**Digitized Clinical Data Transfer using Digital Ledger**

Project Id :

Project Proposal Report

Bachelor of Science Special (Hons) Degree in Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

27th March 2018

**Digitized Clinical Data Transfer using Digital Ledger**

Project Id :

Project Proposal Report

Dias A.M.A.S – IT15029928

W.H.M.L.H Bandara – IT15110312

Widanagamachchi L.L– IT15098474

Geekiyanage C.L –IT13091880

Supervisor: Mr. Lakmal Rupasinghe

Bachelor of Science Special (Hons) Degree in Information Technology

Department of Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

March 2018

**DECLARATION**

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

|  |  |  |  |
| --- | --- | --- | --- |
|  | STUDENT NAME | STUDENT NO. | SIGNATURE |
| 1 | Dias A.M.A.S (GROUP LEADER) | IT15029928 |  |
| 2 | W.H.M.L.H Bandara | IT15110312 |  |
| 3 | Widanagamachchi L.L | IT15098474 |  |
| 4 | Geekiyanage C. L | IT13091880 |  |

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Name of supervisor: Mr. Lakmal Rupasinghe

Signature of supervisor: Date:

# **ABSTRACT**

## TABLE OF CONTENTS

[**ABSTRACT** 4](#_Toc508738914)

[TABLE OF CONTENTS 5](#_Toc508738915)

[**1.** **INTRODUCTION** 6](#_Toc508738916)

[1.1 Background 6](#_Toc508738917)

[1.2 Literature survey 6](#_Toc508738918)

[1.3 Research Gap 6](#_Toc508738919)

[1.4 Research Problem 6](#_Toc508738920)

[**2.** **OBJECTIVES** 7](#_Toc508738921)

[2.1 Main Objective 7](#_Toc508738922)

[2.2 Specific Objectives 7](#_Toc508738923)

[**3.** **METHODOLOGY** 8](#_Toc508738924)

[2.1 System Overview 8](#_Toc508738925)

[3.2 System Architecture 8](#_Toc508738926)

[**4.** **DESCRIPTION OF PERSONAL AND FACILITIES** 11](#_Toc508738927)

[**5.** **BUDGET AND BUDGET JUSTIFICATION** 12](#_Toc508738928)

[**6.** **REFERENCES LIST** 13](#_Toc508738929)

[**7.** **APPENDICES** 14](#_Toc508738930)

# 

# **INTRODUCTION**

## Background

## Literature survey

Low resolution images not clear and hard to get actual information. Template correction method is use for hyperspectral images. There are three parts in this technique, building the template, image match and template correction and object recognition. These techniques apply for high-resolution images, and objects and identify the objects in the images. Image can be consisting of degraded quality due to poor quality equipment, data transmission issues etc. There are several types of noises Impulse noise, Amplifier noise, Short noise which are case for low quality images. Those type of images can be recreated by removing noise or unnecessary data from the original image.[1]

Template matching use to identify similarities between two objects. It can be categorized into area-base approaches and feature-based approaches. In the area-based approach, there are some cases which direct matching of template and target image is hard. Split the template image into sub-images and perform matching can be used as a solution. In native template matching usually without scaling the target image scan.[2]

Template matching technique for searching words in document images, used to search and finding the location of a small part of image in large image and used in Optical Character Recognition and search and convert scanned image to text. This technique can be used to compare single characters and multiple characters from word document. Input image capture and store in image file format JPG/JPEG, PNG, TIFF etc. Characters detect using template string. Converted color image convert into black and white image which is called binary image. Normalized cross correlation, correlation methods and performance index method use for the analyze process. [3]

Detection of text-Based Traffic Signs, First System detects the text-based traffic signs then convert those images to 2-D frame. The dimensions and heights of road signs are fixed. To detect symbol-based traffic signs this system uses MSER and HSV color thresholding. MSER maintain shape of the image and HSV detect additional text signs. Before read the text perspective transformation apply to vertically align text to improve the detection[4]

