Bringing Knowledge Through AI and SMS

Sam Heather
Department of Computer Science,
The University of York,
sam@heather.sh

November 16, 2014

Abstract

In remote Africa, there are millions of disadvantaged and uneducated individuals, who, in the vast-majority, do not have access to the internet and the access to knowledge that this brings. Outside of their immediate friends and family, individuals can not get access to the information they need on anything from their own body to social problems.

In many parts of the world, the number of people without access to the internet, but access to a mobile phone, is significant. This project aims to research and develop a system capable of bringing knowledge through a question and answer based interactive system, in the language natively spoken by the user, through the use of a simple Artificial Intelligence and an SMS interface. The system will be expandable, such that it can be adjusted to handle questions on any knowledge area.

This project raises ethical issues relating to the responsibility of providing accurate information when in a position of trust, the ethics of machine translation and maintaining user privacy.

It is early days so it is normal that many things are missing from the abstract. You should keep in mind that it should include later on few sentences on method/experiment, result and conclusion.

Contents

1 Useful thoughts from Lilian **8**

2 Introduction **8**

2.1 Background of this project 8

2.2 Motivation for this project 8

2.3 Aims of this project 9

2.4 Structure of this report 9

2.5 Assumptions Within This Project 10

3 Literature Review **11**

3.1 Previous work using SMS 11

3.2 Ethical Issues 12

3.2.1 Ethics of Providing Information 12

3.2.2 Ethics of Translation 13

3.2.3 User Privacy and Data Protection 14

3.3 Software Design Life-cycles 16

4 Method and Requirements **17**

4.1 Chosen Software Development Lifecycle 17

4.2 Plan for Software Development 17

4.3 Requirements 17

4.4 Evaluating Success 17

5 Design **18**

5.1 Software Design 18

5.1.1 Micro-system for Protecting User Identity 18

5.2 Platform, Language and Tools 18

6 Implementation **19**

7 Results **20**

8 Discussion **21**

9 Evaluation **22**

10 Extending this project **23**

11 Conclusion **24**

List of Figures

1 Screenshots of Shy iOS App 9
2 Google Now information box 11
3 Hashing a phone number using an imaginary hashing algorithm 15

List of Tables

1 Example of software database: hashed phone numbers paired with a list of previous returned answer ID's. 15

Statement of Ethics

Pretty much the most important part - write this!

1 Useful thoughts from Lilian

Take a quick look at list of other students previous projects to get an idea of how intro/lit review is structured etc.

Remove page number on title page.

Remember to take notes on when I come across problems, for my conclusion.

2 Introduction

2.1 Background of this project

The inspiration for this project originates from a project the author undertook in September 2014 whilst attending a week long 'hackathon' (Yacht Hack, http://toughhackers.com/yacht-hack-2014/), with Julie Markham and Nicholas Hopper, called Shy. On this project, the author co-created a project to prototype a mobile application that would facilitate the immediate answering of questions that fall within certain categories.

One of the initial aims of Shy was to further the transition from e-learning (learning conducted online, usually with the help of a computer) to m-learning (learning using tools available on a mobile device). M-learning is an especially important topic at the moment, due to the high, and increasing, availability of mobile devices worldwide.

A semi-functional prototype was completed for iOS (shown in Figure 1), although question recommendations were evaluated without knowledge of previous material the user had viewed, and as such there was no knowledge of their interests to support explicitly targeted answers for a question they might search for.

Up until this point, the service was restricted to working on iOS smart devices only, with poor quality question/answer matching.

should add at least one more significant paragraph here, not sure what to write about though

2.2 Motivation for this project

Access to knowledge is, in the opinion of the author, a critical part of modern life. It's also a Human Right, under Article 27 of the Universal Declaration of Human Rights [1]. Indeed, we use tools to access knowledge hundreds of times each day, often unaware that we are doing so. Despite this, hundreds of millions of individuals live without this facility. This project aims to investigate and create a technology to give more people access to knowledge, and thus their human rights.

