

# DETERMINING CRITICAL SKILLS AND KNOWLEDGE REQUIREMENTS OF IT PROFESSIONALS BY ANALYZING KEYWORDS IN JOB POSTINGS

Paul J. Kovacs, Robert Morris University, kovacs@rmu.edu  
Gary A. Davis, Robert Morris University, davis@rmu.edu

---

## ABSTRACT

*In a constantly changing job market, Information Technology (IT) practitioners must keep their skills up-to-date. Educational institutions with Informational Technology curricula must also keep current in course offerings and programs of study. One way to maintain relevance is to examine and identify what IT skills and competencies are in demand. Job positions, indicating current expertise needed by companies, are good indicators of the types of IT skills that current and future IT professionals should possess in order to be marketable in a competitive labor market. This paper analyzes keyword data from job postings of a regional trade association to determine what skill sets and competencies are essential.*

**Keywords:** IT Skills requirements, IT Workforce skills requirements, University Corporate Relationships, IT Curriculum Development

## INTRODUCTION

In a constantly changing job market, Information Technology (IT) practitioners must keep their skills up-to-date. Educational institutions with Informational Technology curricula must also keep current in course offerings and programs of study. One way to maintain relevance is to examine and identify what IT skills and competencies are in demand. Job positions, indicating current expertise needed by companies, are good indicators of the types of IT skills that current and future IT professionals should possess in order to be marketable in a competitive labor market.

This study analyzes keywords from Internet job postings of a regional trade association. The objectives of the study were to determine the following:

1. What IT Skill Categories have the largest proportion of total keywords?
2. Out of the top IT Skill Categories, which keywords have the largest proportion?
3. What are the Top Ten Overall Keywords (regardless of IT Skill Category)?

The study results should be useful to educational institutions who want to remain competitive in their curriculum design and IT course offerings. The results should also be beneficial to IT practitioners who seek to advance their skills and further their professional development. Finally, students will find this study helpful in the identification of essential skill sets that are useful in the selection of IT tracks and elective courses.

## RELATED RESEARCH

Previous studies have been undertaken involving the IT workforce and job skills. The survey method was used in a number of studies that targeted IT managers, recruiters, and professionals to determine essential and non essential IT skills and competences. Other survey research was directed at IT faculty and students in higher education to determine their assessment of IS/IT skills in academic programs [1, 2, 3, 4, 6, 8, 12, 13, 15].

Two prior research projects regarding skills and knowledge requirements of IT professionals using on-line job postings and keywords are closely related to this research. Koong, Liu, and Liu [7] selected two Internet job databases Monster.com and HotJob.com as data sources to examine the types of expertise needed in new employees. Their data set consisted of 300 observations conducted over ten weeks from November 2000 to February 2001. 150 jobs were extracted from each database and classified into one of the following five categories: Programming Languages, Web Development, Database, Operating Systems and Networking. Two other variables, Communication Skills and Experience Requirements, were also collected and tallied.

305 job skills were identified in the 150 jobs extracted from Monster.com. 292 job skills were found in the other 150 jobs posted in HotJobs.com. Collectively, the 300 jobs showed a total of 597 job skills. Out of the 300 jobs, 27% fit into the Programming Languages category and 24% into Web Development. These two categories alone accounted for about 51% of the total number of job skills identified and indicated a majority of the jobs posted require excellent coding skills involving such

contemporary languages as C/C++, Java and Visual Basic, as well as SQL and DHTML/HTML

Within the Programming Language Category, about 95% of the job skills were on C/C++, Java and Visual Basic. Six types of Web Development skills were SQL, DHTML/HTML, ASP, JavaScript, VBScript, and COM. SQL and DHTML/HTML were the two dominant job skills in this category.

The five types of Database skills extracted were Oracle, SQL Server, Access, JDBC, and DB2. Oracle was the dominant skill, comprising half of the job skills; two Microsoft products accounted for 39%. DB2 skills were indicated only seven times by recruiters.

