

UNIVERSITI TEKNOLOGI MARA

**Automated Event Extraction From Email Using Pattern
Matching Technique**

Ahmad Tarmizi Bin Razali

Thesis submitted in fulfillment of the requirements for
BSC. (HONS.) INTELLIGENT SYSTEM
Faculty of Information Technology And
Quantitative Science

May 2007

ACKNOWLEDGEMENTS

Alhamdulillah, praise Allah for His Almighty and Graciousness. With His blessings, I was able to complete this project paper within the time duration given.

My greatest appreciation goes to my project supervisor, En Zaki Zakaria, who has given me such a good idea that has opened my minds to what is currently needed in the project with her strong knowledge in building Intelligent System. Without her endless advice and support, I would not have completed this work.

Secondly, my heart appreciation also goes to my lecturer for ITS 690 (IT Project), Pn. Shuzlina Abd. Rahman, who has been tolerant and patient in guiding me in my efforts to complete this project paper. And also full of respect for CS223's lecturer for their guidance during my study.

My thank you also extends to my beloved parents for their support and also to Razali Bin. Ismail and Radziah Bt. Puteh who are very helpful in giving me information and creative idea for this project. Lastly, thanks to all my friends and my classmates.

Thank you very much.

ABSTRACT

Email has become a popular communication media today. It's not only used for communication but also become an announcement media. One such example is announcements of different types of event. In many environments, seminar announcements, meeting announcements, workshop and etc., flood the inbox of users. A significant amount of time is spent for each user to read the message, extract some important information from the announcement and then update their calendar. This process is time consuming and tedious. And because of that, users just leave important announcements in the inbox and then miss them. Therefore, the objective of this research is to develop an event extraction engine to make the process of jotting the information into calendar will be done automatically. The information extraction is done using the pattern matching (also known as knowledge engineering) approach.

TABLE OF CONTENT

ACKNOWLEDGEMENTS	i
ABSTRACT	ii
TABLE OF CONTENT	iii
LIST OF TABLES	vi
LIST OF FIGURES	vi
CHAPTER 1 : INTRODUCTION	1
1.0 Background of Project	1
1.1 Problem Description	2
1.2 Objectives of Project	2
1.3 Scope of Project	3
1.4 Significance of Project	3
1.5 Conclusion	4
CHAPTER 2 : LITERATURE REVIEW	5
2.0 Introduction	5
2.1 The history of email	6
2.3 Modern internet email	7
2.3 Email and event management	8
2.4 Information Extraction: Introduction	11
2.5 Information Extraction: technique and Algorithm	12
2.5.0 Syntactic Rule	14
2.5.1 Natural Language Processing and Information Extraction	16
2.6 Conclusion	20
CHAPTER 3 : METHODOLOGY	21
3.0 Introduction	21
3.1 Knowledge Acquisition	23

CHAPTER 1

INTRODUCTION

1.0 Background of Project

Electronic mail or email is a store is a store and forward method of composing, sending, storing, and receiving messages over electronic communication systems (wikipedia.com). This email can be applied both to the internet email system base on the Simple Mail Transfer Protocol (SMTP) or within intranet which allowed users in one organization to email each other. Since it is fast and easy, it has become a new trend of communication and it will totally replace the traditional mail (which using paper, envelope, and stamp and take a few days to arrive at the receiver) in just few years more. Email provider also increase tremendously and each of them competing each other to provide the best service for their users.

Email has evolved from mere to a way to organizing workflow and managing task communication tool (Corston-Oliver et al., 2004). Nowadays, a lot of feature has been added by email provider to the email to make it more beneficial for the users. Some of the common feature is spam filtering, encryption, calendar, email alert and lot more. Since the feature of the email has been added tremendously, the researcher has taken the opportunity to add another new feature to the email which can be called as "Automated event extraction". The function of this feature is to extract the appropriate information about event announcement from incoming email and store it in the event calendar automatically.