Desirable ICT Graduate Attributes: Theory vs. Practice

J. Debuse
M. Lawley
Faculty of Business
University of the Sunshine Coast
Maroochydore DC, Queensland 4558, Australia
jdebuse@usc.edu.au mlawley1@usc.edu.au

ABSTRACT

The majority of ICT graduates must begin their careers by successfully fulfilling the requirements advertised within online recruitment sites. Although considerable research into employer requirements is commonly undertaken when preparing curricula, studies investigating how well the graduate attributes on which curricula are based match those required by employers have been limited in terms of the techniques used. This study employs an innovative approach of analyzing online ICT employment advertisements in Australia and the United States to determine the key attributes sought by ICT employers, together with the most commonly required skill groupings. A position-based wrapper system was developed to extract the advertisement data, which was then analyzed using a text mining package. The results are benchmarked against those from standard ICT curricula produced by academic and professional bodies. The findings suggest that employers place greatest emphasis upon experience and technological skills; although current curricula meet these requirements, their emphases warrant revision. There also appear to be differences between professional body curricula and the ISCC '99 curriculum which was produced by industry and academia, with the latter appearing to match employment market demands more closely.

Keywords: Graduate Attributes, ICT Curricula, ICT Employment, Data Mining

1. INTRODUCTION

ICT graduates seeking employment are faced with the challenge of addressing the selection criteria of prospective employers, typically through responding to advertisements placed on the internet. Although this is only the first phase in what can be a lengthy recruitment process, graduates must clearly be successful at this stage if they are to have any chance of securing employment within the industry. Their success requires them to demonstrate satisfaction of selection criteria, which to date have only been examined in a very limited fashion, such as technical skill occurrence frequencies (Liu et al., 2003).

Graduate attributes are influenced by five key stakeholder groups: academics, professional bodies, employers, clients and technology providers. Academics are likely to have the greatest influence, since they develop and teach degree curricula, followed by the professional bodies, whose influence includes accrediting degree programs and sponsoring educational conferences. The contribution of employers is likely to be relatively minor, particularly organizations that are too small to participate in activities such as student sponsorship and collaborative curriculum development. The requirements of ICT clients are likely to affect professional bodies and employers, and thus provide a secondary influence. Finally technology providers such as Microsoft and Sun have considerable impact, since employability is often dependent upon possession of skills with their technologies such as .NET and Java.

The effect of this range of influences is reflected in the significant differences in the skills possessed by ICT graduates compared to those required by employers (Scott et al., 2002). Specifically, business (Milton, 2000) and project management (Kim, Hsu and Stern, 2006) skills appear to be particularly important to employers and of lesser importance to academics. Employers also require skills in the latest technologies, which are hard to incorporate into standard curricula that are strongly biased towards fundamental principles (Lightfoot, 1999).

Research to date has been based upon a number of different approaches (Gallivan, Truex III and Kvasny, 2004): surveys (Crews, 2000; Fang et al., 2004; Gallivan, Truex III and Kvasny, 2004; Kim, Hsu and Stern, 2006; Scott et al., 2002; Wilkins and Noll, 2000), combined focus group, interview and surveys (Gallivan, Truex III and Kvasny, 2004) and job advertisement analysis (Gallivan, Truex III and Kvasny, 2004). The analysis of job advertisements has however been limited to simple analyses of skill occurrence frequencies, or augmentations of this approach (Gallivan, Truex III and Kvasny, 2004; Koong, Liu and Liu, 2002; Petrova and Claxton, 2005).

This study therefore aims to address this gap by analyzing large samples of ICT job advertisements to determine the required applicant attributes and then compare these with curricula produced by academia, professional bodies and industry. These sources are obtained from the United States and Australia to allow international comparisons to be performed. The results give insight into the key attributes sought by employers, together with common skill groupings, and thus may be used by educators to ensure offerings meet business demand. Further, modifications to curricula to reflect these requirements may mitigate the current shortage of ICT student enrolments (Granger et al., 2007).

The remainder of the paper proceeds as follows. A brief background to the study is presented, followed by the web content mining approach used. The method is then discussed, followed by the results and discussion, before final conclusions are drawn.

2. BACKGROUND

ICT education draws upon a number of standard curricula produced by academia, industry representatives and professional bodies, such as the Australian Computer Society (ACS) Core Body of Knowledge (Underwood, 1996), ISCC '99 (Lidtke et al., 1999) and IS 2002 (Gorgone et al., 2002); the primary focus of these is ICT knowledge, communication skills and technological competence (Snoke, Underwood and Bruce, 2002). The requirements of industry in terms of employers represent an important additional influence, since these ultimately determine the success of graduates in obtaining employment. Although academia has been found to agree with industry representatives belonging to a professional body on the most important attributes (Snoke and Underwood, 2001), graduates must ultimately fulfill the selection criteria for specific employment positions. These selection criteria are reflected in advertisements that are increasingly placed online, thus representing a valuable store of data that may be mined to determine key attributes that graduates must possess, together with the most beneficial combinations. However, to date only a limited range of analysis approaches have been performed, such as occurrence frequencies for specific technical skills (Gallivan, Truex III and Kvasny, 2004; Koong, Liu and Liu, 2002; Litecky and Arnett, 2001; Litecky, Prabhakar and Arnett, 1996, 2006; Liu et al., 2003; Prabhakar, Litecky and Arnett, 2005; Prabhakar, Litecky and Arnett, 1995), or augmentations of these with manual classifications and searches for specific skill and experience requirements (Gallivan, Truex III and Kvasny, 2004; Hardin, Joshi and Li, 2002; Koong, Liu and Liu, 2002); the latter clearly restrict the quantity of data that can be analyzed.

Previous studies have produced a number of useful skill categorizations, such as: technological and non-technological (Gallivan, Truex III and Kvasny, 2004); technological, systems and business (Hardin, Joshi and Li, 2002); technological, project management, business domain, sourcing and IT administration (Zweig et al., 2006); and technological, communication, teamwork, problem-solving, initiative and enterprise, planning and organisation, self-management and learning (Rundle-Thiele, Bennett and Dann, 2005).