In the Operating Systems category, there was a demand for only two types of operating systems, Unix and Windows (95/98/2000). There was not a single job listing for any of the mainframe-based operating systems. Similarly, the Networking category indicated a need for only two types of skills. About 81% of the Networking-related descriptions were for Windows skills. The other 19% required some type of WAN/LAN expertise. These authors also classified Communications Skills and found that only 27% of the jobs required both oral and written communication skills.

A more recent study reviewed the current IT job market based on advertising on the Internet. In the study, data were collected from Monster.com and verified against the Career Builder website over a three year period from April 2002 to April 2005 [11]. Of the 42 categories provided by Monster.com, the IT jobs were selected from five categories the Information Technology, Computer Hardware, Computer Software, Internet/E-commerce, and Telecommunications. The skill-sets used for the search were selected from 39 technical skills and resulted in the following 14 skill categories: Web Programming, Java, C/C+, Unix, Windows, Oracle, SQL Programming, SQL Server, Visual Basic, .NET Development, ERP, Linux, E-commerce Servers, and Certification. More specifically, the Web Development Skill category included keywords such as Perl, JavaScript, ASP, HTML, DHTML, XML, VBScript, PHP, VB.NET, and Cold Fusion.

The results in the aforementioned study indicate that the need for development work in Web-related applications has changed the demand for these programming skills. These skills were mentioned in 42.6% of job advertisements. Specifically, Java remained important and was required in more than one-fifth of all jobs. However, Java has not replaced

the need for C++ and C# programmers. SQL programming was also required in more than one-fifth of the job openings. In addition, the demand for Microsoft's .NET skills has grown in importance and now represents 13% of skills required by employers.

In the Operating Systems category, the demand for Unix and Windows held steady to Linux. However, over a two year period, the demand for Linux rose from five percent to 13.6% of all advertised job postings. Within the database category, Oracle was still the most demanded skill with over 22% of positions requiring it. SQL server skills grew in demand and represented 16% of all IT jobs in 2005. Additionally, an average of five percent of all positions advertised required industry or vendor certifications.

### **DATA SOURCE**

The data source for the current study was the Pittsburgh Technology Council (PTC). The PTC is a regional trade association that includes an Information Technology cluster. The IT cluster includes participants from more than 750 companies, ranging in size from large corporations to small entrepreneurial organizations. The 750 member companies provide services or products focused on software, hardware, telecommunications and professional services.

The Career Center of the PTC provides an on-line service where companies can post positions and search a resume database. In turn, job seekers from all over the world can post resumes and respond to opportunities in the Pittsburgh region. Additionally, the PTC provides services for companies offering internships to undergraduates and graduate students in Southwestern Pennsylvania. The data for this study were provided by the Educational Network of the PTC. The authors obtained spreadsheet data containing keywords from the Career Center's Job Posting and Resumes for October 2006.

### **RESEARCH METHODS**

In order to answer the research questions, the authors used the spreadsheet keywords extracted from the Job Postings. The original set of data included 153 individual keywords. Some of the keywords directly related to a specific IT skill or content area such as C#, Linux, or HTML. Conversely, some of the keywords such as AS400, Intranet, or Frame Relay did not directly relate to a specific IT skill. The keywords were analyzed to guarantee that no duplication or elimination of data existed. In certain

instances, groupings of related keywords (e.g., Oracle, Oracle 9i, and Oracle Enterprise Manager) were consolidated into the category “Oracle”). After consolidation, 99 unique keywords resulted from the original 153. The 99 keywords were then classified into 14 Categories. Guidelines to develop these 14 Categories, were the result of existing job category classifications from ABET (Accreditation Board of Engineering and Technology), AIS (Association for Information Systems, and ACM (Association for Computing Machinery). The keywords that did not directly relate to an IT skill were grouped in a category called *Other*.

**DATA ANALYSIS**

After consolidating the 99 unique IT keywords into 14 categories, the total number of occurrences for each keyword was calculated. The counts of the unique keywords, as well as the percentage of the category total, are displayed in Table 1: IT Keywords Grouped by Category.

