

# Automatic Cover Letter Generator System from CVs (ACLGS)

Hasan Al Shalabi<sup>1</sup>

<sup>1</sup> College of Information Technology Al-Hussein Bin Talal University

*Received: 7 December 2012 Accepted: 5 January 2013 Published: 15 January 2013*

---

## Abstract

The proposed system comes to overcome the problem of writing a C.V. Cover letter which requires some linguistic skills and a lot of experience in this domain in addition to its cost in term of time and money. The ACLGS solved the problem by developing an auto generated cover letter based on the user C.V. regardless its format. The ACLGS takes the user C.V. and the carrier announcement that contains the job requirements and the skills needed as input. The system solved the problem by building a template as a frame of slots each slot contains a required skill for the job; the system extracted the required information from the user CV and fills the slots in an automatic fashion. The ACLGS applies the Information retrieval methodologies to extract information with intelligence trends to mine the user C.V. in terms of part of speech tags and some of indicator words that the system used to recognize the proper data and required information. In addition, the system specifies a set of features for each slot in the form. The user C.V. clustered into a number of categories (e.g. Personal information, Qualifications, Experience, Skill, Rewords, and Publications). These categories are used as additional features for the extracted information and data. The system took into account the problem of sentence coherence and improves the output document through using pre-specified sentences that inserted into the output document based on the extracted information discovered from the user C.V.

---

**Index terms**— cover letter, curriculum vitae, information retrieval, information extraction.

## 1 Introduction

The cover letter is a letter usually attached to the applicant's CV to summarize the information related to that particular job. It reflects the applicant's personality in a positive way and includes basic information about his/her expertise and qualifications. It should reflect his/her enthusiasm and competency for the job. The content of the letter should be complementary to the CV, translation and adaptation-oriented information, biography realism in addition to the personal touch. A well-constructed letter often motivates the reader to go through the entire content of the CV. Yet, such a well-organized letter requires significant time and effort to have it in an acceptable shape.

A typical CV does not allow for prolonged and detailed sentences or paragraphs. While, on the other hand, a cover letter could be employed to deliver Author : College of Information Technology Al-Hussein Bin Talal University.

detailed and specific information signifying the applicant's capability and interest about the issue for which the letter has been written.

Rule-based information extraction is a twostage process: learning rules and application rules for target information. Information extraction rules are mainly used to indicate the target information and the context constrained environment, such as CIRCUS [??]. extraction rules of the system concept nodes, each concept node specified rules trigger words, activation conditions, hard constraints, soft constraints and the position of the target information. The trigger word is used to indicate that the target information context must contain

keywords, language patterns of activation conditions specified must meet rigid constraint is mandatory semantic constraints, soft constraints is a semantic restrictions, but this restriction is violated. Concept node later AutoSlog [1], CRYSTAL [3].

LIEP [5], PALKKA [2], RAPIER [6] and other extraction rules of the system have a similar end. Shows that as long as the text to meet the rules specify constraints, namely to achieve the purpose of information extraction. Therefore, the learning of the rule itself and extracting key information, information extraction is relegated to a secondary process. Rules epitomize the fusion of domain knowledge and linguistic knowledge; build process of the knowledge acquisition process. According to the manual involvement of the different, the building is divided into three types: the manual preparation of knowledge, knowledge of the semi-automatic acquisition and knowledge rules automatically obtain.

The proposed system takes into consideration many parameters to improve the results in additional to the applicant C.V. the system based its results on the institutes announcement and the job position. The new system gave different results with different sentences which make the output dynamic and not limited to a single template as other research papers. The ACLGS follows the Information retrieval methodologies to extract information with intelligence trends to mine the user C.V. in terms of part of speech tags and some of indicator words that the system used to recognize the proper data and required information II.

## 2 Proposed system

The ACLGS is a new approach of creating cover letter based on processing two documents: the user Two types of cover letters the system serves, one for a faculty position and the other for postdoctorate degree.

A classifier used to identify the job title using the announcement and assign a class for it. Based on that class, the system selects the best template for the cover letter. The classifier builds its decision based on a set of keywords that identifies the appropriate class. We use the CTS Tagger [8] subsystem to identify the part of speech tag (P.O.S). The P.O.S is a significant feature that the system used for information extraction in additional to other features.