3. THE WEB CONTENT MINING APPROACH

A key challenge of obtaining internet employment data for analysis is the inconsistent data structures that must be dealt with; even a single site is unlikely to use the same template indefinitely (Zhang and Simoff, 2006). Systems to perform this data extraction, known as wrappers, fall into two principal categories. Ontology-based systems require that the data is labeled or may be identified through the use of lexical patterns; where this is not the case, position-based software, which uses the HTML document structure to identify data of interest, can be used (Gregg and Walczak, 2006). Although position-based systems are less resilient to structural changes, they offer high accuracy (Chidlovskii, 2002; cited in Gregg and Walczak, 2006) and can even have their extraction rules learnt automatically rather than manually defined (Muslea, Minton and Knoblock, 1999). Position-based systems can also self-repair if errors are detected, and can incorporate ontology-based extraction (Gregg and Walczak, 2006).

This study uses a simple position-based Java wrapper to download all job advertisements from a web site. Since accuracy is paramount, the data is accessed at a single point in time. The system assumes that the web site allows users to search for jobs matching specified criteria and then produces a series of pages, each containing hyperlinks to jobs matching the criteria. The site is also assumed to allow the user to navigate between multiple results pages by using a URL containing a page number value. The user can enter the text corresponding to these hyperlinks and URLs, together with a cookie value if required by the site, and then save these settings for future use.

Once the wrapper has downloaded documents from the internet, a further key challenge is the extraction of useful information from these. Such web content mining is most commonly performed using text mining, which can perform tasks such as clustering, classification, association pattern extraction, topic discovery and tracking (Srivastava, Desikan and Kumar, 2004). The process of text mining can begin with a collection of documents, from which a document is retrieved and preprocessed to check its character set and format; text analysis is then performed to extract useful information, which can be placed in a management information system (Fan et al., 2006). Text mining can be applied to a diverse range of areas, such as biomedicine, government intelligence and education (Fan et al., 2006).

Leximancer (www.leximancer.com; Smith, 2000) is an example of a sophisticated text mining system, and works by mapping word groups, at a multi-sentence resolution, to concepts; these concepts are then made available to the user, together with details of their co-occurrences (Smith, 2003). Although Leximancer concepts are single words, each concept represents a collection of words that tend to appear together within the text. Leximancer searches for concepts by examining the most common words and selecting those that tend to co-occur with larger numbers of other words; these 'seed concepts' are then augmented by adding other words that are sufficiently relevant ("Leximancer Manual (Version 2.23)," 2007).

The Leximancer system is well established; it is available commercially, and papers describing it date back to 2000 (Smith, 2000). Leximancer has previously been used to analyze patient records (Watson, Smith and Watter, 2005), interview data (Finger, Jamieson-Proctor and Watson, 2005), emails (Murphy and Levy, 2006) and discussion board transcripts (de la Varre, Ellaway and Dewhurst, 2005). Although Leximancer can over-generalize when words with low semantic content such as 'and' and 'is' have high frequency, and is sensitive to changes in parameter settings, its validity has been demonstrated extensively (de la Varre, Ellaway and Dewhurst, 2005; Grech, Horberry and Smith, 2002; Rooney, 2005; Smith and Humphreys, 2006). For example, (Grech, Horberry and Smith, 2002) found no significant difference between Leximancer and manual coding on a subset of their data. de la Varre, Ellaway and Dewhurst (2005) confirmed that the conceptual maps produced by Leximancer accurately described documents with known content. Leximancer has also been found to demonstrate reproducibility and yield high coding stability levels (Rooney, 2005). Leximancer is therefore suitable for analysing the downloaded data within this study.

Data can be analysed in a number of different ways using Leximancer: the identification of concepts, which are then grouped into categories (de la Varre, Ellaway and Dewhurst, 2005; Isakhan, 2005); automatic and manually defined concept definition (Grech, Horberry and Smith, 2002); and variations in concept frequencies over time (Scott and Smith, 2005). This study uses automatic concept identification so that the employment requirement themes arising purely from the examined documents can be identified. Category grouping is also performed to allow broader issues to be identified and comparisons across data sets to be simplified.

The employment themes are contrasted with those discovered within standard curricula. These curricula are analyzed with Leximancer using the same parameter settings as the employment data, to ensure consistency across data sets.

4. METHOD

The first part of this study investigated employment market data in Australia and the United States, to allow international comparisons to be performed; the differences between these countries would be expected to be relatively minor (Welsum and Vickery, 2005). The second stage, described subsequently, compared the results with those from standard curricula, specifically the Australian Computer Society (ACS) Core Body of Knowledge (Underwood, 1996), IS 2002 (Gorgone et al., 2002) and ISCC ’99 (Lidtke et al., 1999). The CareerOne web site was used for the Australian data. This is linked to 100 News Limited newspapers, and claims to offer more jobs than any competitor ("About CareerOne," 2006); moreover, unlike competitors such as Seek it allows positions to be sorted in order of relevancy. The US data was gathered from CareerBuilder, a site commanding a 39% market share, compared to 37% and 25% respectively for its competitors Monster and HotJobs (Ruiz, 2006).

The data was obtained from the Career One web site by searching for all IT jobs on 15th August 2007 using the default settings provided on the site, since these would be expected to be used by the majority of users. This meant that positions were at any location within Australia, any contract type (permanent, part-time or temporary) and sorted by the relevancy rating given by the site. The site permits up to 1000 job positions to be viewed before demanding that search criteria be revised; this number of advertisements was therefore downloaded. The search also returned an additional 100 special advertisements, named ‘spotlights’ which appeared to be selected at random and contained numerous duplicates; these were therefore excluded.

The US data was obtained from CareerBuilder by searching for 1000 IT jobs using the default settings provided on 17th August 2007, which provided similar parameters to the Australian data; this means that jobs were sorted by date (most recent first), at any US location, covered all employment types (full time, part time, contractor and intern), industries, salaries and education requirements. The sample size was restricted to 1000 to permit comparability with the Australian sample. The US data contained HTML script and style tags that Leximancer identified as concepts; such tags were therefore removed before analysis was performed.

Leximancer was applied to the downloaded pages for the Australian and US data separately, using default settings, and a list of the most frequently occurring concepts produced. Key concepts were formed into groups and further analysed to determine the principal concepts related to them using a similar approach to Isakhan (2005). The grouping was undertaken manually by examining a sample of the source data from which the concepts were drawn, to determine the likely meaning of the concept. Although variation in source data was expected, if this was sufficiently high to make identification of the most appropriate group difficult then the concept was placed in the miscellaneous group.