As Table 1 shows, keywords in the *Web Development* category (n = 1073) comprised the largest percentage (19.3%) of the total for all IT keyword categories. The category with the second highest number of IT keywords was *Programming Languages* (n = 1025), which made up 18.4% of the total. Finally, the category *Database* had the third highest number of IT keyword occurrences (n = 975), which amounted to 17.5% of all categorized keywords. All 14 categories of IT keywords, along with the corresponding counts and percentage of total are displayed in Table 1.

**Table 1.** IT Keywords Grouped by Category

<i>Category</i>	<i>Count</i>	<i>Percentage</i>
Web Development	1073	19.3
Programming Languages	1025	18.4
Database	975	17.5
Operating Systems	545	9.8
Communication Skills	497	8.9
Networking	419	7.5
Security	275	4.9
Project Management	267	4.8
Business Intelligence	149	2.7
Other	118	2.1
Certification	64	1.2
Systems Analysis & Design	62	1.1
Vendor Software	53	1.0
Multimedia/Graphics	43	0.8
Totals	5565	100.0

To further analyze the IT keywords in the present study, the top three categories (in terms of keyword occurrence) were examined in more detail. Specifically, the count of IT keyword occurrences was computed for each of the following categories: *Web Development*, *Programming Languages*, and *Database*.

As displayed in Table 2: Web Development Keywords, the keyword *HTML* (n = 173) made up the highest percentage (16.1%) of the *Web Development* category. The keyword *XML* was the second most occurring keyword (n = 165), comprising 15.4% of the total for the *Web Development* category. Finally, the keyword *ASP* was the third most prevalent IT keyword (n = 133), constituting 12.4% of the *Web Development* category. All 21 unique IT keywords in the *Web Development* category are listed in Table 2, along with the corresponding counts and the percentage of total.

**Table 2.** Web Development Keywords

<i>Keyword</i>	<i>Count</i>	<i>Percentage</i>
HTML	173	16.1
XML	165	15.4
ASP	133	12.4
JavaScript	79	7.4
J2EE	78	7.3
PERL	78	7.3
Shell Scripting	47	4.4
JSP	46	4.3
CSS	43	4.0
Websphere	41	3.8
PHP	39	3.6
Weblogic	30	2.8
XSLT	27	2.5
DHTML	21	2.0
Cold Fusion	18	1.7
Ajax	14	1.3
VBScript	13	1.2
Java Servlets	11	1.0
Java Server Faces (JSF)	7	0.6
Java Swing	6	0.5
WSAD	4	0.4
Totals	1073	100.0

The category *Programming Languages* had the second highest count of IT keywords (n = 1025). The eight unique IT keywords in the *Programming Languages* category are displayed in Table 3: Programming Languages Keywords. As shown in Table 3, the keyword *Java* (n = 348) represented the highest percentage (34.0%) of the *Programming Languages* category. The keyword *Visual Studio.net*

was the second most occurring keyword (n = 255), representing 24.9% of the total for the *Programming Languages* category. The third most popular IT keyword, C++ (n = 165), made up 16.1% of the *Programming Languages* category. All eight unique IT keywords in the *Programming Languages* category are displayed in Table 3, along with the associated counts and the percentage of total.

**Table 3.** Programming Languages Keywords

<i>Keyword</i>	<i>Count</i>	<i>Percentage</i>
Java	348	34.0
Visual Studio.NET	255	24.9
C++	165	16.1
Visual BASIC	151	14.7
C#	101	9.9
ActiveX	4	0.4
Lisp	1	0.0
Delphi	0	0.0
Totals	1025	100.0

The category with the third highest occurrence of IT keywords was *Database*. The four unique keywords within the *Database* category are shown in Table 4: Database Keywords. As noted in Table 4, *SQL* (n = 614) comprised the highest percentage (63.0%) of all IT keywords in the *Database* category. The keyword *Oracle* was the second highest occurring IT keyword (n = 305) in the *Database* category, at 31.3% of the category total. More specifically, the IT keywords *SQL* and *Oracle* combined, accounted for 94.3% of all IT keywords in the category *Database*. In contrast, the IT keywords *DB2* and *mySQL* (n = 31, n = 25) comprised a very small percentage of the *Database* category total, with 3.2% and 2.5%, respectively.