The project also involves the use of a number of systems to work, including machine translation, an SMS input/output system and a custom system to build a profile

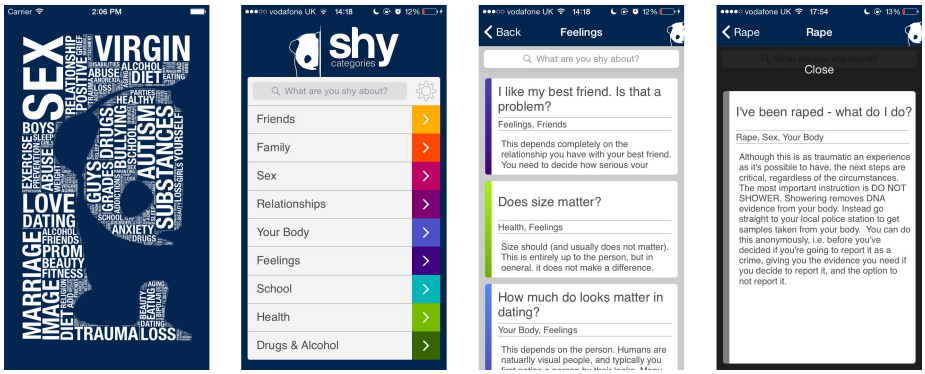

Figure 1: Screenshots of Shy iOS App

on a user, to facilitate high-quality question-answer matching. These, along with complex ethics and privacy issues create an interesting project that draws together many technologies and discussions in a way that can be used as a basis for other projects in the future.

2.3 Aims of this project

The first goal of this project is to assess and address the ethical issues that arise from the software that this project aims to create. These include the responsibility the software has, stemming from it's position of trust, to provide accurate information and protecting the privacy of users by using only necessary information, among others. The author will do this by researching ethical issues relating to the technologies that the project will use, for example machine translation. This will then be used to feed the design process of the software and to specify the expected use-case of it. Finally, the resulting software will be evaluated through experimentation, with volunteers been asked to ask a set of questions within a topic area, in a non-English language, and evaluate the relevance of answers returned.

2.4 Structure of this report

This report starts with a review of pre-existing literature on this topic, in chapter 3, where the author looks at existing software, tools and services of a similar type to those that will be either used in this project or that which this project hopes to create to research the problems that they stumbled across. Comparisons between different software development life-cycles are also discussed.

Chapter 4 describes the method that will be taken to develop the software and service. This covers the software development life-cycle that will be followed, and sets out a plan for when development will take place. Requirements will also be identi-

fied, described and categorised as either functional or non-functional. A method of evaluating the success of the software will also be discussed and chosen.

The design of the software, driven from information collected from the requirements, will be set out in Chapter 5. The individual components that make up the technology will be described and discussed, followed by a description of how these modules will interface with each other to build a complete system. Finally, this section will include a description of the platform, language and other tools that will be used.

described and discussed, or just described?

Chapter 6 contains details about the implementation of the software. This includes information on the tools that were used to create components such as the SMS interface and question/answer matching system.

Results from the evaluation of the system will be presented in Chapter 7. A discussion of these results and their implication on this project is then presented to the reader in Chapter 8.

Chapter 9 contains an evaluation of the software produced and the process taken to complete it. Comparisons will be made to the success criteria, set out for the project in section ¡insert number here¿.

The authors thoughts on potential future work, building on the work produced in this project, are then displayed in chapter 10.

Finally, the main achievements of this project and then the report is concluded in chapter 11 with final thoughts.

2.5 Assumptions Within This Project

This project will make the assumption that users of this service have basic literacy skills in a language supported by the project. Although this is not true for the whole of Africa, expanding the remit of this project to include a 'graphical' user interface that works over SMS is a challenge larger than would fit in an undergraduate dissertation.

3 Literature Review

The point of this is to show that I 'own' the material and subject (competence). Show different points of view, choose a side, express a view on which I prefer and if I agree, critique. Have objective goals.

A significant amount of work has previously been undertaken in the area of making knowledge quickly accessible to us in the form of questions and answers. To take one example, consider the situation of suspecting that a relative was suffering from a disease, for example, Cholera. The initial step would be to search for information before seeking medical assistance. In the connected world, this is easy - a simple Google search returns the answers, as shown in Figure2.