The second part of this study involved comparing the employment market concepts with those from a number of standard curricula (as outlined next), again using Leximancer to perform the extraction. The curricula were formed into two sets: the first contained examples produced exclusively by US and Australian professional bodies, and the second held those formed by collaboration between industry and academia.

The first set of curricula examined contained the Australian Computer Society (ACS) Core Body of Knowledge (Underwood, 1996) and the IS 2002 curriculum. The IS 2002 curriculum was produced by the US-based ACM, AIS and AITP (Gorgone et al., 2002) and is the successor to IS'97 (Davis et al., 1997). The second set of curricula contained the ISCC ’99 example (Lidtke et al., 1999). The concept extraction for both sets was performed to allow the curriculum information to be analyzed using exactly the same methodology as the job advertisement data and thus provide a clearer comparison than could be achieved using existing qualitative and quantitative comparisons of ICT curricula (Lemmen, Mulder and Brinkkemper, 1999).

The data sets used provided the opportunity to undertake comparisons between the following categories: Australian and US employment; employment (both Australian and US) and curriculum (both the ACS Core Body of Knowledge/IS2002 and ISCC ’99 sets); curricula produced by professional bodies and the ISCC ’99 curriculum from academia and industry.

5. RESULTS

5.1 Employment Market Data

5.1.1 Australian data: The overall ranking of the most frequently occurring concepts within the Australian employment data is presented in Table 1, together with the principal components related to the most important concepts (Table 2 and Table 3). The frequency of concept occurrences is given by the absolute count value; this value is expressed as a percentage of the largest absolute count value to obtain the relative count (Murphy and Levy, 2006). Where the concept with largest absolute count value offers little meaning and is removed, such as within Table 1, the maximum relative count will be less than 100%.

The concepts were cleaned to remove terms that offered little meaning, such as 'career' and 'job'. The concept 'experience' was the principal item of interest, followed by 'skills', 'team', 'development', 'technical', 'business', 'years', 'design', 'support', 'systems', 'knowledge', 'solutions', 'technologies', 'application', 'requirements', 'network', 'Java', 'data' and 'service'. The concepts of interest most closely related to 'experience', the most popular example of interest, were 'years', 'development', 'skills' and 'knowledge'. The second most popular concept of interest, 'skills', is most closely related to 'experience', 'team', 'technical' and 'business'.

Concept	Absolute Count	Relative Count
experience	1363	99.1%
skills	1058	77%
team	988	71.9%
development	949	69%
technical	657	47.8%
business	623	45.3%
years	611	44.4%
design	601	43.7%
support	599	43.5%
systems	566	41.1%
knowledge	534	38.8%
solutions	473	34.4%
technologies	410	29.8%
application	376	27.3%
requirements	353	25.6%
network	344	25%
Java	283	20.5%
data	213	15.5%
service	212	15.4%

Table 1: Principal concept ranking for Australian employment advertisement data.

Manual grouping of the concepts revealed five overall skill types: people, containing 'team'; business, containing for example 'solutions' and 'requirements'; technological, containing examples such as 'development' and 'support'; experience, containing items such as 'years' and 'systems'; and theoretical, which had no members but was created to allow other data sets with theoretical items to be covered. Table 4 shows how all the key concepts are divided amongst these sets. The results can also be viewed as a Leximancer concept map (Figure 1). The brightness level of each concept corresponds to its frequency within the data; the closeness of concepts relates to their co-occurrence with other similar concepts (Smith, 2005). Every concept has a related dot, with size corresponding to its connectedness. The absolute position of each dot relative to the center of the map has no meaning ("Leximancer Manual (Version 2.23)," 2007).

5.1.2 US data: The ranking of the most popular components from the US data is presented in Table 5, together with the principal components related to the most important concepts (Table 6 and Table 7). Figure 2 contains the concept map.

All concepts offering little useful information were again removed, such as 'interest' and 'help'. The principal concept of interest was 'experience', followed by 'work', 'skills', 'technical', 'support', 'years', 'design' and 'knowledge'. The concepts of interest most closely related to

Concept	Absolute Count	Relative Count
years	557	40.8%
development	252	18.4%
skills	208	15.2%
knowledge	162	11.8%
design	129	9.4%
systems	124	9%
technologies	115	8.4%
Java	113	8.2%
team	106	7.7%
business	89	6.5%
technical	82	6%
support	72	5.2%
application	71	5.2%
network	63	4.6%
solutions	63	4.6%
data	35	2.5%
service	32	2.3%
requirements	20	1.4%

Table 2: Principal concepts related to 'experience', the most popular Australian employment data concept of interest overall

Concept	Absolute Count	Relative Count
experience	208	19.6%
team	135	12.7%
technical	126	11.9%
business	104	9.8%
knowledge	96	9%
development	91	8.6%
design	90	8.5%
years	85	8%
systems	71	6.7%
support	67	6.3%
technologies	62	5.8%
solutions	47	4.4%
requirements	43	4%
service	41	3.8%
Java	38	3.5%
network	32	3%
application	31	2.9%
data	23	2.1%

Table 3: Principal concepts related to 'skills', the second most popular Australian employment data concept of interest overall

Category	Concepts (rank follows in brackets)			
	Employment data		Standard curricula data	
	Australian	**US**	**ISCC '99**	**IS 2002 / ACS**
Experience	Experience (1) Years (7) Systems (10) Technologies (13) Java (17)	Experience (1) Work (2) Years (6) Design (7)	Work (8)	
Technological skills	Skills (2) Development (4) Technical (5) Design (8) Support (9) Systems (10) Knowledge (11) Technologies (13) Application (14) Network (16) Java (17) Data (18)	Skills (3) Technical (4) Support (5) Design (7) Knowledge (8)	Enterprise (1) Software (4) Design (5) Development (6) Tools (7) Data (10) Computing (13) Computer (16) Business (22)	Development (1) Design (3) Software (4) Technology (5) Knowledge (6) Data (7) Implementation (10) Model (12) Computing (13) Computer (15) Programming (17)
People skills	Team (3)		Team (3) Group (11) Management (21)	Management (2)
Business skills	Business (6) Solutions (12) Requirements (15)		Project (2) Change (15)	Project (9) Organizations (11)
Miscellaneous			Skills (12) Problem(s) (14) Environment (17) Program (18) Time (19)	Skills (8) Program (14) Techniques (16) Context (18)
Theoretical skills			Concepts (9)	

