MAKERERE UNIVERSITY

COLLEGE OF COMPUTING AND INFORMATION SCIENCES

ACCURATE DATA ENTRY SYSTEM

By
CSCD007

DEPARTMENT OF COMPUTER SCIENCE
SCHOOL OF COMPUTING AND INFORMATICS TECHNOLOGY
A Research proposal submitted to the school of computing and informatics technology
in partial fulfillment of the Final Exam and
requirements for the award of the degree of Bachelor of science in computer science

Supervisor: Ernest Mwebaze
$APRIL, 17^{th}, 2018$
prepared by:

No.	Student Name	RegNo	Signature
1	KAZOOBA B LAWRENCE	16/U/5830/PS	
2	AWAT LYNETTE	16/U/3973/PS	
3	GIRAMIA PATRICIA	16/U/4831/EVE	
4	ORIKIRIZA OSCAR	16/U/1055	
5			
6			

Research Proposal,

Contents

1 Introduction

2 Background of Study

Many significant technological changes have occurred in the IT industry since the beginning of the 21st century., Data entry was by means of hardcopy, big data and it was always fed into the books, ledgers. But there has been an outstanding advance in technology with powerful tools like the OdK collect and package, Google cloud computing and platforms and many others which have really helped with entry and computing abilities. But even with all these technology advancements their still some glitches in data entry such as data inconsistency and data invalidation which we seek to curb in this implementation written in this proposal

3 Statement of the problem

The problem being solved or implemented is Users entering inaccurate data which sometimes results into data inconsistency, data invalidation and so this results into very wrong or insufficient and inaccurate entry of data.

4 Objectives

- To curb data inconsistency where same data is just stored n different formats lets say duplication of files.

- To help in data validation ensuring data cleansing and improve data quality.

- To help users curb entering inaccurate data and learn to always have a good data entry system.

4.1 Main objectives

To help users to generally learn and curb any difficulties that entails data entry and totally learns how to enter accurate data.

4.2 Specific objectives

- To help in accuracy in data entry and help users easily enter data

- to help in ease and convenience during data entry..

- To help in data consistency and data validation especially when entering data.

- To generally help all sorts of data users, literate, illiterate and semi-illiterate users who all enter data.

5 Scope

This research is going to contain samples of inaccurate data, how the solution is going to be implemented.. Examples of data inconsistency and also samples of how data can be validated and stored properly so as to avoid unclean data. Sources of big data and how it was done before the 21st century. Examples of the technological advancements of data entry. The research is going to entail all kids of users and the problems they face as they do data entry, such as the very literate, semi-literate and the illiterate.

6 Research Significance of the study

This study or research should help all kinds of users who do data entry to be better when entering data and to avoid all the difficulties that come with data entry and should also provide all the examples that deal with data entry and big data before and after the 21st century. The technological advancements that have come with dealing with data entry , the new ways of data entry that are convenient and their glitches and the solutions that will be implemented.

7 Literature Review

Data validation is a task that is usually performed in all National Statistical Institutes, in all the statistical domains. It is indeed not a new practice, and although it has been performed for many years, nevertheless, procedures.In the 1950's, researchers began to study quality issues, especially for the quality

of products, and a series of definitions, for example, quality is the degree to which a set of inherent characteristics full fill the requirements (General Administration of Quality Supervision, 2008);fitness for use (Wang Strong, 1996); conformance to requirements (Crosby, 1988) were published. Later, with the rapid development of information technology, research turned to the study of the data quality. The Computer Network Information Center of the Chinese Academy of Sciences proposed a data quality assessment method and index system (Data Application Environment Construction and Service of the Chinese Academy of Sciences, 2009) in which data quality is divided into three categories including external form quality, content quality, and the utility of quality. Each category is subdivided into quality characteristics and an evaluation index. In summary, the existing studies focus on two aspects: a series of studies of web data quality and studies in specific areas, such as biology, medicine, geophysics, telecommunications, scientific data.Big data as an emerging technology, acquires more and more attention but also lacks research results in establishing big data quality.