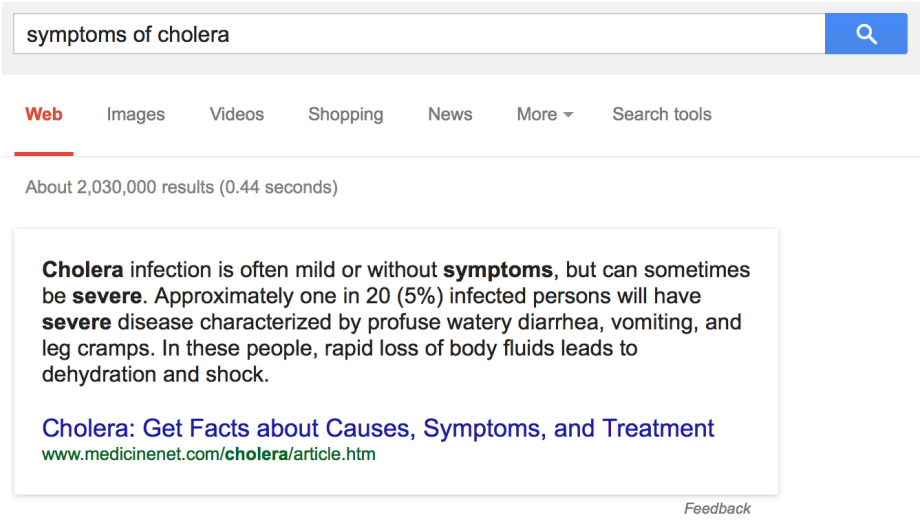

Figure 2: Google Now information box

This is easy to do in the developed world, where up to 75% of the population are connected to the internet [2] (with the population of elderly people been largely responsible for the remaining 25% [3]). This contrasts strongly with the situation in the African continent, where, in 2013, internet usage had only reached 16% [2], a figure largely inflated by South Africa, where 5% of the African population generate 2/3 of internet traffic from the African Continent [2].

3.1 Previous work using SMS

Although the above suggests that the African continent is majoritively disconnected, this is not the case. Because of restrictions in the electricity available, the ways in which a mobile phone can be used in Africa have far surpassed those in the developed

world [4], in-part due to the low power consumption of a mobile phone. Interesting such examples include automated services which SMS HIV/AIDS sufferers, reminding them of medication to take, and providing access to current market information for agriculture and farming, saving farmers from making daily trips (often many kilometres) or relying on out-of-date information from a weekly radio broadcast [5].

Interactive systems have also been developed to operate over SMS. One successful example includes mobile money platforms. These allow for users, from any background, to pay and be paid for goods, and to transfer money across long distances, at negligible cost [5]. One of the most highly adopted services is M-Pesa - which, in Kenya alone, was responsible for £5.7 Billion in transfers in 2012 [1]. M-Pesa gives users a balance linked to a national ID number, from which they can pay for goods by sending an SMS with a cashier (recipient) number or pay outstanding bills in a similar way. Non-subscribers can also use the system, by depositing money with a M-Pesa cashier in exchange for an access code, which can be sent via SMS to a contact, who can subsequently redeem it with their local cashier [5].

This difference in standard use of cell phones is demonstrated in the International Telecommunications Unions's 2013 report [2], which shows that in Europe, for 790 million mobile subscriptions, 53% of subscribers have mobile internet access (422 million), compared to 17% in Africa (93 million have mobile broadband, out of 545 mobile subscribers). This is due to the prohibitively high cost of accessing data services, regardless of the hardware that the user has. In 2012 in Europe, 500 MB of data per month for 12 months cost 1.2% of the average Gross National Income Per Capita (GNI pc). In Africa, the average price was 30x this, at 36.2% of an individuals GNI pc [6].

3.2 Ethical Issues

This project raises a swathe of ethical issues related to translation accuracy, providing information to people in an ethical way and maintaining user privacy.