**Table 4.** Database Keywords

<i>Database Keyword</i>	<i>Count</i>	<i>Percentage</i>
SQL	614	63.0
Oracle	305	31.3
DB2	31	3.2
mySQL	25	2.5
Totals	975	100.0

In an effort to analyze overall keywords (regardless of categorization), the top ten keywords from the current study are listed in Table 5: Top Ten Overall Keywords. As shown in Table 5, the keywords *SQL*, *Communication Skills*, and *Java* were the top three (respectively) in terms of keyword occurrence. The top ten IT keywords, as well as their corresponding

category, count and proportion of total are shown in Table 5.

**Table 5.** Top Ten Overall Keywords

<i>Keyword</i>	<i>Category</i>	<i>Count</i>	<i>%</i>
SQL	Database	614	11.0
Communications Skills	Communication Skills	497	8.9
Java	Programming Languages	348	6.3
Oracle	Database	305	5.5
Project Management	Project Management	267	4.8
Visual Studio.Net	Programming Languages	255	4.6
Unix	Operating Systems	223	4.0
Security	Security	212	3.8
HTML	Web Development	173	3.1
C++	Programming Languages	165	3.0
Other	N/A	3059	45.0
Totals		5568	100.0

## CONCLUSIONS

The present research sought to answer the following three questions: 1) “What IT Skill Categories have the largest proportion of total keywords?” 2) “Out of the top IT Skill Categories, which keywords have the largest proportion?”; and 3) “What are the Top Ten Overall Keywords?” Regarding Question 1, Table 1 shows that *Web Development* was the category that contained the largest proportion of IT keywords. In addition, the categories *Programming Languages* and *Database* contained the second and third largest proportion (respectively) of IT keywords. These findings are similar to the findings of Koong, Liu, and Liu [7] who found that *Programming Languages*, *Web Development*, and *Database* (respectively) had the highest occurrence of keywords. Also, Prabhakar, Litecky, and Arnett [11] found that web and web-related application skills were the skill-keywords in highest demand.

Regarding Research Question 2, the current research took a more detailed look at the *Web Development*, *Programming Languages*, and *Database* categories. Within the *Web Development* category, *HTML*, *XML*, and *ASP* were the top three IT skills, in terms of proportion. Within the *Programming Languages* category, *Java*, *Visual Studio.net*, and *C++* were found to be the top three languages, in terms of largest proportion of keywords.

Within the *Database* category, *SQL* and *Oracle* were the two dominant keywords. In contrast, the demand for *DB2* and *mySQL* database skills was found to be very minimal. While the present study found *SQL* to be the highest occurring keyword in the *Database* category, it should be noted that the more-specific keyword, "MS-SQL Server" was not included in the research data.

Regarding Research Question 3, the current study found a very high demand for the keywords *Communication Skills*, *Project Management*, and *Security*. More specifically, the keyword *Communication Skills* was found to be the second most sought after skill (n = 497) in the entire study. Previous research has identified soft skills (e.g., communication and team building) to be as vital, if not more vital, than technical skills in an IS/IT position [2, 3, 16]. Additionally, the current research indicates that Project Management and Information Security skills are in high demand among IS/IT positions in the 21<sup>st</sup> century.

In addition to the proposed elements of this study, certain ancillary findings resulted from the analysis of the research data. Perhaps the most surprising was the omission of COBOL as a keyword. This was unanticipated in light of the rising number of requests made to the authors' University to fill full-time and part-time/internship positions involving COBOL programming. However, the omission of COBOL may not be unexpected, considering that other mainframe-and IBM related keywords (e.g., JCL, CICS, MVS and DB2) had a very low occurrence in the research data.

Another interesting finding in the current research was the omission of any keywords related to the Microsoft Office suite or other productivity software. Ironically, Lotus and Domino were referenced in the research data; although, both were associated with small levels of demand.

