CHAPTER TWO: **LITERATURE REVIEW**

**2.0 INTRODUCTION**

This chapter presents the theoretical background of the study and the technologies used. The literature review is also discussed.

**2.1 Theoretical background**

2.2.1 Conceptual Background

The idea of mail started in other to enhance communication between two people, organizations, etc. It is done in the sense that the sender writes a letter which will contain the address of the receiver together with his own address is its informal letter, after which he/she proceeds to poll station to drop it in the mail box, letter is appropriately delivered to the receiver address. This had been a traditional way of sending a mail of it takes much time to send and receive mails, this lead to developing and email systems, which is a means by which sender sends a mail to the receiver using an electronic mail which is very fast means of sending mails. But still there is a need to classify the mails. According to [10] MailCat which is an intelligent assistant that helps users organize their e-mail into folders. It uses a text classifier to learn each user’s mail-filing habits. MailCat uses what it learns to predict the three folders in which the user is most likely to place each incoming message. It then provides shortcut buttons to file each message into one of these three folders. When one of MailCat’s predictions is correct, the effort required to file a message is reduced to a single button click. When people use electronic mail, they can communicate even when they are not physically or close [9]. Thus, it is not surprising that most studies report that the use of electronic mail increases organizational communication. [12] Provides a method for displaying information that represents a quantity of received email messages that have been held for processing by a specific entity for a predefined time period in an email response management system. In the implementation, the method includes determining first value associated with a first quantity of received email messages that have been held for processing by the specific entity for a first predefined time period, wherein each email message has been sent from end user. According to [13] Managing email over the long term is challenging because of email’s essential characteristics. An email originates in an electronic format, but it can also exist in a specific type of computer file or file format. Email tends to reside in a proprietary file format in an email system; although there is interaction between email software systems, email exists as a file format that is owned and controlled by a single software company and is not easily exportable to another environment. For these reasons, long term or permanent emails must be created in or converted to non-proprietary formats for preservation.

2.2.1 Technologies used

The technology used in this project is server/client technology. These include (Xampp, PHP, CSS, MySQL, and HTML) and relational database technology.

Xampp is used in this system as a local host that help the developer to test the application locally to know how the system will behave when deployed online. It is an integrated server package of MySQL, and PhpMyAdmin, that run from removable drive. It is a database, File Transfer Protocol (FTP) and mail server package for windows. It requires no configuration and has control panel starting/stopping mode.

HTML is a language for describing web pages. It stands for Hypertext Mark-up **Language** and it is the language upon which all web pages on the "www" (World Wide Web) are based. The HTML enables your web browser to be viewed and also enhance user computer interface which add interaction and relationship attributes to build and organize its codes.

PHP (Hypertext Pre-processor) is a server side scripting language that can be written into your HTML scripts and used to create dynamic web pages.

**CSS** (Cascading Style Sheets) is a collection of formatting rules that control the appearance of content in a web page. Using CSS styles to format a page separates content from presentation.

MySQL is an Open Source relational database management system (RDBMS) that uses Structured Query Language (SQL). In English, MYSQL is a server-side program (which means it is installed on a server, not onto your computer) where you can store your databases and access them in the internet.

**2.2 Review of related literature**

According to [2] Email based communication over the course of globalization in recent years has transformed into an all-encompassing form of interaction and requires automatic processes to control email correspondence in an environment of increasing email database. Relevance characteristics defining class of email in general includes the topic of the mail and the sender of the email along with the body of email. Intelligent reply algorithms can be employed in which machine learning methods can accommodate email content using probabilistic methods to classify context and nature of email. This helps in correct selection of template for email reply. Still redundant information can cause errors in classifying an email. Natural Language Processing (NLP) possess potential in optimizing text classification due to its direct relation with language structure. An enhancement is presented in this research to address email management issues by incorporating optimized information extraction for email classification along with generating relevant dictionaries as emails vary in categories and increases in volume. The open hypothesis of this research is that the underlying concept to fan email is communicating a message in form of text. It is observed that NLP techniques improve performance of Intelligent Email Reply algorithm enhancing its ability to classify and generate email responses with minimal errors using probabilistic methods. Improved algorithm is functionally automated with machine learning techniques to assist email users who find it difficult to manage bulk variety of emails.

