoma
  - d. Asthma
  
6. Which of these is not a preventive measure against HIV/AIDS?
  - a. Use latex condom when engaging in sexual intercourse
  - b. Avoid multiple sexual partners
  - c. Being faithful to each other
  - d. Stay healthy by eating nutritious food
  
7. Which of the following is a mode of transmission of HIV?
  - a. Tears

- b. Saliva
  - c. Urine
  - d. Human Bite
8. In the absence of prophylactic antiretroviral therapy to mothers during pregnancy, labour and delivery and the foetus following birth, the probability of transmission of HIV from mother to infant in developing countries is?
- a. 15-25%
  - b. 50-60%
  - c. 25-35%
  - d. 100%
9. In the absence of prophylactic antiretroviral therapy to mothers during pregnancy, labour and delivery and the foetus following birth, the probability of transmission of HIV from mother to infant in industrialised countries is?
- a. 15-25%
  - b. 25-35%
  - c. 50-60%
  - d. 100%
10. Which of these is the chief predictor of heterosexual transmission of HIV?
- a. Duration of sexual intercourse
  - b. Genetic makeup
  - c. Presence of other sexually transmitted disease
  - d. Plasma viremia



## Appendix C: User Interface Screenshot



### HIV - AIDS m-Learning System

News Message board Quick reports FAQs

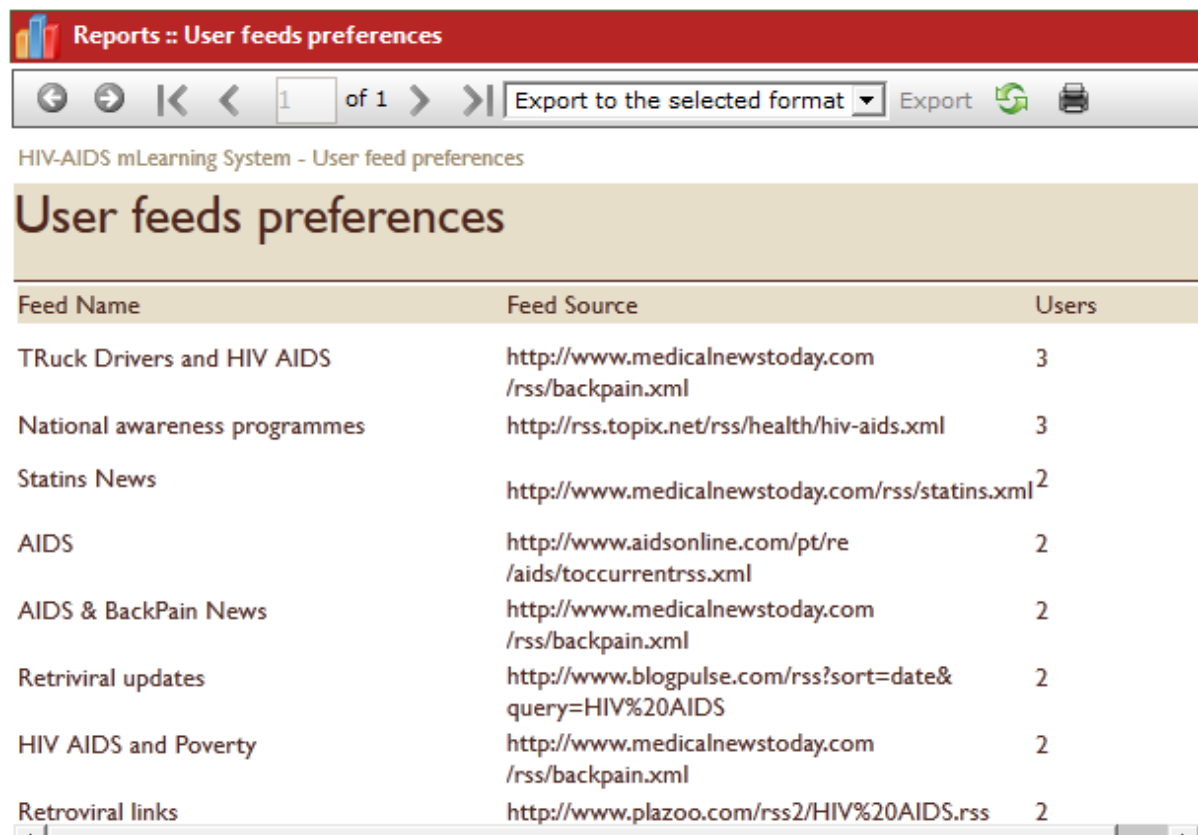
Administration  
 Feed Categories  
 Audit trail  
 User accounts  
 Settings  
 Preferences  
 Exception log