A final interesting discovery related to the *Multimedia/Graphics* category. This category had the lowest proportion of keywords within the 14 categories examined in the study. Keywords such as *Flash* and *Adobe* had a very low level of occurrence. However, this finding may not be unusual, since the majority of IS/IT positions do not require multimedia/graphics skills.

Many recommendations for University curricula can be made from the current findings of this paper. Specifically, liaisons should be established and encouraged between industry and academia. Faculty

and student membership in IT/IS professional associations should be advocated. Finally, academia should promote student internships that more closely match student skills to corporate needs.

The current study has generated the need for further areas of research. In particular, surveys should be devised and administered to students to determine which mix of enrolled skills match with industry norms. In addition, follow-up studies concerning IT/IS skills should be undertaken on a timely basis to detect advancing technological directions and developments. Finally, corporate survey data concerning IT skills and competencies should be made available to all University students.

## REFERENCES

1. Athey, S., & Plotnicki, J. (1998). The evaluation of job opportunities for IT professionals. *Journal of Computer Information Systems, Spring*, 71-88.
2. Bailey, J., & Mitchell, R. B. (2007). Industry perceptions of the competencies needed by computer programmers: technical, business, and soft skills. *Journal of Computer Information Systems, (Winter)*, 28-33.
3. Cappel, J. J. (2001). Entry-level IS job skills: A survey of employers. *Journal of Computer Information Systems, 42(2)*, 76-83.
4. Davis, G. A., Caputo, D., Kovacs, P., & Turchek, J. (2005). Identifying competencies for the IT workforce: a quantitative study. *Issues in Information Systems Journal, 6(1)*, 339-345.
5. IS2002, Model Curriculum and Guidelines for Undergraduate Degree Programs in Information Systems, 2002, curriculum effort of Association for Computing Machinery (ACM), Association for Information Systems (AIS), Association of Information Technology Professionals (AITP)
6. Jiang, J., Udeh, I., & Hayajneh, A. (1994). Employers' expectations of incoming graduates: from recruiters' views. *Journal of Computer Information Systems, (Spring)*, 57-59.
7. Koong, K. S., et al. (2002). A study of the demand for information technology professionals in selected Internet job portals. *Journal of Information Systems Education, 13(1)*, 21-29.
8. Kovacs, P. J., Caputo, D., Turchek, J., & Davis, G. A. (2006). A survey to define the skill sets of selected information technology professionals. *Issues in Information Systems Journal, 7(1)*, 242-246.
9. Lee, D., Trauth, E., & Farewell, D. (1995). Critical skills and knowledge requirements of IS professionals: a joint academic/industry investigation. *MIS Quarterly, 19(3)*, 313-340.

10. Maier, J. L., Clark, W. J., & Remington, W. J. (1998). A longitudinal study of the management information systems (MIS) job market. *Journal of Computer Information Systems*, 39(1), 37-42.
11. Prabhakar B, Litecky, C, & Arnett K. IT skills in a tough job market, Association for Computing Machinery. *Communications of the ACM*. New York: Oct 2005. Vol. 48, Iss. 10; p. 91
12. Richards, T., Yellen, R., Kappelman, L., & Guynes, S. (1998). Information managers' perceptions of IS job skills. *Journal of Computer Information Systems*, 38(3), 53-57.
13. Tang, H. L., Lee, S., & Koh, S. (2001). Educational gaps as perceived by IS educators. *Journal of Computer Information Systems*, (Winter), 76-84.
14. Todd, P., McKeen, J., & Gallupe, R. B. (1995). The evolution of IS job skills: a content analysis of IS job advertisements from 1970 to 1990. *MIS Quarterly*, 19(1), 1-27.
15. Trauth, E., Farwell, D., & Lee, D. (1993). The IS expectation gap: industry expectations versus academic preparation. *MIS Quarterly*, 17(3), 293-303.
16. Van Slyke, C., Kittner, M., & Cheney, P. (1998). Skill requirements for entry-level information systems graduates. *Journal of Information Systems Education*, (Fall), 6-10.