The algorithm starts by pre-processing the input documents as a required step in order to get good results. This step partitions the C.V. into many segments as in the algorithm (1) below.

## 3 Proposed System Algorithm (1):

The segments of the C.V. are (Personal information, Qualifications, Experience, Membership, Publications, Supervision, Awards and Patents). We know that there is no unified C.V. template but the system identifies these parts based on a set of features .Table (1) lists all the subjects that will be searched in the C.V. and the synonyms that may be written.

## 4 Global Journal of Computer Science and Technology

Volume XIII Issue III Version I Two more steps implemented in the preprocessing step are the tokenization and word tagging. Based on the classifier the right template will be selected. The template contains many slots with identified features that will be filled by the system as in figure ??? for Post-Doctorate. The following table (2) displays the rules and the features used to extract the required information in order to place it in the blanks (Slots). A set of P.O.S patterns was extracted by examining the C.V. These patterns used to be one of the features that help in extracting the required items from the C.V. Where the number of the C.Vs used in the dataset is about 100 document form the field of Academic Faculty members especially in the domain of Information technology. One more feature adapted to extract the required information which defines the set of keywords that are the indicator of the existing of important words in the C.V. these keywords called as indicator words as shown in table (2). The indicator words are frequently written before the required information that the system tries to extract.

The algorithm takes into account the calculation of the user (Faculty member) experience years. In some C.Vs the user didn't write the total experience years so the algorithm extract that value by accumulating the years of experience. The algorithm starts by calculating the period of each job especially that the users wrote the experience of each job in the C.V. So we find the period of the experiment by subtracting the second value from the first one, and finally we accumulate all these periods to give the total number experiment years.

The algorithm takes into consideration the information exists in the carrier announcement document that much the user information and used as a feature to be searched in the user C.V. One of the data that the algorithm looks for is the University or College and department name to be inserted in the beginning of the Created Cover Letter and the job title that can be extracted by the set of features that described in table (2) above.

The system provides a set of sentences for each paragraph in the cover letter. These sentences clustered into three categories for the three paragraphs that cover letter consists of. The system selected randomly by the system in order to make results vary as much as possible as in table (3).

---

## 5 First Paragraph Sentences

I am interested in a (type of work) position in your (company, agency). I believe that my interest, experience and education support my ability to learn and produce in this area. I am interested in applying for a (teaching position, opportunity in your school district). I will be/am certified to teach (subject or grades).

## 6 Second Paragraph Sentences

My educational background, experience in this area, and my sincere interest in the challenges offered support my belief that I have the qualifications you seek.

During the past four years of college, I have developed through education and experience a strong desire to find an entry level opportunity in (work area). I feel that I am equipped with educational preparation and valuable experience that supports my qualification for a career in \_\_\_\_\_.

## 7 Conclusions and Future Works

The need of cover letters, the difficulties that the applicants faced and the cost of writing the cover letter by experts motivated us to design a system to auto generate the cover letter. ACLGS takes into consideration many parameters to improve the results in addition to the applicant C.V. the system based its results on the institutes announcement and the job position. The new system gave different results with different sentences which make the output dynamic and not limited to a single template as other research papers.

For future work, to improve our proposed system in order to get more valuable and accurate outputs by adding more sentence database to generate completely different output. And implementing the research on Arabic language. <sup>1</sup>



Figure 1: T © 2013

1

2  
D D D D )  
(

Figure 2: Table 1 :

---

<sup>1</sup>© 2013 Global Journals Inc. (US) Global Journal of Computer Science and Technology

2

013  
2  
Year  
C

Figure 3: Table 2 :

A position with your <institute> would provide the kind of opportunity and challenge I seek.

## .1 Third Paragraph Sentences

Enclosed is a resume describing my employment and educational background for your consideration.

Enclosed is a resume describing my education and employment background in support of my qualifications for your staff opportunity.

