Analysis of Knowledge and Skills Needed for Software Architects

Gyul Hee Kim, Min A Kim and Yeong Kwang Han*
Keimyung University
1095 Dalgubeol-daero, Dalseo-Gu, Daegu 42601, Republic of Korea
[e-mail: sundayhiphop@daum.net]
*Corresponding author: Yeong Kwang Han

Abstract

As the impact of IT on business grows, businesses that use IT are becoming more important than ever. Although many companies seek software architects at various business sites, little agreement has been made on definition of a software architect. The purpose of this research is to explore the definition of a software architect by examining how it is accepted by these industries and businesses by way of identifying what skills and knowledge are required for the job. A content analysis was conducted on 6,863 job advertisements. Through the window of job postings collected, this study attempted to look into knowledge and skills needed for software architects in eleven countries. As a result, software architects were shown to be highly experienced professionals with advanced degrees. The main areas of profession were software architecture, development methodology, and problem solving.

Keywords: Software architects, job advertisement, content analysis, skills

1. Introduction

A software architect is a responsible architect and top developer for system development based on software technology, creating a software architecture that combines technology, business, and social influences. As organizations start to realize the importance of developing software architecture, they are beginning to pay attention to the human resources which possess the proper skills and knowledge to enable these goals [1]. As a result, the demand for software architects will increase as the senior software developer. Businesses will continue to want a software architect who understands and connects the characteristics of job skills required for specific positions, including software engineers, programmers, and managers.

However, being a relatively new job title, a software architect is very vague in understanding what exactly it is and what it does. We attempt to examine how software architects are accepted by industries and businesses as well as identify what skills and knowledge are demanded. In order to understand the reality of the knowledge and skills that organizations require, this research collects and analyzes job advertisements posted on major job seeking web sites.

Most previous studies have been conducted by using small datasets collected from various sources such as job advertisements published on offline newspapers. However, nowadays many job advertisements are posted on the websites. Since they are abundant, it would be better to use them for job requirement analysis.

2. Data

We collected job ads as suggested by Kim and his colleagues [2]. Job advertisements have been collected using a customized web crawling

system, called WebMon, for job post collection. A set of webpages that contain hyperlinks for job advertisements are registered for crawling and WebMon periodically revisits them every one hour and identifies the newly posted job advertisements. Job advertisements collected from the websites are formatted with HTML and thus it is necessary to extract pure text content from them after removing HTML tags, comments and other redundant contents. Generally job advertisements are repeatedly posted on the same webpage in order to attract candidate applicants by displaying them in the first page. Therefore, it is necessary to remove duplicate job from the collected advertisements before conducting text analysis.

In this study, a program developed based on the Java language was used to extract data necessary for data analysis in the online job advertisement text collected by the web crawler. The text read to analyze the job advertisement was converted to lower case with touppercase and divided into individual vocabularies using split. Since some words are meaningful when multiple words are joined in series, we have implemented an n-gram algorithm to generate a series of consecutive words to extract such data. Frequently appearing words(a, the, etc.) among the generated words or word combinations and words identified as meaningless information were removed using an if statement.

We selected a total of 12 countries that posted architects job advertisements in English through well-known job advertisement websites. The countries were grouped into six groups based on geographical and economic proximity. We collected the job advertisements for about 5 months from April 1, 2015 to October 31, 2015 with some interruptions which caused by technical problems. A total of 206,585 advertisments was collected from these countries. Some of job advertisements are duplicated since the advertisers post their advertisement repeatedly. Therefore, we removed duplicated job advertisements from the collected data set. As a result, only 6,863 survived the extensive filtration process. The entire data set was further filtered to allow only ads which specifically contain the term “architects” in its job title.

