COMPETENCES DESIRED BY EMPLOYERS OF HUMAN RESOURCES MANAGEMENT IN LOUISIANA: IMPLICATIONS FOR IMPROVING ELECTRONIC RESUMES FOR JOB SEEKERS

Adelina Kaliba[1], MBA

2018 graduate of the College of Business at Southern University and A&M College

ABSTRACT

A resume is a medium that summarizes job seekers' competences, and it is the first item that captures an employers' attention. Since most employers are using employment-related online search engines, to be matched with qualified applicants, resumes should contain keywords that are related to job descriptions. I analyzed descriptions of 1,059 human resources jobs advertised on indeed.com by employers located in Alexandria, Baton Rouge, Lafayette, Monroe, New Orleans, and Shreveport–the main metropolitan cities in Louisiana. I used different data exploratory and visualization techniques to summarize the data and then applied latent semantic analysis (LSA) to extract vocabularies that employers use to highlight the required competencies of a prospective employee. Application of Microsoft Office tools and experience in working in a business-oriented environment were respectively important hard and soft skills wanted by most employers. Resumes that highlight these skills are likely to be matched with advertised jobs and therefore increases the likelihood for job interviews.

Keywords: job descriptions, resume, human resources**,** latent semantic analysis**,** text mining.

[1] Contact: akaliba92@gmail.com; LinkedIn: linkedin.com/akaliba

INTRODUCTION AND CONTEXT

The objective of this study is to identify competencies in demand by employers located in Louisiana metropolitan cities focusing on human resources related jobs. The aim is to suggest combinations of keywords to include in a resume that will increase the likelihood of a job seeker to be invited to multiple job interviews. Parameswara (2015) defines human resources as the people or personnel who are employed within the organization. According to Armstrong (2014), the profession of Human Resources Management (HRM or HR) is a strategic, integrated, and coherent approach to the employment, development, and well-being of the people working in organizations. Employees are an organization’s greatest asset because without them, essential functions or process could not be completed or achieved. Haslinda (2009) argues that the management aspect of human resources is the process of managing human talents to achieve the organization’s objectives, and goals. Armstrong (2014) explains that human resources management is a complex profession that encompasses the application of policies and practices in the field of organizational design and development, performance measurement, reward and compensation, and the provision of services that enhance the well-being of employees. Employers need talented and devoted human resource employees to perform these complex tasks.

As discussed in Aslam (2013) and Burma (2014), the profession of HRM will always be in demand because it is a role that has become indispensable for 21st century modern businesses. Organizations consider the HR department as playing a significant role in staffing, training, and helping to manage employees so that people and the organization is performing at maximum capability. Because of increasing competitiveness, and the global and technology driven landscape, organizations are always in need of skilled and competent HR professionals working in human resource management. These unique challenges require applicants seeking HRM related employment to have acquired specific competences[2], either through academics or work experience, which will enable them to successfully perform in today's challenging HR roles. White (1959) explains that to demonstrate competence, employees must be able to perform certain tasks or skills with a required level of proficiency and specific skills or tasks can be described in terms of specific behaviors at distinct levels of proficiency. Generally, competencies include both hard (technical) and soft skills (intrapersonal and interpersonal skills). Laker and Powell (2011) show that hard skills are technical and domain specific, such as the ability to use a certain computer language or perform statistical analysis. The soft skills are personal attributes that help employees engage and collaborate in a team. Hard skills are somewhat quantifiable and encompass the declarative and procedural knowledge that is important for a position and can be reflected by applicants past work experiences, degrees, and certificates. Soft skills provide an indication of how individuals will tend to behave and interact with others. Since soft skills are attributes of individuals, they are harder to measure (Heckman and Kautza, 2012). Soft skills are not discipline specific but cut across all professions and job positions, especially when team work is important.

To attract talented and motivated employees, employers are increasingly using job search engines such as indeed.com for recruitment. According to their website, indeed is the most comprehensive search engine for jobs and offers free access to millions of jobs from thousands of company websites and job boards. As a huge aggregator of postings from across the web, indeed.com consolidates listings from many job boards into one place. It also compiles information from various company career pages and allows job seekers to search locally or globally. The engine helps employers to manage the recruiting and screening process and engages job seekers with

[2] Competencies are generally a set of skills and abilities (technical as well as behavioral) which are required for desired level of performance acquired through education and experience or on job training. Competencies integrate knowledge, skills, judgment, and personal attributes that people need to perform a job effectively.

employers from interview to hire. It uses advanced technology to promote the jobs, target skilled job seekers, and find candidates that match an employer's needs. Generally, the job search engine uses three steps: collect information from the resume, put collected information into database, and use matching algorithm to sort and transmit the relevant resume to the prospective employer.

The reliance on job search engines by employers poses one fundamental challenge for job seekers: the employer never sees the applicant's resume and the machines are entrusted with screening potential applicants. Depending on the text retrieval and analysis method used by the search engine, screening is therefore based on the similarity of structure of words, or on meaning of words in the resume, or a combination of both (i.e., semantic of the resume). For a job seeker, the problem is writing a resume that matches with the many job descriptions to increase the chance for an interview. The research question is how to align competencies presented by the resume with the competencies needed by various employers? To answer this question for Louisiana's job seekers, I collected 1,059 human resources job descriptions and desired experience from indeed.com and used latent semantic analysis (LSA) as a basis for information retrieval. The job descriptions and desired experiences contain a variety of information with distinct levels of details. Extracting important criteria will lead to identifying a pattern of competences (hard and soft skills) required by employers, and job seekers could therefore write and submit a generic electronic resume with better job matching results.

First introduced in Dumais et al. (1988) and expanded by Deerwester et al. (1990) as a technique for improving information retrieval, LSA is a machine learning method for representing the meaning of words, sentences, and texts. LSA induces a high-dimensional semantic space from reading an exceptionally large amount of texts. The meaning of words and texts can be represented as vectors in this space and hence can be compared automatically and objectively. Landauer and

Dumais (1997) indicated that the underlying idea of LSA is that the aggregate of all the word contexts in which a given word does and does not appear provides a set of mutual constraints that largely determines the similarity of meaning of words and sets of words to each other. Application of LSA to identify keywords includes Müller et al. (2016) study, which uses LSA to analyze business process management-related job advertisements to develop a typology of individual competences for this profession. Ahmad and Laroche (2016) used LSA to explore the topics of positive and negative reviews and identifying topics that were helpful to potential customers. D'Haen et al (2016) used the approach to identify better prospects and improve sales call efficiency via web data. A study by Meire, Ballings, and Poel (2016) use LSA to map important predictors and associated relationship that help to identify sentiment in Facebook posts. In the literature, I found limited information on using LSA to better the resume of job seekers in the human resources discipline. This paper will close some of the gaps existing in the literature and will help job seekers to write competitive and viable resumes.