Table 4: Concept groupings from the four data sets

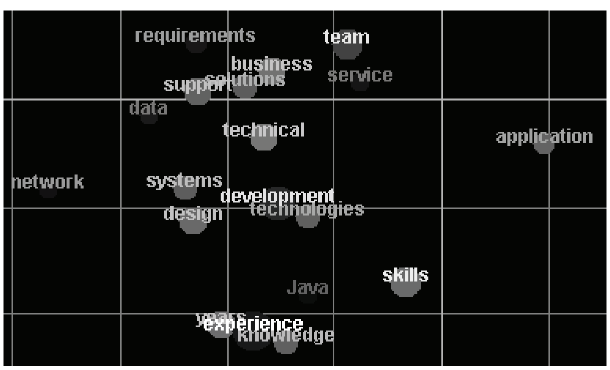


Figure 1: Concept map for Australian employment advertisement data

Concept	Absolute Count	Relative Count
experience	1848	60.9%
work	1584	52.2%
skills	1273	41.9%
technical	1209	39.8%
support	970	32%
years	863	28.4%
design	660	21.7%
knowledge	496	0.16%

Table 5: Principal concept ranking for US employment advertisement data

Concept	Absolute Count	Relative Count
years	736	39.8%
work	309	16.7%
skills	269	14.5%
technical	171	9.2%
knowledge	160	8.6%
support	149	8%
design	138	7.4%

Table 6: Principal concepts related to 'experience', the most popular US employment data concept of interest overall

Concept	Absolute Count	Relative Count
experience	309	19.5%
technical	229	14.4%
skills	190	11.9%
years	123	7.7%
support	103	6.5%
design	78	4.9%
knowledge	49	3%

Table 7: Principal concepts related to 'work', the second most popular US employment data concept of interest overall

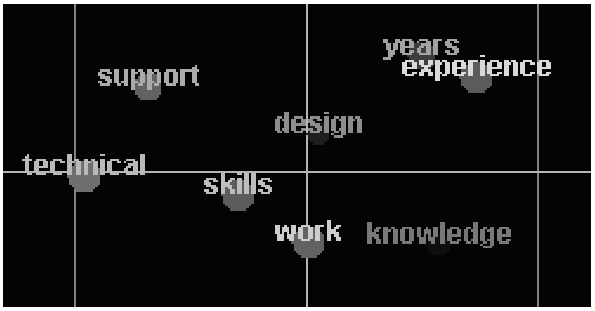


Figure 2: Concept map for US employment advertisement data

'experience' were 'years', 'work', 'skills', 'technical', 'knowledge', 'support' and 'design'; those related to 'work' were 'experience', 'technical', 'skills', 'years', 'support', 'design' and 'knowledge'. The manual grouping of the key concepts into the previously identified skill types is shown in Table 4.

Concept	Absolute Count	Relative Count
enterprise	102	39.3%
project	96	37%
team	76	29.3%
software	75	28.9%
design	73	28.1%
development	73	28.1%
tools	72	27.7%
work	70	27%
concepts	67	25.8%
data	65	25%
group	58	22.3%
skills	49	18.9%
computing	45	17.3%
problems	43	16.6%
change	43	16.6%
computer	41	15.8%
environment	40	15.4%
program	39	15%
time	35	13.5%
problem	34	13.1%
management	32	12.3%
business	29	11.1%

Table 8: Principal concept ranking for ISCC '99 curriculum data

5.2 Standard curricula data

The employment data results may be contrasted with those produced by applying Leximancer to the standard ISCC '99 curriculum, a collaborative project by US academia and representatives from primarily US based companies (Lidtke and Stokes, 1999; Lidtke et al., 1999), presented in full within Table 8 and displayed within a concept map (Figure 3). Again, many of the concepts were obvious and thus rejected, such as 'information', 'system(s)' and 'students'. The key concepts remaining from the original group were, in order: 'enterprise', 'project', 'team', 'software', 'design', 'development', 'tools', 'work', 'concepts', 'data', 'group', 'skills', 'computing', 'problem(s)', 'change', 'computer', 'environment', the broad 'program' and 'time', 'management' and 'business'.

In addition to the combined academia and industry, further comparison may be performed with two key curricula produced by professional bodies, namely the Australian Computer Society (ACS) Core Body of Knowledge (Underwood, 1996) and the IS 2002 curriculum (Gorgone et al., 2002) produced by the US based ACM, AIS and AITP. These results are presented in full within Table 9. Figure 4 contains the concept map, with only the top five concepts displayed for legibility. Again concepts such as 'system(s)', 'area(s)' and 'curriculum' were rejected. The key remaining concepts from the original group were, in order: 'development', 'management', 'design', 'software', 'technology', 'knowledge', 'data', 'skills', 'project', 'implementation', 'organizations', 'model', 'computing', the broad 'program', 'computer', 'techniques', 'programming' and 'context'.

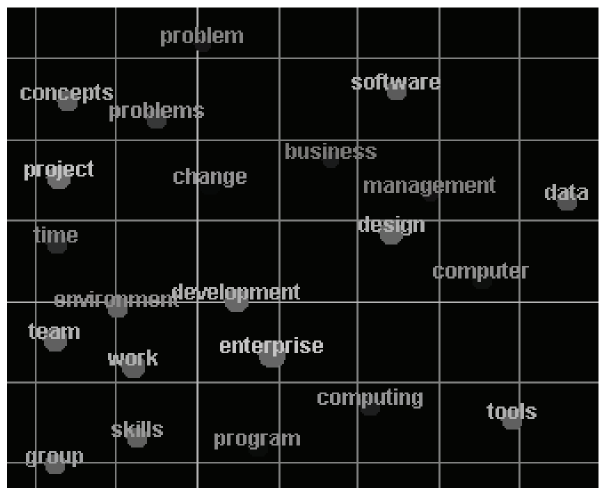


Figure 3: Concept map for ISCC '99 curriculum data

Concept	Absolute Count	Relative Count
development	126	46.3%
management	120	44.1%
design	112	41.1%
software	112	41.1%
technology	101	37.1%
knowledge	93	34.1%
data	65	23.8%
skills	55	20.2%
project	51	18.7%
implementation	48	17.6%
organizations	48	17.6%
model	41	15%
computing	40	14.7%
program	40	14.7%
computer	39	14.3%
techniques	36	13.2%
programming	33	12.1%
context	23	8.4%