OCR belongs to machine recognition technique performing automatic identification. It identifies objects automatically from the data entered to the system.[5] Optical Character Recognition (OCR) is a technique which is used to convert scanned image into editable text format. The results depend on text pre-processing and segmentation algorithms. The performance can be measured using accuracy and error rate. Images from different formats process and remove noises to enhance the quality. Segmentation algorithm use to extract words from the image. Then OCR used to classify characters and patterns. The hand written and printed characters can be recognized using OCR but performance depend on the quality of input document. [6]

In The Text to Speech (TTS) conversion technology used to converts scanned or printed text image, handwritten text into editable text .It get the image and convert color image to gray scale image by threshold operation and Character extracted and resized according to the template size.[7]

Blockchain is using distributed data storage, it is a decentralized data transferring method. Each block chain is encrypted and contains useful information, a reference to the preceding block and an answer to a mathematical problem which is used to validate the information that block contains. A copy of the block chain is stored in every computer. When a block is added to the block chain it cannot be deleted and also it can be accessed by every computer in the network.[8]

Bitcoin-NG is using to overcome transaction latencies and bandwidth problems. It is a blockchain protocol that serializes transactions it allows better latency and bandwidth without sacrificing other properties.[9]

Sidechaining is a mechanism which allows block chain to be securely used within a completely separate blockchain. Sidechains are independent and runs parallel to the main block chain. If a sidechain got hacked it won’t damage other chains, damage will be limited only to that chain. Sidechain can be move back to the main chain if necessary.[10]

There are three types of block chains public, private and consortium blockchains. In public blockchains there is no middle man. Each transaction is verified and synced with every node affiliated with the blockchain before it is written to the system. In private block chains there is a middle man; the efficiency is high and blockchain will be completed faster. Consortium blockchains are partially private; it operates under the leadership of a group instead of a single entity. Efficiency and the transaction privacy is high in consortium blockchains.[11]

Proof of work is a piece of data which is difficult to produce but easy for others to verify and which satisfies certain requirements. This method cause lot of electricity and computational power. Proof of stake is a different way of validating which is using low electricity and computational power.[12]

Ethereum is a public blockchain based distributed computing platform. It can work as a computer; butthe performance will be slower than most of current computers since it has a transaction time of around 12 seconds.[13]

Proof of Interoperability is an alternative method for network consensus that avoids some of the disadvantages of Proof of Work. It is designed to leverage the effort required to reach network consensus to do something intrinsically valuable.[14]

Blockchain enables people to access the records of every transaction they make, as it permanently records transaction history at every node of the blockchain. Blockchain transactions are recorded using public keys and private keys.[15]

Practical byzantine fault tolerance (PBFT) is a replication algorithm to tolerate byzantine faults. PBFT requires that every node is known to the network.[16]

## Research Gap

## Research Problem

In Sri Lanka most of the Hospitals store Clinical Data such as infections, drugs used for certain earlier diseases in papers. When patient go for a medical checkup doctor first check patient history. For an example, if the patient is suffering from diabetes problem then he will look for exceptions like diabetes, cholesterol or any allergies for certain medicines. After going through the patient history, he/she will come to a conclusion on how the patient should be treating, what kind of tests should be done first and what kind of treatment he or she must follow.

As well as that If the person is a suffering from STD (Sexual Transmitted Disease) then he/she will identify by a number which is given by hospital. Those people do not like to present in publicly. Assume a patient is transferred by one hospital to another hospital. According to the above scenario doctor needs to check patient history before the treatments. In that situation the hospital which patient is transferred by, have to send all the paper documents that belongs to particular patient. That practice is not good enough for the patient who is suffering from STD, and VIP patients, because that is not a secure way to transfer clinical records of a patient.

Moreover, Most of the Doctors letters are not clear and different from one to other. Therefore, there is a high probability to read that document in wrong way. It might be a huge risk.

.

# **OBJECTIVES**

## Main Objective

We came up with a solution to overcome those issues. One solution is to scan the paper document and identify letters of different doctors using image processing. Then feed them to the Algorithm and get the actual outcomes (whole data in the document) using Machine Learning. After that convert digitized data into HL7 (Health Level Seven) standard. Since records are in HL7 format it will be easy to transfer data among two Hospitals and keep the interoperability between two applications.