This study focuses on Louisiana because, overall, the state has experienced a slow but a steady job growth. In addition, job competences wanted by employers, particularly those caused by the technological advance, are usually regional specific. Donovan (2017) estimates that in the six-month period that ended in June 2017, Louisiana gained an average of 3,300 jobs per month, outpacing the national rate of job growth. The analysis also shows that, in 2017, three-fourths of working-age adults in the state where employed. With the economy in Louisiana slowly progressing, HRM professionals will be greatly needed to support and to oversee the talent management and productivity strategies that ultimately impact every organization's bottom line. As discussed in Lawrence (2017), the improving economy has contributed to job growth in the HR field. Furthermore, at about 4.5 percent, the unemployment rate for HR professionals is about half of the national unemployment rate. Armed with a resume that aligned with demanded competences, human resources job seekers are likely to experience employability and success in their job search.

This paper proceeds as follows: in the following section I present brief literature on what LSA is, what can be done with LSA, and how I used the technique to achieve the objective of the paper. In section 3, I discuss the source of data and method used to extract information from indeed.com and methods for data analyses. I respectively discuss the results and present the implications arising from this study in section 4 and 5.

A REVIEW OF LATENT SEMANTIC ANALYSIS AND APPLICATION

To support effective labor market policies, educational and occupational proxies are the two reference points commonly applied in skills demand and supply analyses. Although these two proxies might help to produce employment projections for the industries at a global level, Manacorda and Manning (1999, 2007) argue that transitions specific to an industry or specialization cannot be accurately reflected. Cedefop (2015) suggest using indirect measures that have advantages of being relatively easy to use and indicate the kind of competencies required through employers' job descriptions. Preference for online recruitment has created a demand for job-search engines. Websites are being created, updated, and actively promoted to invite employers and jobseekers into virtual interaction. These activities produce enormous amounts of text that can be retrieved, analyzed, and used to map competencies demanded by employers. Using job descriptions and desired experiences to mine competencies has the advantage of being more flexible and up-to-date. As for each position, employers describe actual competencies in demand for each academic field and occupation.

There are two-main text mining approaches that use job descriptions to extract required competencies of each job posting: clustering and n-gram analysis. Text categorization using clustering techniques as applied in Han, Karypis, and Kumar (2001), is a classification technique for deciding whether a document belongs to a set of pre-specified classes of documents. The key element of this approach as discussed in Lowe (1995) is the availability of a similar measure that can identify textual neighbors within a document. A major drawback of this approach is a tendency of using all features when computing similar distances. In many textual data, however, only smaller number of the total vocabulary may be useful in categorizing documents. The n-gram analysis is based on modeling the relationship between documents and their n-gram phrases, which are a set of co-occurring words within a given text (Ceska, Hanak, and Tesar, 2007).

As one of the important form n-gram[3] formulation, latent Semantic Analysis (LSA) investigates the hidden associations between the documents and their unique n-gram phrases derived from the Singular Value Decomposition (SVD) computations (Simmons and Estes, 2006). Therefore, LSA is the process of finding the meaning behind a series of words, used to gain an understanding of the attitudes, opinions and emotions expressed by the text. It is a method for open-ended text analysis using sophisticated statistical and mathematical algorithms (Papadimitriou et al, 2000) to reveal subtle textual meaning using an automated approach that eliminates potential human bias and permits rapid coding of copious amounts of data. According to Landauer and Dumais (1996) and Sun et al. (2008), LSA is widely used in applications of information retrieval, spam filtering, and automated essay scoring. Landauer, Foltz, and Laham (1998) argue that apart from extracting and standing for the contextual-usage meaning of words in

[3] According to Wikipedia and in the fields of computational linguistics and probability, an n-gram is a contiguous sequence of n-items from a given sample of text or speech. The items can be phonemes, syllables, letters, words, or base pairs according to the application. They are basically a set of co-occurring words within a given text.

the text (perception), LSA also produces measures of word-word, word-passage, and passage-passage relations that are well correlated with several human cognitive phenomena involving association or semantic similarity. A main advantage of LSA is that the usage and meaning of words in a sentences or texts and the results are represented as vectors in space and hence can be compared automatically and objectively. This is based on the idea that text does have a higher order. Latent semantic structure is often obscured by the use of synonyms or polysemy, and the use of SVD can discover these latent structures.

In their study, Litecky et al. (2010) use several techniques that incorporated web and text mining techniques to retrieve advertisements and abstract facts from text to estimate skills demand from online vacancies. The study aimed at extracting specific competences from job descriptions, and therefore evidence about skills needs in software engineering. The data was extracted from online repositories, parsed, and filtered based on a set of predefined keywords. They used hierarchical agglomerative clustering to identify groups of skills producing coherent job definitions. Ahmed, Capretz and Campbell (2012) used a similar approach to mine demanded soft skills from job advertisement in software development disciplines. In their study, Kurekova, Haita, and Beblavy (2012) analyzed job advertisements to identify and quantify skills and personal attributes in demand in the Slovak labor market. The study demonstrates how mining of job postings can identify competences demanded by employers. The results showed that there are significant differences in skills demand for relatively similar jobs with slightly different job titles. Wang et al. (2008) employed clustering and n-gram analysis to segment online advertisements by geography and job type to identify competences suitable for disabled Chinese jobseekers. Debortoli et al. (2014) applied latent semantic analysis to develop competency taxonomies among business intelligence and big data job advertisements in the United States.