Table 9: Principal concept ranking for ACS Core Body of Knowledge and IS 2002 curriculum data

Figure 4: Concept map for ACS Core Body of Knowledge and IS 2002 curriculum data

5.3 Combined data

The key concepts from each of the four data sets analysed are shown in Table 4. The concepts have been manually divided into categories, and for each data set the concepts are shown in order of their rank. The ranking of each category is determined by its highest ranked member; the categories are displayed in descending order of highest item rank within the employment data. For example, experience is the highest ranked of all the Australian employment data, and thus experience is the highest ranked category.

The concept categorisations used have some overlap with those of (Gallivan, Truex III and Kvasny, 2004), who define overall groupings into technological and non-technological skills, which are then further subdivided. The experience and theoretical categories used here are not included, but the technological example is present, and people and business skills appear similar to the non-technological examples of interpersonal/leadership and organisation. The technological group presented here is the only example with sufficient members to warrant further subdivision using the scheme of (Gallivan, Truex III and Kvasny, 2004). However, the only specific subgroups that contain concepts are: hardware, containing ‘computer’; programming languages, containing ‘Java’; application use, containing ‘application’; and communication, containing ‘network’. The operating systems, application development environment and CASE tools subgroups contain nothing. An alternative categorisation is presented by (Hardin, Joshi and Li, 2002), who have a business category but divide technical skills into systems examples, such as design and implementation, and technical items such as network and database. An additional categorisation into technical, project management, business domain, sourcing and IT administration also exists (Zweig et al., 2006).

The findings can also be analysed using the skill categories defined by the Australian DEST (Federal Department of Education, Science and Technology) study of employers (Rundle-Thiele, Bennett and Dann, 2005): communication; teamwork; problem-solving; initiative and enterprise; planning and organisation; self-management; learning; and technology. The technology category matches the example used here, and teamwork matches people skills; however, no other DEST categories are present here, and the experience, business and theoretical and business skills identified here are not recognised by DEST.

The results can be examined initially in terms of employment data. The US and Australian employment data sets show strong similarities, supporting previous research (Welsum and Vickery, 2005), with experience and technological skills valued most highly and theoretical skills being absent. However, the people and business skills required by Australian employers are absent within the US set, suggesting that the US positions are biased more towards early career stage positions (Lee et al., 2001).

The curriculum data is the next area of focus. Both curriculum data sets give technological skills highest priority, with people skills also highly rated and business skills present to varying degrees. However, the academia/industry curriculum (ISCC ’99) places higher priority on business skills than the curriculum produced by professional bodies (IS2002/ACS); it also includes experience and theoretical skills that are absent from the professional body curricula.

Comparisons between curricula and employment data suggest that they are well matched in terms of technical skills, with all giving them high priority and containing a wide range of examples. Industry and the curricula all place heavy emphasis on skills relating to software development and, with the exception of the US data, with only support and networks being absent from the curricula. Theoretical skills are also given low priority throughout, with the only presence being a single example within the academia/industry curriculum (ISCC ’99). However, experience differs significantly between these groups; although given highest priority within the employment data, only the ‘work’ example, at a relatively low level, exists within the academia/industry curriculum. The people and business skills required by Australian employers are well represented in the curricula.

6. DISCUSSION

The results suggest that standard curricula offer the key technical skills required by industry, along with the business and people examples required within Australia. However, the emphasis placed on experience would clearly benefit from adjustment, particularly in the professional body curricula, which appear less suitable than the industry-influenced ISCC ’99 and do not include experience. These findings show some support for previous research suggesting that ICT graduates are supplied the skills required by employers (Petrova and Claxton, 2005), and that curricula such as IS 2002 can form useful inputs into ICT curriculum design (Dwyer and Knapp, 2004).

The results show some agreement with a recent survey of primarily US employers (Bullen, Abraham and Galup, 2007), with business and project management skills, which are likely to be correlated with experience, valued more highly than technical examples. The authors echo the suggestions presented here of increasing the emphasis upon experience for students, and advocate internships; these are also suggested for faculty to increase their business awareness. Similarly, experience has been found to be rated very highly by employers, being contained in over 80% of job advertisements (Koong, Liu and Liu, 2002), and through internships leading to increases in job offers (Fang et al., 2004). This experience is particularly important when job opportunities are scarce (Sutcliffe, Chan and Nakayama, 2005).

The technical skills identified here have also proved to be important to recent ICT graduates in their jobs (Davis, 2003); however, the most popular skills nominated by the graduates were non-technical examples that were not identified in any of the sources within this study, such as thinking, a desire to learn and personal characteristics.

A survey of academics and industry representatives (Snoke and Underwood, 2001) supports the findings presented here that teamwork and technical skills are important. However, theoretical knowledge is given much higher priority, particularly by academics, whilst business skills are ranked less highly and experience is omitted. Further, many of the most important skills are non-technical in nature, such as learning, information retrieval and oral communication. Indeed, communication skills in general have been reported to be increasing in popularity (Crews, 2000).

The focus on technical skills by employers matches the findings of (Gallivan, Truex III and Kvasny, 2004; Hardin, Joshi and Li, 2002; Koong, Liu and Liu, 2002; Litecky and Arnett, 2001; Litecky, Arnett and Prabhakar, 2004; Litecky, Prabhakar and Arnett, 1996, 2006; Liu et al., 2003; Prabhakar, Litecky and Arnett, 2005; Prabhakar, Litecky and Arnett, 1995). A possible explanation for why this occurs despite organizational emphasis on non-technical examples is that technical skills are easier to screen (Gallivan, Truex III and Kvasny, 2004) and advertise (Litecky, Arnett and Prabhakar, 2004; Litecky, Prabhakar and Arnett, 2006). Indeed, Litecky, Prabhakar & Arnett (2004) hypothesize that recruitment initially filters applicants based on technical skills, for example through job advertisements, and then examines the non-technical skills through approaches such as interviews; the results presented here thus support this model.