3.2.1 Ethics of Providing Information

One significant issue for this project stems from providing information that may affect an individual or lead them to take a harmful action. **In a similar way to that which a teacher has a responsibility to teach accurate information to a pupil, due to their position of trust, any service relied upon by a user must equally provide accurate information. The result of not doing this could be providing inaccurate information that leads to a user carrying out an action that causes harm**.

this needs to be waaaayyyyy expanded, but I struggled to find information that was useful. There's lots on remote education, and on basic ethics of teaching, but

[1] Data from Safaricom, M-Pesa operator in Kenya - actual value 817,085,000,000 Kenyan Shillings, converted to GBP on 1st November 2014 at rate of approximately 0.007.

I struggled to find anything specifically relating to ethics of ensuring information you provide is correct (actually I found some material that pointed the other way).

3.2.2 Ethics of Translation

A significant part of this project is represented by the support of multiple languages. It is clear that translation on demand, at scale, needs to be automated by some kind of machine or algorithmic translation.

This project involves two blocks of translation. These are:

- The translation of the user's input into the language of the system (English)
- The translation of the answer to a users question from the system language to their local language.

The second of these raises some issues. To understand these, it is necessary to first understand the two categories of machine translation.

Rule-based machine translation effectively treats human language in a similar style to programming languages. Formal grammars and lexicons are used to represent words that exist in either the source or target language, structures representing the translation of individual words or groups of words. Map structures map individual or groups of words to their translated counterparts, sometimes with multiple results (a 1-many map), from which rules decide which is selected. Maps and rule sets are created by trained computational linguists [7].

More commonly, Statistical Machine Translation (SMT) is used (for example, this is used by Google and Microsoft Bing Translate) [7, 8]. Statistical machine translation learns maps between strings of words of potentially non-equal lengths from pre-existing original texts and their trusted human translations. The accuracy and breadth of language support for translation increases as more source material is analysed by the system, as potentially erroneous or low quality translations can be identified and marked. SMT is also dependent on the quality of the human translation on the input material [7].

The aforementioned issue that is present in statistical machine learning systems comes from the knowledge we assume an individual has and derives from a word. In human translation, this is solved by the translators knowledge of the difference in material culture, allowing them to append necessary information to the resulting translation that the recipient might find useful. In statistical machine learning, cultural awareness of material knowledge is a separate problem on it's own, on this scale. **formal reference to this, probably again from Kenny, but should be able to find better**. To take one example, from Melby (2006),

> "when translating a French menu, a human translator might stop to think that an English speaker in France would appreciate being told that a steak

> tartare is served entirely raw, even if this information is not contained in the original text (because French people might be assumed to know this already). Such a translator would be aware of differences in material culture, and would be able to empathise with the English speaker who might choose to avoid the dish, given more information" [7, 9]

This issue is a result of the translation engine not been capable of taking as input, and using, a complete representation of the expected cultural differences between those who speak the input language and those who speak the output language [9].

Lilian: quote taken from [7], but quote is interpreted from [9] - so I think referencing both is correct, but for this reason rather than last reason we discussed?

below should really be expanded and be in the discussion. Explain the decision to keep translated answers in the database, which are pre-approved, to ensure no cultural misunderstandings.

Because the potential use of this system, regardless of whether this is within this project or by a third-party using the results of this project upon completion, could include distributing information that may be used by an individual making an important decision, cultural confusion such as this could have potentially catastrophic effects.

3.2.3 User Privacy and Data Protection

This project uses a simple Artificial Intelligence (AI) to keep a record of information that has already been returned to a user, to attempt to save sending repetitive information (saving the cost of additional SMS). This raises another ethical issue though - a privacy issue. The material that a user has researched is likely to be sensitive - such is the nature of health related questions. This means that the information on a user has to be kept securely, in a way that an individual user can not be identified.