As defined by Landauer and Dumais (1997) LSA is therefore a theory and method for extracting and representing the contextual-usage meaning of words by statistical computations applied to a large corpus of text. Therefore, LSA can be applied in two ways: obtaining approximate estimates of the contextual usage of words and extract common factors that represent concepts hidden in a text. We use both approaches to extract competencies demanded by the employers of human resource managers in Louisiana’s main metropolitan cities which include Alexandria, Baton Rouge, Lafayette, Monroe, New Orleans, and Shreveport.

SOURCE OF DATA AND DATA ANALYSES

All data were extracted from indeed.com. As a starting point, the engine asks to enter the tittle of the job (e.g., human resources) and the location of the job (e.g., Baton Rouge). The results include a list of human resource jobs located in the selected area. The search engine has been optimized over time to return listings relevant to a given query, so it would know a job like “staffing manager or recruiter” is relevant to the “human resources” query. All advertised jobs have a title, the name of the employers, a job description, and desired experience. Some employers include a salary range. The data was extracted on July 22, 2018 using rvest package (Wickham, 2016) that work in the R software environment (R Core Team, 2018). After the first page at https://www.indeed.com, the second page “~/jobs?q=Human+Resources&l=Baton+Rouge%2C+LA&start=10” search for human resources jobs in Baton Rouge. You need only need to add numbers after “start=”, in increments of ten to move from the second page to another. Each page contains about 14 to 15 job postings and I used this strategy to extract all job posting from each page.

For the six metropolitan cities in Louisiana, there was 1,059 human resources job listings: 52 jobs in Alexandria, 400 jobs in Baton Rouge, 105 jobs in Lafayette, 48 jobs in Monroe, 364 jobs in New Orleans, and 90 jobs in Shreveport. There was also a combination of 37 unique hard skills desired by employers. The combination of unique and desired skills was respectively 21 in Alexandria, 134 in Baton Rouge, 34 in Lafayette, 3 in Monroe, 133 in New Orleans, and 54 in Shreveport.

I followed three steps to achieve the study objectives. In the first step, I used extracted job descriptions and desired experiences to set up a term-frequency matrix. The second step involved transforming the term frequencies into a term-document matrix using various weighting schemes[4]. In the third step I performed singular value decomposition (SVD) on the matrix to reduce the dimensionality. Specifically, to set up the text as a term-document matrix, this study started the analysis with text preprocessing procedures. Before creating the matrix, the text preprocessing procedure consists of term reduction by removing stop words and tokenization[5].The job description text generated a dictionary of 12,343 tokenized terms and the desired experience text generated a dictionary of 1,899 tokenized terms. Term reduction and tokenization was implemented using a tidytext package (Silge and Robinson, 2016). These two dictionaries produced respectively two term-frequency matrices with 1,515 columns and 3,299 rows (terms) and columns and rows (terms).

While there are many methods to define term weights as explained in Han and Kamber (2006), the traditional *tf-idf* weighting was used to transform the raw term frequencies in the matrix before conducting LSA. The *tf-idf* weighting scheme combines term frequency (*tf*) and inverse document frequency (*idf*) together such that:

$$W_{ij} = tf_{ij}\ [\ idf_i], \tag{1}$$

[4] Term frequency matrix shows how frequently an expression (term, word) occurs in a document and a term-frequency matrix measures the association of a term with respect to a given document and an entry in the matrix contains the number of times a term occurs in a document.

[5] Stop words are words with little meaning in natural language processing such as "a”, "the", "and", "an", and similar words and tokenization divides documents into a set of terms.

where tf_{ij} is term frequency (TF) and idf_i is the inverse document frequency (IDF) of term *i*. The IDF represents the importance of a term and is defined as:

$$idf_i = log2\left(\frac{N}{df_i}\right) + 1. \quad (2)$$

In Equation 2, *N* is the total number of documents and df_i is the document frequency (DF) of term *i*. The IDF in Equation (2) implies that the discriminative power of a term will be decreased if it occurs in many documents or the importance of a term will increase if it appears in a limited number of documents. The reasoning behind the *tf-idf* weighting is that a term occurring frequently in a document but rarely in the rest of the collection is more important.

LSA can be considered as an application of reduced-order Singular Value Decomposition (SVD) of W_{ij} in Equation 1. Practically, SVD decomposes a weighted term-document matrix W_{ij} into the product of three other matrices:

$$W_{ij} = [W_o][S_o][C_o], \quad (3)$$

where, Equation (3), W_0 is the value-weighted matrix of terms and C_0 is the transpose matrix of documents, which are the matrices of left and right singular vectors of SVD and S_0 is the diagonal matrix of singular values. In Equation 3, W_0 has the same number of rows as W_{ij}, similarly C_0 has the same number of columns as W_{ij}, and the S_0 is a square matrix with non-zero entries only along one central diagonal and sorted in decreasing order.