An alternative explanation for the contrast between the business skills reported to be required by employers and the technical skills advertised is the seniority level of the positions. The entry level skills required by senior ICT executives have been found to be predominantly technical, with the balance shifting to favour business and project management examples as seniority increases (Zweig et al., 2006). Thus, the job advertisements analysed here may be skewed towards more junior positions, although the Australian positions appear to be more senior.

The business skills reported to be required by industry (Hardin, Joshi and Li, 2002; Milton, 2000; Zweig et al., 2006) appear only to be present within the Australian and not the US employment data here. Similarly, the industry requirement for project management skills (Kim, Hsu and Stern, 2006) appears not to hold for the Australian and US employment data. Other identified critical skills such as quality assurance, ERP, end user computing and security also appeared to be missed, although examples such as networking and support were present. The absence of ERP skills is surprising, particularly since their popularity with employers has lead to their incorporation within curricula (Boyle and Strong, 2006; Seethamraju, 2007); this may be because ERP skills attract relatively higher salary levels (Sager et al., 2006) but are less prevalent in more commonly occurring lower level roles such as programmers (Prabhakar, Litecky and Arnett, 2005). Communication skills, a secondary factor in recruitment (Koong, Liu and Liu, 2002), appear unimportant within the data sets examined here.

A recent study (Lee, 2005a) suggests that employers seek candidates with a range of technical (architecture/networks, hardware, software, problem solving and development), business and people (management and social) skills. Their results match the findings presented here in terms of technical examples being most important, with business and people skills also of key relevance.

The study has a number of limitations. Firstly only a single site is used in each country for analysis; however, this reduces the risk of data duplication through advertisements that appear on multiple sites, and a large data set is still available. Secondly, this approach has been criticized for missing large sized employers, who typically use their own corporate websites and provide several times the number of positions posted on commercial ICT job sites (Lee, 2005a). The effect of this drawback on the study appears to be relatively minor; the Career One website is used by large companies such as Virgin Blue, together with organizations listed on the Australian Stock Exchange S&P/ASX 100 list such as Coca-Cola Amatil, BHP Billiton and Rio Tinto. Further, almost all Fortune 100 companies supplement their corporate websites with at least one of the three major job sites (Careerbuilder, Hotjobs and Monster) (Lee, 2005b). Next, internet employment advertisements are clearly not exact representations of the positions to which they relate and may even exist independently from actual jobs. However, they are unlikely to be significantly removed from the required content of most positions and more importantly represent the first hurdle that must be cleared by graduates seeking employment.

Further, given the aim of comparing graduate attributes between two countries, the study used comparable data sources; that is, the most popular career websites in each country, during the same week. However, this means that alternative popular sources of job advertisements in each country have not been considered. For example, although university recruitment postings are popular in the US, they are not commonly used in Australia and thus were not considered.

An additional limitation is seasonality; the data was gathered in a single week, and thus did not take into account seasonal variations. However, only two of the four data sets were time-dependent; the remainder were standard curricula, with no seasonal variations. Further, the data was gathered in August, which fell outside major holidays such as Christmas or Easter. In addition, although in the case of the US data August is a summer holiday period, it was still possible to obtain 1000 advertisements; thus, whilst this is not a major recruiting period we would not expect the data to differ significantly from other ‘regular’ periods. Graduate recruitment campaigns are an additional potential source of seasonal variation. However, the data did not specifically target graduate positions; instead, all IT positions were gathered, which limited the extent to which this could affect the results. Further, graduation and recruitment campaign dates vary between institutions and across countries, and the occurrence of either in August appears to be unlikely.

While there are some potential ambiguities in the concepts produced by Leximancer, such as the inclusion of irrelevant background information in advertisements and lack of differentiation between essential, desirable and negated criteria, the potential impact of these has been minimized by the initial manual examination of key concepts.

Despite these limitations, this study does offer new insight into international ICT employment requirements, going beyond existing skill frequency based approaches by employing an innovative web content mining approach. This has facilitated the discovery of concepts emerging directly from the data, and the contrasting of these between categories such as countries, employers and curricula, and curriculum providers.

Overall this analysis suggests that current curricula match employers' emphasis upon technological skills, together with the people and business examples required in Australia. However, revision is required to increase their experiential content, particularly for the professional body examples.

7. CONCLUSIONS

This study has extended existing analysis of international ICT employment data through the use of a novel data mining approach. Further, through analysing curricula with the same technique, their appropriateness has been comprehensively examined. The curricula appear to be well suited to employers' skill demands, particularly with respect to technological examples. However, the key gap identified by this study is experience, and therefore the incorporation of this into curricula through approaches such as internships and real world projects would be strongly recommended.

8. REFERENCES

"About CareerOne," (2006), Retrieved 25 January, 2007, from http://www.careerone.com.au/jobs/about-careerone.

Boyle, T. A. and Strong, S. E. (2006), "Skill Requirements of ERP Graduates," Journal of Information Systems Education, Vol. 17, No. 4, pp. 403-12.

Bullen, C. V., Abraham, T. and Galup, S. (2007), "IT workforce trends: implications for curriculum and hiring," Communications of the AIS, Vol. 20, No. 34, pp. 545-54.

Chidlovskii, B. (2002), "Automatic repairing of Web wrappers by combining redundant views", Proceedings of the IEEE International Conference on Tools with Artificial Intelligence, Washington DC.

Crews, T. B. (2000), "Comparison of Communication Skills Needed by Information Systems Undergraduates and Graduates as Perceived by Information Systems Professionals," Journal of Information Systems Education, Vol. 11, No. 3-4, pp. 151-5.

Davis, D. C. (2003), "Job Titles, Tasks, and Experiences of Information Systems and Technologies Graduates from a Midwestern University," Journal of Information Systems Education, Vol. 14, No. 1, pp. 59-68.