Email has been one of the most commonly used tool for communication in the recent years and email management has evolved as a major challenge due to prevailing situation of online email congestion. [5] Presents a novel algorithm for automatic email response methodology in an Email Management System to minimize email overload. The proposed model uses Bayes classifier to categorize emails into classes and generate suitable replies to these classes using information extraction and template filling. Our research aims to intelligently automate email response using Naïve Bayesian classification and formulate probabilistic dictionaries for accurate information extraction. This research will help in reducing email overload and unavoidable congestion by employing novel email response architecture for an email management system.

The Semantic Web was designed to represent the enormous data that is existing on the World Wide Web in a machine readable format. [4] Research shows the long period of time that was spent on the Emails for communication and information exchange. Adding the semantics to the existing Email systems could not only provide for the valuable usage of time and resources, but also refreshes the meaning of Email communication. The presented research work examines the ontology extraction process from the Email systems adopting scalable pattern rules that is based on the extracted techniques. The proposed architecture is designed to handle the unstructured Emails and the ontologies that are extracted from the Email which is divided into four main components as follows: the Ontology Learning Component, the Management Component, the Semantic Email Component and the Client Side Plugin.

[7] Develops a ripper classification algorithm which is used in automatic filtering of email. Its architecture is based on rule based structure to sort email. Ripper has the ability to automatically generate rules for selecting keywords instead of manual selection. Its advantage is that it is fast and able to deal with a large set of email attributes. However keyword extraction rules have to be constructed for every possible class and it is easy for emails to be mixed up or irrelevant attribute information is extracted. These extraction rules only makes binary decisions, so there predictions are not fully deterministic as strict rules may cause emails to get mixed up in wrong classes. This is caused by attributes competing against each other for possession of an email messages. Such a system is also unable to learn adaptively. Whenever the attributes of the email are changed, the rules will be recreated from ground up in order accurately assign an email class. This task needs time to complete and other dependent task like automatic email reply is directly affected.

[3] Describes a series of interviews that examine the ways that professional office workers use electronic mail to manage their daily work. The purpose is to generate hypotheses for future research. A number of implications for the design of flexible email systems are discussed. Two principal claims are made, first, the use of electronic mail in strikingly diverse, although not infinitely so; Individuals vary both in objective measures of email use and in preferred strategies for managing work electronically. Feelings of control are similarly diverse and are related to the size of the user’s inbox, number of folders, and subscriptions to distribution lists. This diversity implies that one’s own experiences with electronic mail are unlikely to provide sufficient understanding of other’s uses of email. According to the literature, email designers should thus seek flexible primitives that capture the important dimensions of use and provide flexibility for a wide range of users. The second claim is that electronic mail is more than just a communication system. Users archive messages for subject retrieval, prioritize messages to sequence work activities, and delegate tasks via email. Taxonomy of work management is proposed in which email is used for information management, time management, and task management activities.

In this paper [8], the connection between the distinctive association techniques and the time important to utilize a specific methodology is shown by a numerical model in light of keystroke-level examination. The model gauges that time utilization for filing and recovering email messages for singular clients. Other than clarifying why clients create diverse techniques to sort out email messages for singular clients. Other than clarifying why clients create different techniques to sort out email messages, the model can likewise be utilized to prompt clients separately when to begin utilizing envelopes, clean messages, take in the inquiry functionality and utilizing channels to store messages.

The development of the proposed system will help the directors or managers of corporate organization to plan ahead of events, by effective scheduling of a mail which that will be sent before or on the day of the event. The only thing that the user needs to do is to compose a mail by giving it time and date then save it in the database. Once the date and time is due, the mail will be sent appropriately to the specified receiver.