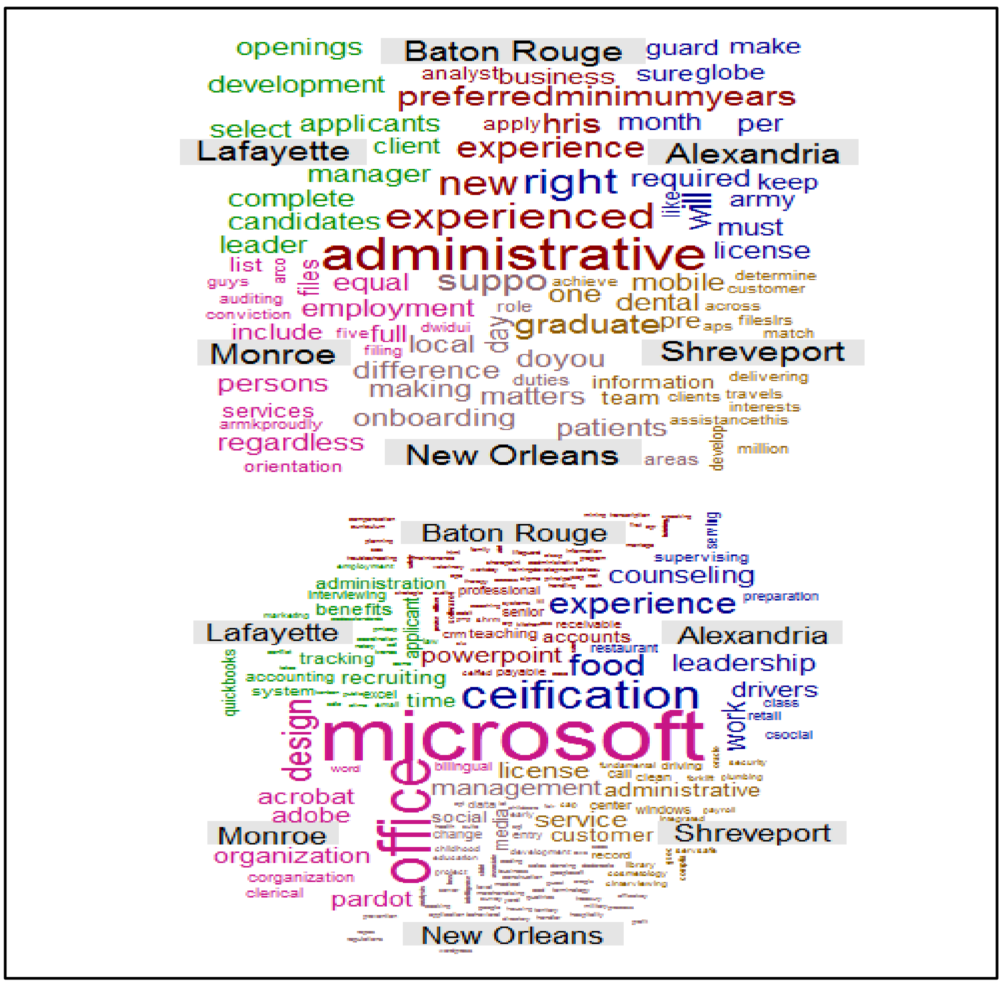
Under LSA and using SVD techniques, the dimensionality of the W_{ij} in Equation (3) is reduced by keeping the first *k*-largest coefficients in the diagonal matrix S_0 and setting the remaining smaller ones to zero. The zero rows and columns of S_0 is then deleted to get a new diagonal matrix *S*. Similarly, the corresponding columns of W_0 and C_0 are removed to obtain *W* and *C* respectively. The product of the simplified matrices is a new matrix $\widehat{W}_{ij}$ such that $\widehat{W}_{ij} = [W][S][C]$, which is the *k*-rank matrix with the best possible least-squares-fit to W_{ij}. The results of

SVD therefore include one set of *k*-factor loading for the terms and one for the documents. Refer to Deerwester et al. (1990) and Hofman (2017) for technical details of applying SVD during latent semantic analysis and LSA package (Wild, 2015) for estimating Equation (1) through (4).

RESULTS AND DISCUSSION

Figure 1 is a comparison cloud that shows words that have different frequencies among the six metropolitan cities and the size of each word indicates its frequency or importance. Results in Figure 1 assist with exploratory textual analysis by identifying words that frequently appear in job description and desired experience and communicating the most salient words in both texts reporting stage.

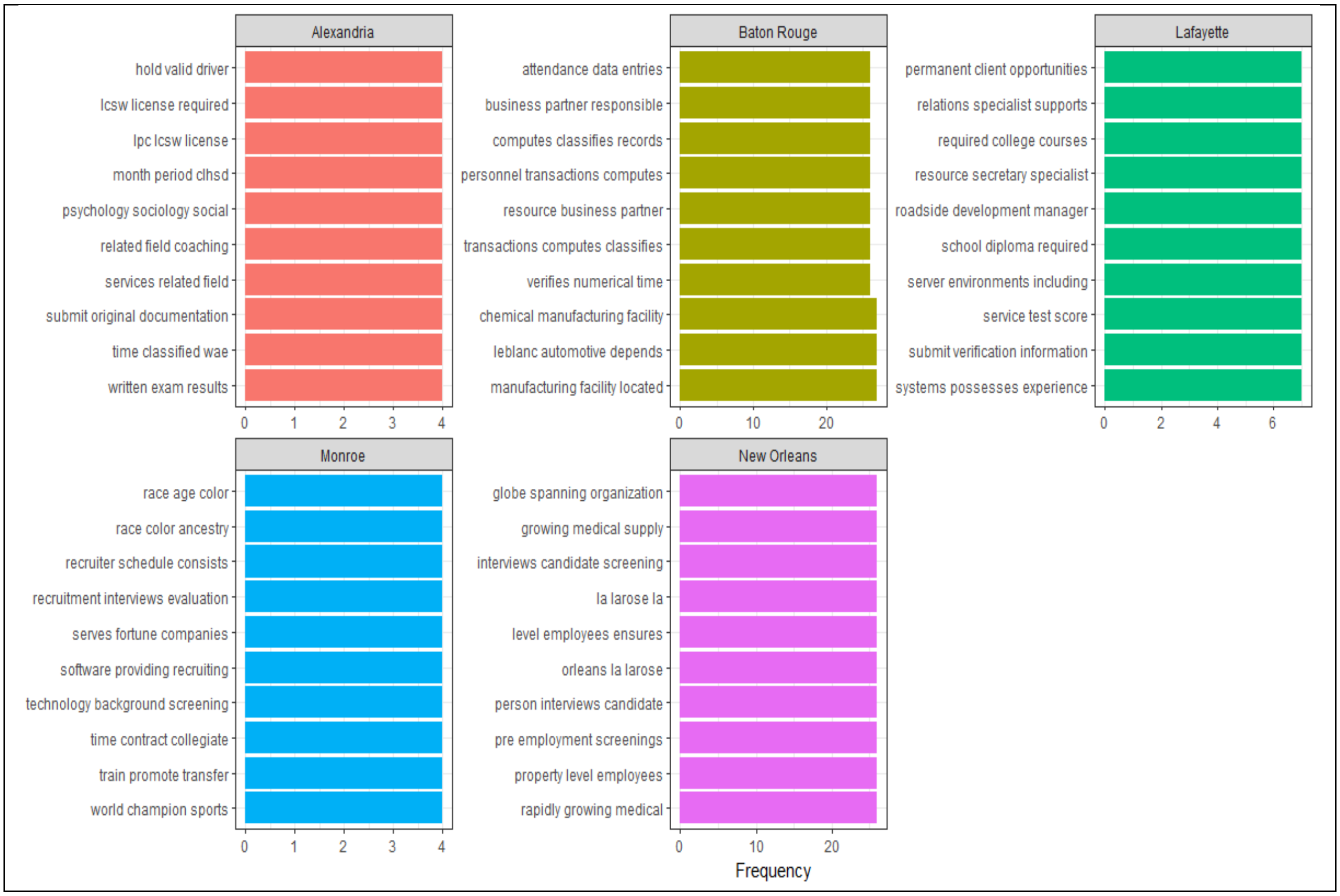
Figure 1: Word cloud from job descriptions and desired experiences