A summary of the collection results can be seen in **Table 1**. No data from USA and Vietnam was included

Table 1. Summary of data collection

Area	Country	Collected	Filtered
Asia	Malaysia,	41,748	569
	Singapore,		
	Indonesia,		
	Thailand		
	Hong Kong		
	Vietnam		
Australia	Australia	8,703	446
Europe	England	65,498	1605
	Ireland		
Gulf	Gulf	7,555	176
India	India	49,987	3,185
Canada	Canada	30,117	882
Total		206,585	6,863

3. Results

3.1 Knowledge and Skills Analysis

We performed content analysis that was developed by Kim and Lee [3]. The overall distribution of the skills and knowledge categories for software architects are summarized in **Fig. 1**. With the exception of hardware, seven of the eight categories were mentioned at least once in 80% or more of the job advertisement. As shown in Figure 1, the software architect must be exceedingly well rounded. Problem Solving, Management, Development Methodology, Business, Software, and Network/Architecture categories were exceptionally high with more than 90% of the job advertisements mentioning related skills. The overall top 10 skills which were mentioned by more than 80% of the collected job advertisements are listed in **Table 2** to elicit the popular skill traits of a software architect.

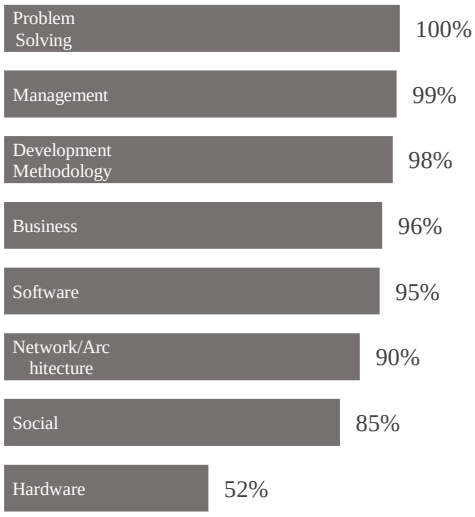

Fig. 1. Software Architect Skill Distribution

Table 2. Top 10 Skills

Top 10 Skills	Num. of Ads	% of Ads
Quantitative	6,863	100%
Knowledge of Methodologies	6,841	100%
General Knowledge of Software	5,758	84%
General Knowledge of Development	5,749	84%
Organizing	5,652	82%
Design	5,399	79%
Leadership	5,300	77%
Interpersonal	5,130	75%
Industry Specific	5,022	73%
General Management	5,022	73%

4. Software Architect Skills

From the analysis results, we can see that software architecture, development methodology, and problem solving are the top required and distinguished skills of a software architect. Although software skills within the technical skills category had the highest percentage after systems skills, business skills on average was higher than technical skills. Management skills was the highest category under business skills with a 99% percentage. This reflects the need for software architects to be comfortable working with others such as coworkers or team members, as well as communicating well with decision makers. As for the business subcategory, it is interesting to see that industry specific(domain) knowledge is the highest in rank. General knowledge of business also showed a response of over 60% adding to the notion that software architects must be well rounded. Software architects need to be well rounded with not only expert hard skills, but a high importance in soft skills as well. The highest ranking of soft skills for Software architect is interpersonal skills. Software architect must be able to work closely with company executives to provide insight on a strategic level. Both team and executive relationships are supported by the Software architect's ability to communicate well and understand business. They must be highly competent in communicating with people outside their field, or the non-experts, in a clear and effective manner while addressing solutions to the problems faced by the company.

5. Conclusions

5.1 Implications

The implications of this study can be presented in a two-fold. Firstly, this research is one of the first study to investigate the newly emerging job role of a software architect from an IT researcher perspective. Secondly, the research was implemented by developing a software tool for automatically collecting data from the top US job sites. This can be considered a substantial advancement compared to the numerous past studies which depended on manual data collection and preparation. By developing a flexible software tool. it will not only be able to establish an easily deployable framework for consistent and continuous future research on software architects. But it can also be used on any other job categories as well.

5.2 The challenge

In this research, we reported job requirement analysis results using online job advertisements. Also it would be interesting if we analyze the

contents of job advertisements. However, this research does not consider content analysis of job advertisements. This will be our future work.

References

[1] L. Bass, "*Software architecture in practice*," Pearson Education India, 2007.

[2] Y. S. Kim, C. K. Lee, J. H. Lee, E. H. Kim, and S. J. Shim, "An Empirical Analysis of IT Job Requirements Using Online Job Postings", *Proceedings of the 4th International Conference on Smart Media and Applications*, pp. 36-39, 2016.

[3] J. Kim, C. K. Lee, "An Empirical Analysis of Requirements for Data Scientists Using Online Job Postings," *International Journal of Software Engineering and Its Applications*. vol. 10, no. 4, pp. 161-172. 2016.