Feeds you may like  
 Prostitution and HIV AIDS  
 Impact of HIV AIDS on Work Life

My News Feeds

Retroviral links	Retroviral updates	Medical News Today (HIV-AIDS)
<p><a href="#">Life Skills Coordinator (Seattle (Univ District))</a>            Sat, 02 Jun 2012 04:31:00 GMT</p> <p>Are you seeking a career that allows you to engage way? Do you have experience working with homeless? If you are a Life Skills Coordinator at YouthCare you would join our team helping at-risk youth prepare for ...</p> <p><a href="#">HIV Health Counselor - Shanti L.I.F.E. Program (ter)</a>            Sat, 02 Jun 2012 00:26:00 GMT</p> <p>JOB OPPORTUNITY Title: HIV Health Counselor De Full Time, Non-Exempt</p> <p>Shanti is primarily dedicated to enhancing the qua</p> <p><a href="#">Day Treatment Clinician (san rafael)</a>            Fri, 01 Jun 2012 22:10:00 GMT</p>	<p>--- Oops!! Could not reach source, try later ---</p>	<p>Sexually transmitted infections (STIs), also known as Sexually Transmitted Diseases (STDs) or Venereal Diseases (VD) are diseases that are passed on from one person to another through sexual contact, and sometimes by genital contact - the infection can be passed on via vaginal intercourse, oral sex, and anal sex...</p> <p><a href="#">HIV Drug Resistance Often Caused By Mutations</a>            Tue, 12 Jun 2012 11:00:00 PDT</p> <p>Harvard researchers have made a critical discovery that could pave the way for new, more effective HIV treatments. The researchers discovered pre-existing mutations in a small number of HIV patients, which can lead to the virus developing a resistance to the drugs that are administered in order to slow the progression of the disease...</p>
Statins News	AIDS & BackPain News	AIDS
<p><a href="#">Side Effects Of Statins</a>            Tue, 12 Jun 2012 02:00:00 PDT</p> <p>In a study of more than 1,000 adults, researchers</p>	<p><a href="#">Researcher Calls For Expansion Of Programs To Identify Potential Drug Abusers And Protect Pain Patients</a>            Fri, 01 Jun 2012 00:00:00 PDT</p>	<p>--- Oops!! Could not reach source, try later ---</p>

## Appendix D: User Feed Preferences



The screenshot shows a web application interface with a red header bar containing the text 'Reports :: User feeds preferences'. Below the header is a navigation bar with a search box containing '1 of 1', a dropdown menu for 'Export to the selected format', and buttons for 'Export', a refresh icon, and a printer icon. The main content area has a title 'User feeds preferences' and a table with three columns: 'Feed Name', 'Feed Source', and 'Users'.

Feed Name	Feed Source	Users
TRuck Drivers and HIV AIDS	<a href="http://www.medicalnewstoday.com/rss/backpain.xml">http://www.medicalnewstoday.com/rss/backpain.xml</a>	3
National awareness programmes	<a href="http://rss.topix.net/rss/health/hiv-aids.xml">http://rss.topix.net/rss/health/hiv-aids.xml</a>	3
Statins News	<a href="http://www.medicalnewstoday.com/rss/statins.xml">http://www.medicalnewstoday.com/rss/statins.xml</a>	2
AIDS	<a href="http://www.aidsonline.com/pt/re/aidstoccurrentrss.xml">http://www.aidsonline.com/pt/re/aidstoccurrentrss.xml</a>	2
AIDS & BackPain News	<a href="http://www.medicalnewstoday.com/rss/backpain.xml">http://www.medicalnewstoday.com/rss/backpain.xml</a>	2
Retriviral updates	<a href="http://www.blogpulse.com/rss?sort=date&amp;query=HIV%20AIDS">http://www.blogpulse.com/rss?sort=date&amp;query=HIV%20AIDS</a>	2
HIV AIDS and Poverty	<a href="http://www.medicalnewstoday.com/rss/backpain.xml">http://www.medicalnewstoday.com/rss/backpain.xml</a>	2
Retroviral links	<a href="http://www.plazoo.com/rss2/HIV%20AIDS.rss">http://www.plazoo.com/rss2/HIV%20AIDS.rss</a>	2