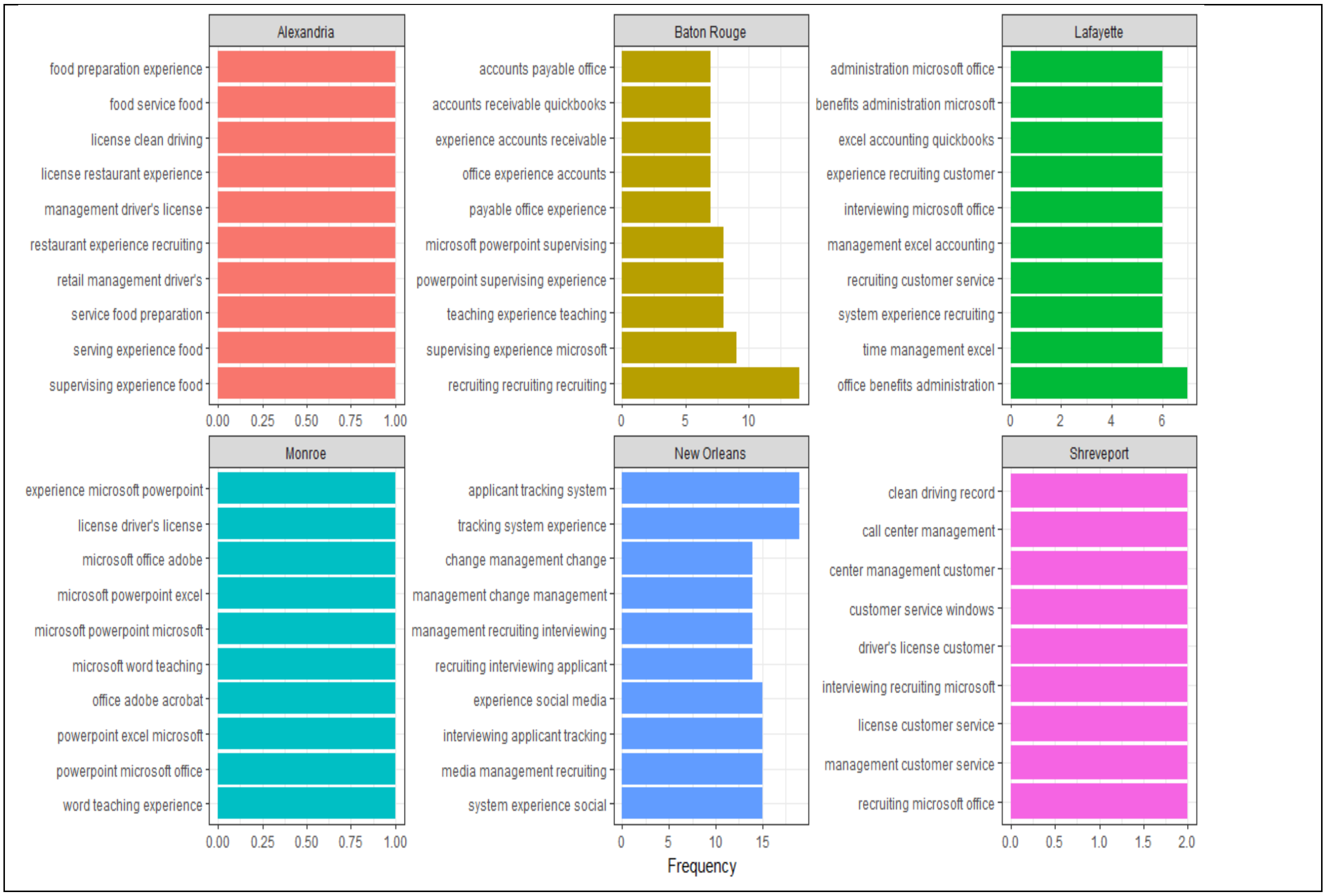


The top part of Figure 1 shows results from the job descriptions and the lower part shows the results from the desired experiences. Words that appears in the central location of Figure 1 are commonly mentioned in all metropolitan cities, and words in the periphery are associated with a specific metropolitan area. As it can be seen in Figure 1, administrative experience was frequently mentioned by most of the job descriptions, and Microsoft Office Certification was an important hard skill desired by most employers. For each metropolitan area and as regard to job description keywords, for Monroe it was full employment; team and information in Shreveport; local, making, and matters in New Orleans; experience, minimum years in Baton Rouge; right and license in Alexandria. Regarding desired expertise it is adobe and acrobat in Monroe; social media in New Orleans; recruiting and administrative benefits in Lafayette; professional in Baton Rouge; and counseling leadership in Alexandria. Although word clouds are engaging in terms of revealing essential key words, providing an overall sense of the text has several drawbacks. Word clouds are less accurate because of design features and may lead to loss of information due lack of context especially if slightly different words are used for the same idea so much that the meaning of individual words are lost. Because of these limitations, I use the results from the word clouds for exploratory qualitative analysis and I expand the analysis to identify co-occurring words and the results are presented in Figure 2.

Results in Figure 2 show ten trigrams (a group of three consecutive words) from both the job descriptions and desired experiences, which are a special case of the n-gram, where n is 3. The results in Figure 2 show the relationship between words that tend to follow others immediately, or that tend to co-occur within the job description or desired experience, which are useful for frequency analysis. The first part of Figure 2 shows the trigrams from job descriptions and the second part shows trigrams from desired experiences. While frequencies of different trigram from job descriptions differs across the six metropolitan cities, the first ten trigrams have equal distribution within the metropolitan cities. The first 10 bigrams appeared 4 times for job descriptions in Alexandria, Lafayette, and Monroe, and appeared 20 times in Baton Rouge and New Orleans. There were no valid trigrams for Shreveport. These generated trigrams also varied by metropolitan cities without specific patterns.