Davis, G., Gorgone, J., Couger, J., Feinstein, D. and Longenecker, H. (1997), "IS 97 Model Curriculum and Guidelines for Undergraduate Degree Programs in Information Systems," Retrieved 9 December, 2004, from www.aisnet.org/Curriculum/Is97.pdf.

de la Varre, C., Ellaway, R. and Dewhurst, D. (2005), "Analysis of the large-scale use of online discussion boards in an undergraduate medical course", Proceedings of the 12th International Conference of the Association for Learning Technology, Manchester.

Dwyer, C. and Knapp, C. A. (2004), "How Useful is IS 2002? A Case Study Applying the Model Curriculum," Journal of Information Systems Education, Vol. 15, No. 4, pp. 409-16.

Fan, W., Wallace, L., Rich, S. and Zhang, Z. (2006), "Tapping the power of text mining," Communications of the ACM, Vol. 49, No. 9, pp. 77-82.

Fang, X., Lee, S., Lee, T. E. and Huang, W. (2004), "Critical Factors Affecting Job Offers for New MIS Graduates," Journal of Information Systems Education, Vol. 15, No. 2, pp. 189-203.

Finger, G., Jamieson-Proctor, R. and Watson, G. (2005), "Measuring Learning with ICTs: An external evaluation of Education Queensland's ICT Curriculum Integration Performance Measurement Instrument", Proceedings of the AARE Conference, Parramatta.

Gallivan, M. J., Truex III, D. P. and Kvasny, L. (2004), "Changing Patterns in IT Skill Sets 1988-2003: A Content Analysis of Classified Advertising," The DATA BASE for Advances in Information Systems, Vol. 35, No. 3, pp. 64-87.

Gorgone, J., Davis, G., Valacich, J., Topi, H., Feinstein, D. and Longenecker, H. (2002), "IS 2002 Model Curriculum and Guidelines for Undergraduate Degree Programs in Information Systems," Retrieved 9 December, 2004, from www.aisnet.org/Curriculum/IS2002-12-31.pdf.

Granger, M. J., Dick, G., McKinnell Jacobson, C. and Van Slyke, C. (2007), "Information Systems Enrollments: Challenges and Strategies," Journal of Information Systems Education, Vol. 18, No. 3, pp. 303-11.

Grech, M. R., Horberry, T. and Smith, A. (2002), "Human Error in Maritime Operations: Analyses of Accident Reports using the Leximancer Tool", Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, Baltimore, MD.

Gregg, D. G. and Walczak, S. (2006), "Adaptive Web Information Extraction," Communications of the ACM, Vol. 49, No. 5, pp. 78-84.

Hardin, A. M., Joshi, K. D. and Li, X. (2002), "Business as usual? IS job skill requirements during the internet era", Proceedings of the Eighth Americas Conference on Information Systems.

Isakhan, B. (2005), "From Despotism to Democracy: Reporting Iraq's January 2005 election in the Australian and Middle Eastern print media", Proceedings of the Journalism Education Association Conference.

Kim, Y., Hsu, J. and Stern, M. (2006), "An Update on the IS/IT Skills Gap," Journal of Information Systems Education, Vol. 17, No. 4, pp. 395-402.

Koong, K., Liu, L. and Liu, X. (2002), "Study of the Demand for Information Technology Professionals in Selected Internet Job Portals," Journal of Information Systems Education, Vol. 13, No. 1, pp. 21-38.

Lee, C. (2005a), "Analysis of Skill Requirements for Systems Analysts in Fortune 500 Organizations," Journal of Computer Information Systems, Vol. 45, No. 4, pp. 84-92.

Lee, I. (2005b), "The Evolution of E-Recruiting: A Content Analysis of Fortune 100 Career Web Sites," Journal of Electronic Commerce in Organizations, Vol. 3, No. 3, pp. 57-68.

Lee, S., Yen, D., Havelka, D. and Koh, S. (2001), "Evolution of IS professionals' competency: An exploratory study," Journal of Computer Information Systems, Vol. 41, No. 4, pp. 21-30.

Lemmen, K., Mulder, F. and Brinkkemper, S. (1999), "Information Systems Curricula Evaluated by a Method Engineering Framework," Requirements Engineering, Vol. 4, pp. 210-20.

"Leximancer Manual (Version 2.23)," (2007), Retrieved 23 September, 2008, from http://web.archive.org/web/20070411204124/www.leximancer.com/documents/Leximancer2_Manual.pdf.

Lidtke, D. and Stokes, G. (1999), "An information systems-centric curriculum, ISCC '99," Journal of Systems and Software, Vol. 49, No. 2-3, pp. 171-5.

Lidtke, D., Stokes, G., Haines, J. and Mulder, M. (1999), "ISCC'99 An Information Systems-Centric Curriculum '99 Program Guidelines for Educating the Next Generation of Information Systems Specialists, in Collaboration with Industry," Retrieved 24 July 2006, from http://www.iscc.unomaha.edu/TableOfContents.html.

Lightfoot, J. M. (1999), "Fads versus fundamentals: The dilemma for information systems curriculum design," Journal of Education for Business, Vol. 75, No. 1, pp. 43-50.

Litecky, C. R. and Arnett, K. P. (2001), "An update on measurement of IT job skills for managers and professionals", Proceedings of the Seventh Americas Conference on Information Systems.

Litecky, C. R., Prabhakar, B. and Arnett, K. P. (1996), "MIS job market: Shaken but not stirred," Journal of Systems Management, Vol. 47, No. 4, pp. 50-3.

Litecky, C. R., Arnett, K. P. and Prabhakar, B. (2004), "The paradox of soft skills versus technical skills in IS hiring," Journal of Computer Information Systems, Vol. 45, No. 1, pp. 69-76.

Litecky, C. R., Prabhakar, B. and Arnett, K. P. (2006), "The IT/IS Job Market: A Longitudinal Perspective", Proceedings of the ACM SIGMIS CPR Conference on Computer Personnel Research, Claremont, California, April 13-15.

Liu, X., Liu, L., Lu, J. and Koong, K. (2003), "An Examination of Job Skills Posted on Internet Databases: Implications for Information Systems Degree Programs," Journal of Education for Business, Vol. 78, No. 4, pp. 191-6.

Milton, T. (2000), "Cross training the answer to e-commerce staff shortages," Computer Weekly.

Murphy, M. and Levy, M. (2006), "Politeness in Intercultural Email Communication: Australian and Korean Perspectives," Journal of Intercultural Communication, No. 12.