One way of doing this is using a technology called hashing. Ideal hashing is the technique of taking data as an input and generating an output value (called a 'hash') that is unique to an input. Hash functions are ideally one way functions, where given input data X, the output hash will always be Y, whenever and however the hashing algorithm is executed. This allows a users identify (phone number) to be obscured (hashed), in such a way that the mobile number can not be recovered from the database. When a question is received, the phone number (represented by variable X in the above example) can be hashed and the data for this number looked up from the database, without the database application ever been aware of the phone number of the user. In the example in Figure 3, a phone number (X) is used as input to an imaginary hashing function and the output hash (Y) is shown on the right hand side. As just described, this number can then be used as the key for the user information database, shown in Table 1 to retrieve the information needed for the AI.

reference the above from security engineering ross anderson, chapter 10 and 21?

In practise it is possible for the output of a hashing algorithm not to be unique to a particular input. This is because the range of possible input values is much smaller than

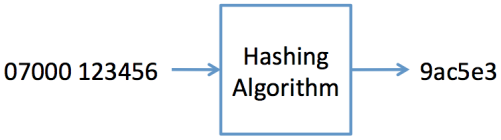

Figure 3: Hashing a phone number using an imaginary hashing algorithm

userId	userData
8f64B2	{previousQuestions:[13,301,170,577]}
9ac5e3	{previousQuestions:[441,56]}
f9dd7e	{previousQuestions:[301,623,89,280,364,621,209]}

Table 1: Example of software database: hashed phone numbers paired with a list of previous returned answer ID's.

the range of output hashes, and because all valid inputs must map to an output, there is repetition in outputs [10]. This property of hashing algorithms can be a security risk for some use-cases, where hashes are used to verify that data has not changed since been hashed. In this case, a third-party may be attempting to generate alternative, malicious data with the same hash as some pre-existing data, to swap them unnoticed [11]. In the use-case of this project however, hashes are only been used to anonymise phone numbers to some unidentifiable string of characters and numbers, meaning that the above-described collision risk is not a security risk.

Two properties that this project does need from it's hashing algorithm are collision resistance and irreversibility. Collision resistance the property of a hashing algorithm that determines the probability of the same output hash for two distinct random inputs [10]. A high collision resistance is necessary to ensure that two users of the system with distinct phone numbers don't get mapped to the same row in the database, resulting in poor quality question/answer matching.

Irreversibility is a necessary property of the hashing algorithm chosen for this project. This ensures that the key used in the database can not be used to retrieve the phone number of the user, thus protecting their identity. Commonly used hashing algorithm families, such as the MD family (e.g. MD5) and SHA family are all designed to be one-way [12][2].

[2]Bruce Schneier is a fellow at Harvard's Berkman center for internet and society. He has been posting regularly for his newsletter and then blog since 1998 (over 16 years) and has published material related to this field throughout this time. He has been involved in the creation of other cryptographic algorithms, such as the Skein hashing algorithm, blowfish block cipher, so may have a vested interest in dis-crediting other algorithms such as MD5 and SHA. However, the essay cited above is written as a result of the Crypto 2004 Conference in California, Schneier has a world-renowned reputation for cryptography and computer security and the algorithm that he suggests should be used at the end is not one that he has involvement in. For these reasons, I believe this source is reliable.

Although hashing algorithms are computationally one-way, they can suffer from another type of vulnerability affecting the security of the original input data. When hashes are used for short strings of information (passwords, for example), a simple way to try and retrieve the original data is to compute a dictionary of the hashes of lots of common passwords, from which matches can be found. This idea has been extended to produce Precomputed Hash Chains and subsequently Rainbow Tables. These are highly efficient data structures that consist of chains of processed input messages and their hashes (with a reduce function to reach the next item in the chain), allowing original input data to be looked-up from it's hash. Both Precomputed Hash Chains and Rainbow Tables tend to be Terabyte's in size, and as such they only exist in a significant form for the most common hash functions (including MD5 and SHA) [13]. Within this project, this raises privacy implications, as hashes within the database would be 'convertible' to phone numbers.

perhaps the below should move to method and requirements - TODO needs to be cleaned up as well.