Figure 2: Distributions of trigrams from job descriptions and desired experiences

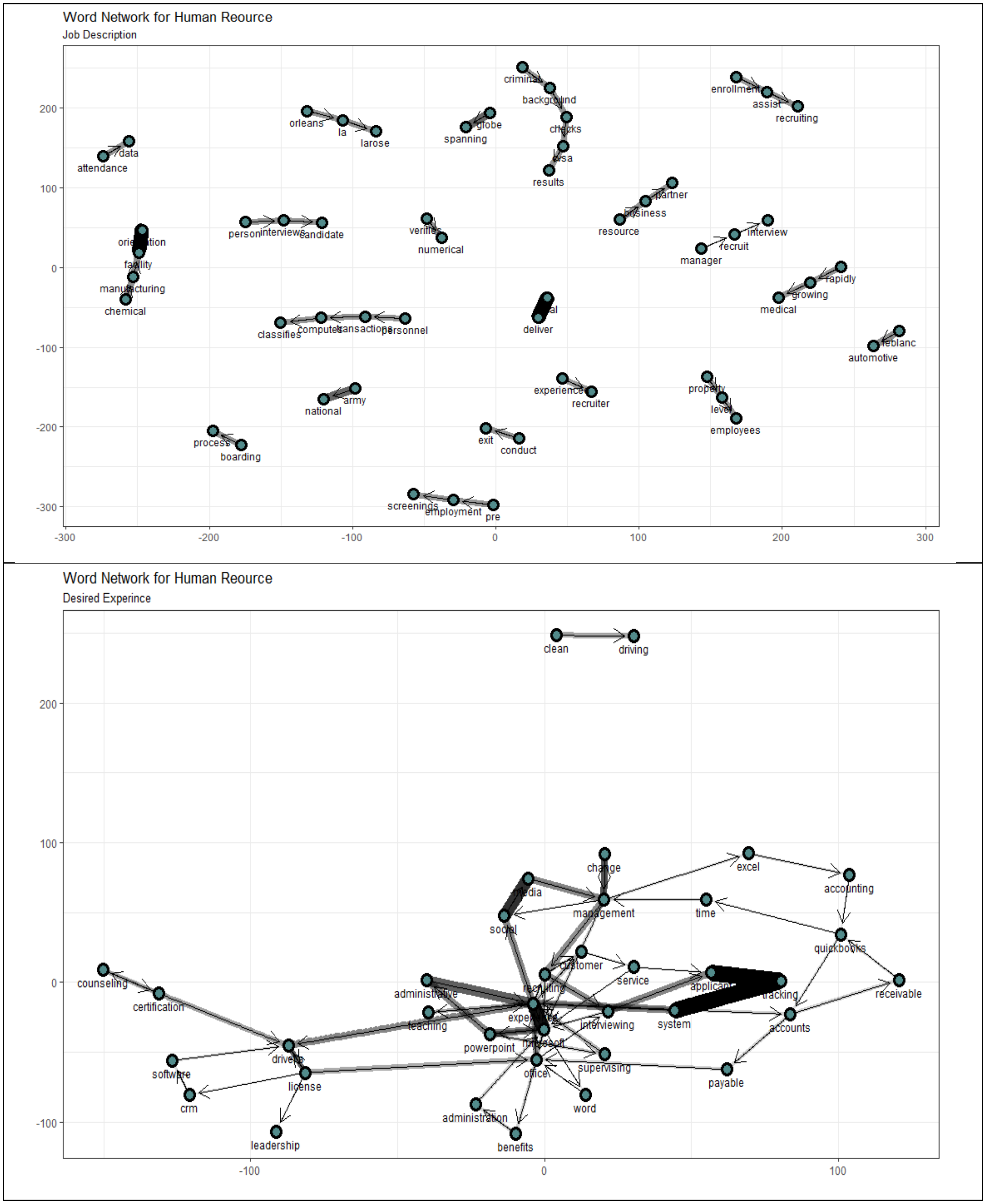




The first 10 trigram for desired experience appeared once in Alexandria and Monroe, occurred twice in Shreveport and six times for most of trigrams in Lafayette. The occurrences in Baton Rouge varies from 7 to 17 and varied from 14 to 21 in New Orleans. The trigrams in Alexandria were mainly related to the restaurant industry, in Monroe and Lafayette the relationship is towards application of Microsoft Office. In Baton Rouge, New Orleans, and Shreveport the desired experience is hinged on application of Microsoft Office with other skills such accounting, recruiting, and customers services.

Although n-grams analysis is very flexible and provide resilience against minor errors in the text by allowing matches on portions of words rather than requiring the entire word during matching, it is possible that a resulting n-gram can be ambiguous. Results of n-gram analysis can be extended to network analysis that provides both a visual and a mathematical analysis of word relationships as shown in Figure 3.

Figure 3: Networks of the trigrams from job descriptions and desired experiences



In Figure 3, apart from mapping and measuring of relationships individual words, network analysis shows the flows between words and other connected information as shown in Figure 3. To avoid information clutter, we present the network analysis results for the entire sample. The nodes in the network are the words while the links indicate relationships or flows among the nodes. In Figure 3, the first part represents the network presentation for the job description and the second part present the results for the desired experience. There are about 20 unconnected networks. The longest network has six nodes and is related to the criminal background check. Although a background check is often the last step taken by employers to help ensure a sound hiring decision and protect the employer from several potential risks, for many employers, a background check is a reliable way of verifying claims made by job seekers during the hiring process, so they are qualified for the jobs they are applying for. Two networks had four nodes and are respectively related to chemical facility orientation with emphasis to facility orientation and using computers for classified personal. The information contained in other networks is vague, but some networks indicates different job description and required soft skills such as pre-employment screening and experience in recruiting and interviewing.