Muslea, I., Minton, S. and Knoblock, C. (1999), "A hierarchical approach to wrapper induction", Proceedings of the 3rd Conference on Autonomous Agents.

Petrova, K. and Claxton, G. (2005), "Building Student Skills and Capabilities in Information Technology and eBusiness: A Moving Target," Journal of Information Systems Education, Vol. 16, No. 1, pp. 27-41.

Prabhakar, B., Litecky, C. R. and Arnett, K. P. (1995), "Boom times ahead!," Journal of Systems Management, Vol. 46, No. 1, pp. 24-8.

Prabhakar, B., Litecky, C. and Arnett, K. (2005), "IT skills in a tough job market," Communications of the ACM, Vol. 48, No. 10, pp. 91-4.

Rooney, D. (2005), "Knowledge, economy, technology and society: The politics of discourse," Telematics and informatics, Vol. 22, No. 4, pp. 405-22.

Ruiz, G. (2006), "Newspapers, job boards step up partnerships," Workforce Management, Vol. 85, No. 23, pp. 17-8.

Rundle-Thiele, S., Bennett, R. and Dann, S. (2005), "The successful preparation and development of future marketing professionals: a recommended methodological framework," Journal for Advancement of Marketing Education, Vol. 7, pp. 27-36.

Sager, J., Mensching, J., Corbitt, G. and Connolly, J. (2006), "Market Power of ERP Education - An Investigative Analysis," Journal of Information Systems Education, Vol. 17, No. 2, pp. 151-61.

Scott, E., Alger, R., Pequeno, S. and Sessions, N. (2002), "The Skills Gap as Observed between IS Graduates and the Systems Development Industry: A South African Experience", Proceedings of the IS2002 Conference, Cork, Ireland.

Scott, N. and Smith, A. E. (2005), "Use of automated content analysis techniques for event image assessment," Tourism Recreation Research, Vol. 30, No. 2, pp. 87-91.

Seethamraju, R. (2007), "Enterprise Systems (ES) Software in Business School Curriculum - Evaluation of Design and Delivery," Journal of Information Systems Education, Vol. 18, No. 1, pp. 69-83.

Smith, A. E. (2000), "Machine Mapping of Document Collections: the Leximancer System", Proceedings of the Fifth Australasian Document Computing Symposium, Sunshine Coast, Australia.

Smith, A. E. (2003), "Automatic Extraction of Semantic Networks from Text using Leximancer", Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, Edmonton.

Smith, A. E. (2005), "Leximancer Manual," Retrieved 16 January, 2008, from http://www.leximancer.com/documents/Leximancer2_Manual.pdf.

Smith, A. E. and Humphreys, M. S. (2006), "Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping," Behavior Research Methods, Vol. 38, No. 2, pp. 262-79.

Snoke, R. and Underwood, A. (2001), "Generic Attributes of IS Graduates: A Comparison of Australian Industry and Academic Views," Journal of Global Information Management, Vol. 9, No. 2, pp. 34-41.

Snoke, R., Underwood, A. and Bruce, C. (2002), "An Australian view of generic attributes coverage in undergraduate programs of study: An information systems case study", Proceedings of the Proceedings of HERDSA 2002.

Srivastava, J., Desikan, P. and Kumar, V. (2004), "Web Mining - Concepts, Applications and Research Directions," in Data Mining: Next Generation Challenges and Future Directions, H. Kargupta, A. Joshi, K. Sivakumar and Y. Yesha (eds.), MIT Press, Cambridge MA, pp. 405-23.

Sutcliffe, N., Chan, S. S. and Nakayama, M. (2005), "A Competency Based MSIS Curriculum," Journal of Information Systems Education, Vol. 16, No. 3, pp. 301-10.

Underwood, A. (1996), "The ACS Core Body of Knowledge for Information Technology Professionals," Australian Computer Society.

Watson, M., Smith, A. E. and Watter, S. (2005), "Leximancer Concept Mapping of Patient Case Studies," Lecture Notes in Computer Science, Vol. 3683, pp. 1232-8.

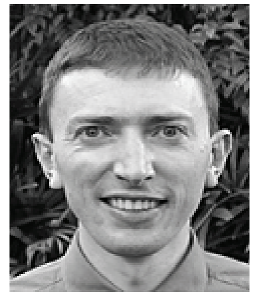
Welsum, D. and Vickery, G. (2005), "New perspectives on ICT skills and employment / Working Party on the Information Economy," Retrieved 11 January, 2008, from http://www.oecd.org/dataoecd/26/35/34769393.pdf.

Wilkins, M. L. and Noll, C. L. (2000), "Critical Skills of IS Professionals: Developing a Curriculum for the Future," Journal of Information Systems Education, Vol. 11, No. 3-4, pp. 105-13.

Zhang, D. and Simoff, S. J. (2006), "Informing the Curious Negotiator: Automatic News Extraction from the Internet," in Data Mining Theory, Methodology, Techniques and Applications, G. J. Williams and S. J. Simoff (eds.), Springer, Berlin, pp. 176-91.

Zweig, P., Kaiser, K. M., Beath, C. M., Bullen, C., Gallagher, K. P., Goles, T., Howland, J., Simon, J. C., Abbott, P., Abraham, T., Carmel, E., Evaristo, R., Hawk, S., Lacity, M., Gallivan, M., Kelly, S., Mooney, J. G., Ranganathan, C., Rottman, J. W., Ryan, T. and Wion, R. (2006), "The Information Technology Workforce: Trends and Implications 2005-2008," MIS Quarterly Executive, Vol. 5, No. 2, pp. 101-8

AUTHOR BIOGRAPHIES



Justin Debuse is a Lecturer in Information Systems at the University of the Sunshine Coast. He received his PhD from the University of East Anglia. His research encompasses topics such as data mining, automated assessment, optimization and visual programming. He has published in journals such as Systems Research and Behavioral Science, Computer Journal, Journal of Applied Intelligence and Journal of Intelligent Information Systems.

Meredith Lawley is an Associate Professor of Marketing at the University of the Sunshine Coast. She received her PhD from the University of Southern Queensland. Her research interests include consumer behavior, international education, e-business, e-marketing and services marketing.