Since the patient clinical documents need to be secure, we suggest transferring data through blockchain, because there is a huge risk of stealing or corrupt those records in digital word nowadays. For an Example WannaCry attack for European hospitals. But in the blockchain if someone wants to hack patient data he /she need to hack all the chain which is not possible, because blockchain is decentralized and immutable cryptography.

## Specific Objectives

# **METHODOLOGY**

## 2.1 System Overview

## 3.2 System Architecture

## 

## 

*Figure 3.2.2.1 – Extensible Database Abstraction Architecture*

## 

# **DESCRIPTION OF PERSONAL AND FACILITIES**

# **BUDGET AND BUDGET JUSTIFICATION**

# **REFERENCES LIST**

[1] D. Gupta and S. Goswami, “Object Recognition based on Template Matching and Correlation Method in Hyperspectral Images Divya Gupta,” *Int. J. Comput. Appl.*, vol. 166, no. 11, pp. 38–43, 2017.

[2] N. S. Hashemi, R. B. Aghdam, A. S. B. Ghiasi, and P. Fatemi, “Template Matching Advances and Applications in Image Analysis,” *Am. Sci. Res. J. Eng. Technol. Sci.*, pp. 91–96, 2016.

[3] V. S and S. A, “Template Matching Technique for Searching Words in Document Images,” *Int. J. Cybern. Informatics*, vol. 4, no. 6, pp. 25–35, 2015.

[4] J. Greenhalgh and M. Mirmehdi, “Recognizing text-based traffic signs,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1360–1369, 2015.

[5] A. Chaudhuri, K. Mandaviya, P. Badelia, and S. K Ghosh, *Optical Character Recognition Systems for Different Languages with Soft Computing*, vol. 352. 2017.

[6] V. S and S. A, “Performance Comparison of OCR Tools,” *Int. J. UbiComp*, vol. 6, no. 3, pp. 19–30, 2015.

[7] G. K. Sagar, “Real Time Implementation of Optical Character Recognition Based TTS System using Raspberry pi,” no. 7, pp. 149–156, 2017.

[8] A. Wright and P. De Filippi, “Decentralized Blockchain Technology and the Rise of Lex Cryptographia.”

[9] I. Eyal, A. E. Gencer, E. G. Sirer, and R. van Renesse, “Bitcoin-NG: A Scalable Blockchain Protocol,” 2015.

[10] A. Back, M. Corallo, and L. Dashjr, “Enabling blockchain innovations with pegged sidechains,” *URL http//www.*, pp. 1–25, 2014.

[11] J. Kang, R. Yu, X. Huang, S. Maharjan, Y. Zhang, and E. Hossain, “Enabling Localized Peer-to-Peer Electricity Trading among Plug-in Hybrid Electric Vehicles Using Consortium Blockchains,” *IEEE Trans. Ind. Informatics*, vol. 13, no. 6, pp. 3154–3164, 2017.

[12] I.-C. Lin and T.-C. Liao, “A Survey of Blockchain Security Issues and Challenges,” *Int. J. Netw. Secur.*, vol. 1919, no. 55, pp. 653–659, 2017.

[13] S. Huh, S. Cho, and S. Kim, “Managing IoT devices using blockchain platform,” *Int. Conf. Adv. Commun. Technol. ICACT*, pp. 464–467, 2017.

[14] K. Peterson, R. Deeduvanu, P. Kanjamala, and K. Boles, “A Blockchain-Based Approach to Health Information Exchange Networks,” *Mayo Clin.*, no. 1, p. 10, 2016.

[15] J. Sun, J. Yan, and K. Z. K. Zhang, “Blockchain-based sharing services: What blockchain technology can contribute to smart cities,” *Financ. Innov.*, vol. 2, no. 1, p. 26, 2016.

[16] Z. Zheng, S. Xie, H.-N. Dai, and H. Wang, “Blockchain Challenges and Opportunities: A Survey,” *Int. J. Web Grid Serv.*, no. January, pp. 1–24, 2017.

# **APPENDICES**