Network analysis results for the desired experience has two distinct networks. The first network has two nodes and is related to a need for clean driving records. The second and the bigger network in centered on experience and Microsoft with some satellites around driver, office, tracking, QuickBooks, and management terms. Strong (thickest) outgoing arrows from experience are pointing towards Microsoft that connects to Microsoft Office that connects to office and PowerPoint, interviewing that connect to applicants, and social that connect to media. The strong incoming arrows from administrative experience and system that originates from applicant tracking. Experience in social media management obviously a high need, followed by experience in interviewing applicants, and experience in power points and use of Microsoft office for employee and employees' benefits management. Other skills are an application of QuickBooks and Microsoft Excel for accounting purposes, teaching for certifications in counseling, and ability to use customer relationship management (CRM) software. Leadership and supervising do not appear in a loop but plays a supporting role in office management and training.

Network analysis has an advantage of providing a convenient representation and summary of words as a team functioning. However basic network analysis does not help to distinguish between effective communication and breakdown. Communication breakdowns, which is attributed to long multi-hop communications along the network because of the small number of words in the network, and the substantial number of communications (arrows) among them. The most useful feature of LSA is representing words and documents by means of vectors in a semantic space, therefore, allowing for the computation of word and document similarities. Neighborhood computations that allow for simple similarity computations LSA space must be a matrix object in which each row represents a word vector of the space, with row names specifying these words as character strings. This implies that all word vectors have the same dimensionality. We first created a term-by-document frequency matrix with and applied the weighting as explained in Equation (1) and (2) and conducted SVD and performing a dimensionality reduction with LSA function in LSA package that creates an LSA space. I recovered W0 in Equation (3) and run the cosine similarity of all words in the matrix. I use the results in Figure 1 to determine the 5 nearest words for the most prominent terms. For job description, I use administrative, experience, new, HRIS (human resources information systems), years, preferred, graduate, employment, persons, and minimum. For desired experience we use Microsoft, office, certification, food, experience, counseling, design, recruiting, organizations, service, and customer. The results are presented in Table 1.

Table 1: Closest Neighbors Using the Term Matrix

Job Description			Desired Experiences		
TERM	**Neighbor**	**Cosine Distance**	**TERM**	**Neighbor**	**Cosine Distance**
administrative	minimum	0.8167	experience	serving	0.9912
	HRIS	0.8161		preparation	0.9912
	software	0.8126		planning	0.9892
	years	0.7571		strategic	0.9892
	dental	0.7475		ITIL	0.9892
experience	preferred	0.7439	certification	lifeguard	0.9943
	achieve	0.7379		cpr	0.9914
	across	0.7379		childhood	0.9907
	customers	0.7379		early	0.9907
	greater	0.7379		education	0.9907
new	hire	0.8944	counseling	work	0.9994
	information	0.6963		six	0.9949
	closely	0.6678		sigma	0.9949
	execution	0.6678		certification	0.9862
	functional	0.6678		aid	0.9851
administrative	minimum	1.0000	Food	aid	0.9959
	software	0.9994		teaching	0.9884
	years	0.9815		catering	0.9853
	excel	0.8797		production	0.9853
	knowledge	0.8797		program	0.9821
years	software	0.9817	design	access	0.9997
	HRIS	0.9815		workers	0.9983
	minimum	0.9814		business	0.9977
	minimum	0.8167		serving	0.9912
	knowledge	0.8832		active	0.9958
preferred	excel	0.8540	recruiting	system	0.9901
	knowledge	0.8540		applicant	0.9893
	Microsoft	0.8540		tracking	0.9893
	outlook	0.8540		interviewing	0.9828
	PowerPoint	0.8540		taleo	0.9804
graduate	training	0.9990	organization	skills	1.0000
	HRIS	0.8161		preparation	0.9912

	school	0.9788		tableau	0.9999
	semester	0.9753		analysis	0.9997
	substitute	0.9753		administration	0.9997
employment	activities	0.9035	guest	handler	1.0000
	conducting	0.9035		hospitality	1.0000
	coordinate	0.9035		profit	1.0000
	letter	0.9035		services	1.0000
	manage	0.9035		territory	0.9848
person	phone	1.0000	customer	management	0.9995
	evaluation	0.9989		service	0.9966
	psychology	0.9989		change	0.9959
	screening	0.9988		ITIL	0.9718
	candidate	0.9935		planning	0.9718
minimum	HRIS	1.0000	leadership	notary	0.9961
	software	0.9995		plumbing	0.9930
	years	0.9814		license	0.9888
	years	0.7571		strategic	0.9892
	knowledge	0.8779		certification	0.9847

The largest values show closely related terms and decrease in value means decreasing relationships. Words with similar values mean that they co-occur across all documents in the matrix. For example, for job description, we can logically see that administrative term closely co-occurs with minimum and another logical word is years and HRIS that stand for human resource information. This means that for HR job-seekers applying for the administrative job, need to emphasize previous experience in administrative positions and experience in using the HRIS system. While experience in job description co-occurs with words are related to customer services under desired experience the word aligns with preparing strategic plans and the experience in using ITIL and acronym for Information Technology Infrastructure Library, which is set of detailed practices for IT service management (ITSM) that focuses on aligning IT services with the needs of business. Reading along and across Table 1 three observations emerge. While job descriptions are more diverse and may present desired both hard and soft skills, desired experiences focus on hard skills demanded by employers. Overall, both job descriptions and desired experience have an emphasis on hard skills. Proficiency in the use of Microsoft Office tools and other software such as HRIS and ITIL creates a comparative advantage for HR jobseekers.

SUMMARY AND CONCLUSION

Using different data visualization tools and latent semantic analysis, I was able to determine important and valued skills a person seeking a position in human resources should include in the resume. By extracting human resources job descriptions from indeed.com, parsing, and aggregating the information into a usable data frame, I was able to see that the job site provides useful information to job seekers. As employers continue to use job-search engines for job advertisements, job seekers are faced with a difficult task of creating a resume that maximizes the likelihood for the job interviews. The results suggest that individuals seeking administrative jobs the resume should highlight their experience in using human resources management software (hard skills) and managing working environments (team work management) an important soft skill. For fresh graduates or unexperienced human resources personnel, the resume should highly emphasize the ability to use Microsoft Office tools and spread sheet-based software such as QuickBooks. Soft skills related to counseling, teaching, interviewing, and event organizations will have a higher competitive advantage.

REFERENCE

Ahmed F, Capretz, L.F. and P. Campbell (2012). Evaluating the demand for soft skills in software development. IT Professional archive 14(1): 44-49.