8 Methodology

8.1 Description of System

An activity aimed at verifying whether the value of a data item comes from the given.(finite or infinite). In this definition, the validation activity is referred to a single data item without any explicit mention to the verification of consistency among different data items.If the definition is interpreted as stating that validation is the verification that values of single variables belong to set of prefixed sets of values (domains) it is too strict since important activities generally considered part of data validation are left out. On the other hand, if the acceptance/rejection of a data item were intended as the final action deriving from some complex procedure of error localization, the previous definition of validation would be too inclusive since it would involve also phases of the editing process not strictly related to the validation process.

8.2 Development of data entry system

Data Validation is an activity verifying whether or not a combination of values is a member of a set of acceptable combinations. The set of 'acceptable

values' may be a set of possible values for a single field. But under this definition it may also be a set of valid value combinations for a record, column, or larger collection of data. Therefore, credibility has become an important quality dimension. However, social media data are usually unstructured, and their consistency and integrity are not suitable for evaluation.However, due to the lack of uniform standards, data storage software and data formats vary widely. Thus, it is difficult to regard consistency as a quality dimension, and the needs of regarding timeliness and completeness as data quality dimensions are not high.

8.3   Data Collection Tools/techniques

Each quality dimension needs different measurement tools, techniques, and processes, which leads to differences in assessment times, costs, and human resources. In a clear understanding of the work required to assess each dimension, choosing those dimensions that meet the needs can well define a project's scope. The preliminary assessment results of data quality dimensions determine the baseline while the remaining assessment as a part of the business process is used for continuous detection and information improvement. There are many ways to collect data through questionnaires,surveys,interviews

8.4   sample design

the sample design for the data entry system is designed in several interfaces that are captioned and shown below in the different figures.it helps the users to easily use the system.
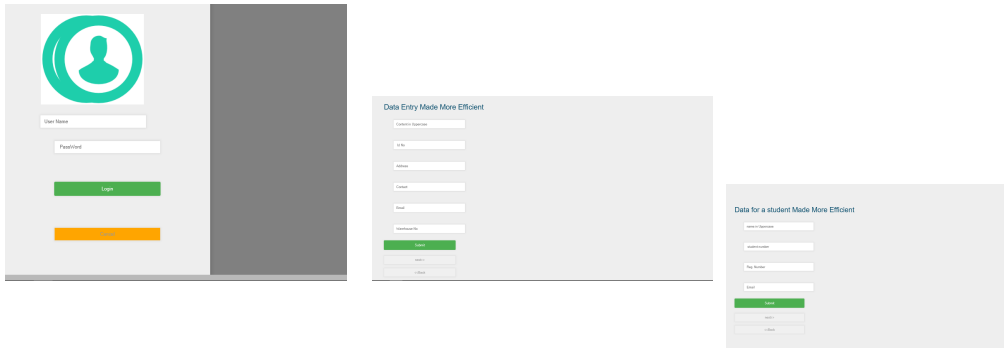
Figure 1: Samples of the interfaces

9 References

1 .Crosby, P. B. (1988) Quality is Free: The Art of Making Quality Certain, New York: McGraw-Hill.

2 .Data Application Environment Construction and Service of Chinese Academy of Sciences (2009) Data Qual-ity Evaluation Method and Index System. Retrieved October 30, 2013 from the World Wide Web: http:// www.csdb.cn/upload/101205/1012052021536150.pdf

3 .Demchenko, Y., Grosso, P., de Laat, C., et al. (2013) Addressing Big Data Issues in Scientific Data Infrastruc-ture. Procedures of the 2013 International Conference on Collaboration Technologies and Systems, California: ACM, pp 4855.

4 .Feng, Z. Y., Guo, X. H., Zeng, D. J., et al. (2013) On the research frontiers of business management in the con-text of Big Data. Journal of Management Sciences in China 16(01), pp 19.

5 .Gantz J Reinsel,D.(2012) THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Big-gest Growth in the Far East. Retrieved February, 2013 from the World Wide Web: http://www.emc.com/ collateral/analyst-reports/idc-digital-universe-western-europe.pdf