In modern password authentication, a tool called bcrypt is used, which 'salts' passwords (additional unique data is added to each password before hashing), nullifying the use of Precomputed Hash Chains and subsequently Rainbow Tables. However, this would prevent efficient lookup of phone numbers in the communications part of the application (where answers and sent back to the user). As such, the best way of protecting a users phone number in the event of a database hack is to use a combination of a largely uncommon hashing algorithm, on the basis of it is unlikely anybody will have wasted vast computation power for a table that's useful in such a small number of situations, a hashing algorithm with a large output size, again because it makes a table more infeasible (because of size constraints) and a further change to the output hash, specific to this application. This way, a hacker would need to know this change (from the source code) as well as the password database, to be able to make use of a hashing table. The above is commonly implemented by hashing the result of one hash with another hashing algorithm.

3.3 Software Design Life-cycles

Describe and compare some software design life-cycles

4 Method and Requirements

4.1 Chosen Software Development Lifecycle

4.2 Plan for Software Development

4.3 Requirements

4.4 Evaluating Success

5 Design

5.1 Software Design

5.1.1 Micro-system for Protecting User Identity

explain the system that uses a middle-application to parse numbers into anonymous hashes, which are used in the database, protecting the user's identity for all but the second their query is been processed

5.2 Platform, Language and Tools

6 Implementation

This section will contain information on how data is stored, the structure of databases, and the format of and method through which data is sent to an SMS provider.

7 Results

This section’s content...

8 Discussion

This section’s content...

9 Evaluation

This section's content...

10 Extending this project

Section not complete If the author was able to expand this project, areas for expansion include:

- Addressing the issue of literacy
- Automating the expansion of the database ...

11 Conclusion

This section's content...

Melby reference not right, publisher weird - ask Lilian in next session. Url Commented in TEX here

References

[1] U. N. international community, "The universal declaration of human rights," 1948. [Online]. Available: http://www.un.org/en/documents/udhr

[2] I. T. U. (ITU), "Key ict indicators for developed and developing countries and the world (totals and penetration rates)," Geneva, 2013. [Online]. Available: http://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2013/ITU_Key_2005-2013_ICT_data.xls

[3] O. for National Statistics, "Internet access quarterly update, q1 2014 release," Online report, London, Tech. Rep., 2014. [Online]. Available: http://www.ons.gov.uk/ons/rel/rdit2/internet-access-quarterly-update/q1-2014/index.html

[4] K. Fox. (2011, Jul.) Africa's mobile economic revolution. [Online]. Available: http://www.theguardian.com/technology/2011/jul/24/mobile-phones-africa-microfinance-farming

[5] J. C. Aker and I. M. Mbiti, "Mobile phones and economic development in africa," *The Journal of Economic Perspectives*, vol. 24, no. 3, pp. pp. 207–232, 2010. [Online]. Available: http://www.jstor.org/stable/20799163

[6] I. T. Union, "Measuring the information society," International Telecommunication Union, Geneva, Tech. Rep., 2013. [Online]. Available: http://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2013/MIS2013_without_Annex_4.pdf

[7] D. Kenny, "The ethics of machine translation," 2011.

[8] G. Inc. (2014) Google research: Find out how our translations are created. [Online]. Available: http://translate.google.co.uk/about/intl/en_ALL/

[9] A. K. Melby, "Why cant a computer translate more like a person," 2006.

[10] S. Goldwasser and M. Bellare. University Lecture Notes. Accessed: 2014-11-15. [Online]. Available: http://cseweb.ucsd.edu/~mihir/papers/gb.pdf

[11] R. Anderson, *Security Engineering*, 2nd ed. Indiana, United States: Wiley Publishing, Inc., 2008.

[12] B. Schneier. (2004) Cryptanalysis of md5 and sha: Time for a new standard. Accessed: 2014-11-16. [Online]. Available: https://www.schneier.com/essays/archives/2004/08/cryptanalysis_of_md5.html

[13] C. Teat and S. Peltsverger, "The security of cryptographic hashes," in *Proceedings of the 49th Annual Southeast Regional Conference*, ser. ACM-SE '11. New York, NY, USA: ACM, 2011, pp. 103–108. [Online]. Available: http://doi.acm.org/10.1145/2016039.2016072