If you will review the enclosed resume you will see that I have had a strong education and varied experience which is compatible with (supportive of) the requirements of this position. Table ?? : Sample Set of Stored sentences [9,4] In the post-processing step we try to give better results and put the final cover letter in different formats and content, the algorithm adds some sentences that depend on the C.V. and the announcement. Each user has different skills and may have different highlights in his C.V. according to that the algorithm will select a suitable sentence from the database to fit in.

## .2 III.

## .3 Results

The following is an example of a cover letter generated by the system for a faculty member applicant as in figure (4).

[Howard] *A Dozen Sentences That Should Appear In Your (Academic) Job Application Letter*, Philip N Howard . <http://www.grad.washington.edu/mentoring/memos/dozen-sentences.shtml>.Date:11.1.2013

[Kim and Moldovan] 'Acquisition of linguistic patterns for knowledge-based information <'. J Kim , D Moldovan . telephone: + 962 799889571. *Address: P.O. Box 20*.

[at Technology University in< department of computer science>. My thesis was entitled < Automatic Keyword Extraction Using 'at Technology University in< department of computer science>. My thesis was entitled < Automatic Keyword Extraction Using Combined Methods >. Samples of My publication are as follows: < Data Gathering for Periodic Sensor Applications, Text Summarization Extraction System (TSES), developing a Virtual Laboratory for a Communication and Computer Networking Course > Enclosed is a resume describing my employment and educational background for, I would appreciate an opportunity to discuss my qualifications in an interview at your convenience. I look forward to hearing from you. Yours sincerely, <Rafeeq Al-Hashemi> extraction'. *IEEE Transactions on Knowledge and Data Engineering* Nov-2006. 1995. 10 (5) p. . (I was awarded my Ph.D. by the)

[Riloff ()] 'Automatically constructing a dictionary for information extraction tasks'. E Riloff . *Proceedings of the 11th National Conference of Artificial Intelligence (AAAI-93)*, (the 11th National Conference of Artificial Intelligence (AAAI-93)) 1993. AAAI Press / The MIT Press. p. .

[CST's Part-Of-Speech tagger] [http://cst.dk/online/pos\\_tagger/uk/](http://cst.dk/online/pos_tagger/uk/).Date:2/12/2012 *CST's Part-Of-Speech tagger*, Brill. (with adaptations)

[Slonnenger and Kurtz ()] *Formal Syntax and Semantics of Programming Language*, K Slonnenger , B Kurtz . 1995. Addison-Wesley Publishing Company.

[Huffman ()] *Learning information extraction patterns from examples. Connections, statistical, and symbolic Approaches to Learning for Natural Language Processing*, S Huffman . 1996. Spingler-Verlag.

[Soderland ()] 'Learning information extraction rules for semi-structured and free text'. S Soderland . *Journal of Machine Learning*, 1999. 34 p. .

[My research and teaching interests fit extremely well with the requirements of this post and with existing members of staff. I have  
'My research and teaching interests fit extremely well with the requirements of this post and with existing members of staff. I have extensive teaching experience in the <department of computer science> <at Al-Hussein Bin Talal University>, most of it focused on < Artificial Intelligence, Distributed Database, Computation Theory, Computer Technology, Network Security, Image Processing, Genetic Algorithm, Software Project Management, Object Oriented Programming, Logic Design. I have taught programming languages such as C++, Java, PROLOG, and Visual Basic. Net.>. My work provides a useful link between < Artificial Intelligence, Data Mining > in the department, encouraging research and teaching collaborations. I have more than 4 years of Experience in administration (as Dean'. *experience in this area, and my sincere interest in the challenges offered support my belief that I have the qualifications you seek*, (I am on several committees, Reviewing Activities)

[Sample Sentences to Use When Writing A Cover Letter] *Sample Sentences to Use When Writing A Cover Letter*, <http://www.cazenovia.edu/default.aspx?tabid=1426>.Date:11.1.2013

[Lehnert et al. ()] 'UMass/Hughes: Description of the CIRCUS System Used for MUC-5'. W Lehnert , J Mccarthy , S Soderland , E Riloff , C Cardie , J Peterson , F Feng , C Dolan , S Goldman . *MUC*, 1993.