Ahmad, S. N. and M and Laroche. (2017). Analyzing electronic word of mouth: A social commerce construct. International Journal of Information Management 37(3): 202-213.

Armstrong, M. T. (2014). Armstorng's handbook of human resources management practice.11th Edition. Kogan Page: London and Philadelphia.

Aslam, H. A. (2013). Importance of Human Resources Management in 21st Century: A Theoretical Perspective. International Journal of Human Resources Studies 3(3): http://www.macrothink.org/journal/index.php/ijhrs/article/view/6255.

Burma, Z. (2014). Human resources management and its importance for today's organizations. International Journal of Education and Social Science 1(2): www.ijessnet.com

Ceska, Z., Hanak, I. and R. Tesar (2007). Teraman: A tool for N-gram extraction from large datasets. Proceedings of the 2007 IEEE International Conference on Intelligent Computer Communication and Processing.

Cedefop (European Center for Development of Vocational Training) (2015). Skills, qualifications, and jobs in the eu: the making of a perfect match and evidence from Cedefop's European skills and jobs survey Luxembourg: Publications Office. Cedefop reference series; No 103. http://dx.doi.org/10.2801/606129.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. and R Harshman (1990). Indexing by latent semantic analysis. Journal of the American Society for Information Science, 41, 391–407.

Donovan, J. A. (2017, July 5). State of working Louisiana. Louisiana Budget Project. http://www.labudget.org/lbp/wp-content/uploads/2017/09/SWLAFINAL-4.pdf.

D'Haen, J., Poel, D. V. d., Thorleuchter, D. and D.F. Benoit (2016). Integrating expert knowledge and multilingual Web crawling data in a lead qualification system. Decision Support Systems 82: 69-78.

Dumais, S., Platt, J., Heckerman, D. and M. Sahami (1998). Inductive learning algorithms and representations for text categorization. In Proceedings of CIKM'98, 7th ACM International Conference on Information and Knowledge Management (pp. 148–155).

Fridolin, W. (2015). LSA: Latent Semantic Analysis. R package version 0.73.1. https://CRAN.R-project.org/package=lsa.

Han E., Karypis, G. and V. Kumar (2001). Text categorization using weight adjusted k-nearest neighbor classification. 5th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), pp.53-65, 2001.

Haslinda, A. (2009). Evolving terms of human resources management and development. Journal of International Social Research 2(9): 180–187.

Heckman, J.J and T.D. Kautz (2012). Hard Evidence on Soft Skills. NBER Working Paper No. 18121. http://www.nber.org/papers/w18121.

Hofman, T. (2007). Probabilistic latent semantic indexing. ACM SIGIR Forum - SIGIR Test-of-Time Awardees 1978-2001. 51(2): 211-218. Doi: 10.1145/3130348.3130370.

Kureková, L., Haita C., and M. Beblavý (2013). Conceptualizing low skill: a new approach. Sociológia 45 (3): 247-266.

Lowe, D.G. (1995). Similarity metric learning for a variable-kernel classifier. Neural Computation 7(1): 72:85.

Laker, D.R. and J. L. Powell (2011). The differences between hard and soft skills and their relative impact on training transfer. Human Resources Development Quarterly 22(1): 111-122.

Lawrence, B. (2017). Human resource jobs outlook remains robust. Human Resources Certification Institute: https://www.hrci.org/community/blogs-and-announcements.

Landauer, T. K. and S.T. Dumais (1997). A solution to Plato's problem: the latent semantic analysis theory of the acquisition, induction, and representation of knowledge. Psychological Review 104: 211–240.

Landauer, T. K., Foltz, P. W. and D. Laham, D. (1998). An introduction to latent semantic analysis. Discourse Processes 25(2&3): 259–284.

Litecky C., Aken, A., Ahmad, A. and H. J. Nelson (2010). Mining for computing jobs. Software IEEE 27(1): 78-85.

Manacorda, M. and A. Manning (1999). Just can't get enough: more on skill-biased change and labor market performance; London School of Economics and Political Science: London, UK, 1999.

Manacorda, M. and A. Manning (2007). Shifts in the demand and supply of skills in the OECD: a single index model with a continuous distribution of skills. Oxford Bulletin of Economics and Statistics 69(5): 635-666.

Meire, M., Ballings, M. and D.V.D. Poel (2016). The added value of auxiliary data in sentiment analysis of Facebook posts. Decision Support Systems 89: 98-112.

Müller O.,Schmiedel T., Gorbacheva E. and J. vom Brocke (2016). Towards a typology of business process management professionals: identifying patterns of competences through latent semantic analysis. Enterprise Information Systems 10(1): 50-80.

Papadimitriou, C.H., Raghavan, P., Tamaki, H. and S. Vempala (2000). Latent semantic indexing: a probabilistic analysis. Journal of Computer and System Sciences, 61, 2000, 217- 235.

Parameswari, B. Y. and V. Yugandhar (2015). The role of human resources management in organizations. International Journal of Engineering Technology, Management and Applied Science 3(7): ISSN 2349-4476.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.URL https://www.R-project.org/.

Simmons, S. Z. and Z. Estes (2006). Using latent semantic analysis to estimate similarity. In

Proceedings of the 28th Annual Conference of the Cognitive Science Society (Hillsdale, NJ: Lawrence Earlbaum Associates), 2169–2173.

Sun J., Zhang Q., Yuan Z., Huang W., Yan X., Dong J. (2008) Research of spam filtering system based on LSA and SHA. In: Sun F., Zhang J., Tan Y., Cao J., Yu W. (eds) Advances in Neural Networks - ISNN 2008. ISNN 2008. Lecture Notes in Computer Science, vol 5264. Springer, Berlin, Heidelberg.

Wang C.H., Zhang, M., Ma, S.P. and L. Y. Ru (2008). Automatic hot event detection using both media and user attention. Journal of Computational Information Systems, 4(3): 985-992.

Wickham, H. (2016). rvest: Easily Harvest (Scrape) Web Pages. R package version 0.3.2. https://CRAN.R-project.org/package=rvest.

White, R. W. (1959). Motivation reconsidered: The concept of competence. Psychological Review 66(5): 297-333. http://dx.doi.org/10.1037/h0040934.