Explicitly demanded skills – investigating job ads

Karl Ledermüller[*]

Abstract **The following study reported here wants to gain understandings of frequency, combinations and clusters of skills, which are explicitly stated in job announcements. These findings can be used to think about curricular decisions in a global and local higher education marketplace. Data Clustering also gives the possibility to show "niche curricula". Furthermore, fields for meaningful learning activities and combinations of skills, which are not covered by their curriculum, are shown to several stakeholder groups of the educational process.**

Keywords – **employability, job advertisements, content analysis, cluster skills, knowledge society**

I. INTRODUCTION

Due to the fact, that education could be seen as a product [15] under a total quality management (TQM) approach, graduates could be seen as finished products and employers may be the customers of this "product" education. Of course education in general enfolds several stakeholders with different objectives [16]. Graduates and employers see the "fitness for purpose", which is the match of competencies and function, as the quality criterion for education. [9]

It may be an advantage for the "quality" of education in general to know more about this decision process. The aim of this study is to identify, model and analyse skill related attributes which are explicitly headed in job advertisements. It is true that job advertisements could be criticized, that they search for unrealistic candidates, but as Clyde [6] points out: "Statements in job advertisements represent the knowledge and skills that employers would like to have and are prepared to pay for". So this special form of text data seems to contain information about demanded skills in different jobs.

Text analysis of job advertisements can, therefore, provide a picture of demanded skills in several professions. Of course, the demand for skills of graduates is limited to the perception of the industry, because these texts are normally written for jobs in companies. Several studies were realized on analyzing job ads [2], [3], [13], [19].

This paper reports the preliminary results of a content analysis of 1200 job advertisements from the Career Center of the Vienna University of Economics and Business (WU). The job ads were collected from 2004 to 2008 and contain jobs for young business graduates. Automated text analysis using the R tm (text mining) package and a self written VBA for excel tool helped to classify the skills and show frequencies according to skill related words. The classification of skills has its theoretical framework in [17] and [4]. In a second step 484 finance related job ads were manually analyzed with content analysis in the logic of the classification. The result of the content analysis was a binary skill document matrix.

[*]Karl Ledermüller works as a research and teaching associate at the Institute of Operations Research, Department of Finance and Accounting, Vienna University of Economics and Business (WU), Austria

The following quantitative findings can be structured into three dimensions: The top demanded skills by frequency; skills related to professions and cluster effects of skills calculated by several multivariate clustering methods.

II. THEORETICAL BACKGROUND

According to the higher education stakeholder model, there are groups of stakeholders, who are interested in (high) employability of graduates due to the perceived quality of higher education.

Eagle/Brennan [18] show the quality perspective of several stakeholder groups for higher education based on [16].

TABLE I
QUALITY PERSPECTIVES [18]

Stakeholder group	Quality perspective
Funding bodies and society at large	Value for money; return on investment
Current and prospective students	High standards in order to gain an advantage in future employment
Employers	Competencies of graduates should match the functions required in employment
Academics and administrators within universities	Consistency, recognition of, and respect for, the challenges of educating a diverse student body

Several of the quality perspectives above indicate a demand for high employability of graduates. Value for money and return on investment of education could have the meaning (of course in just one dimension) that graduates should have the possibility to find and be able to keep a proper job due to the learned skills (value for money) and have the possibility to return the investment in form of earnings and taxes. To gain an advantage in future employment (expected employability) and match the functions required in employment (employability) is the quality criteria which explicitly states employability as a quality dimension. There seems to be empirical evidence, that students see a degree as a possibility for a better career [14]. As a consequence it can be assumed, that employability in general is an important issue in education.

Fugate and others developed a heuristic model for individual employability. They define employability as "a form of work specific active adaptability that enables workers to identify and realize career opportunities" [8] p.16.

They integrated the dimensions of Career Identity, Personal Adaptability and Social and Human Capital.

Under Career Identity they understand a representation of career experiences and aspirations. It is the past and present career that forms the future of one's career. [8]

Personal Adaptability is the flexibility of a person facing a changing working environment. Adaptable persons are willing and able to change personal factors. [8]

Social Capital can be measured in very general terms as the persons we know and how "important" they are. Human Capital consists of the skills and knowledge we have. [8] (Also see [5])
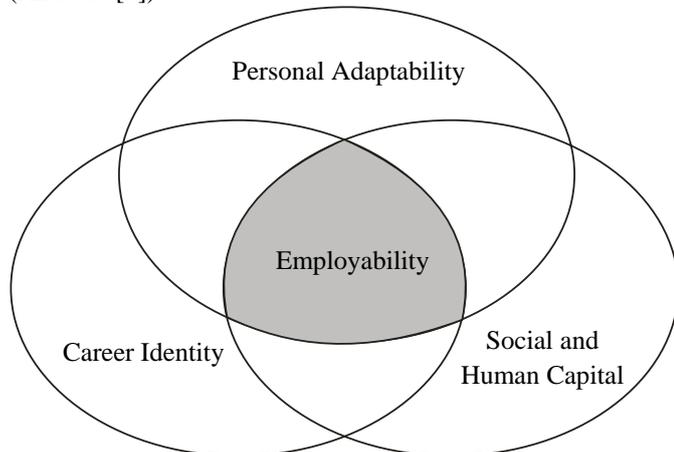

Fig.1. Heuristic model of employability [8]

According to [8] employers seek for graduates who match the functions required in employment. In other words, graduates with high employability are in demand. Job advertisements seem to be a data source for employability related factors, because "Statements in job advertisements represent the knowledge and skills that employers would like to have and are prepared to pay for." [6] Thus, they explicitly state factors of employability. For this reason, job advertisements seem to be a proper data source for employability studies.

III. LIMITATION OF THE DATA SOURCE "JOB ADVERTISEMENTS"

The investigated data of the study was a text corpus of over 1200 job advertisements for business graduates from the Career Center of the Vienna University of Business and Economics (WU). The 484 finance related job advertisements were analyzed by content analysis. A majority of the job advertisements was written in German (70%) the minority was written in English.

It seems to occur, that not all demanded skills are part of a job advertisement. There seem to be skills that are implicitly assumed and skills that are not expected from young graduates.

This effect can be shown for demanded language skills. Even though most of the job advertisements are written in German, language skills in German have a very low frequency. It can also be seen, that 50% of the ads with an explicit demand for German skills also have a demand for a foreign language except English.
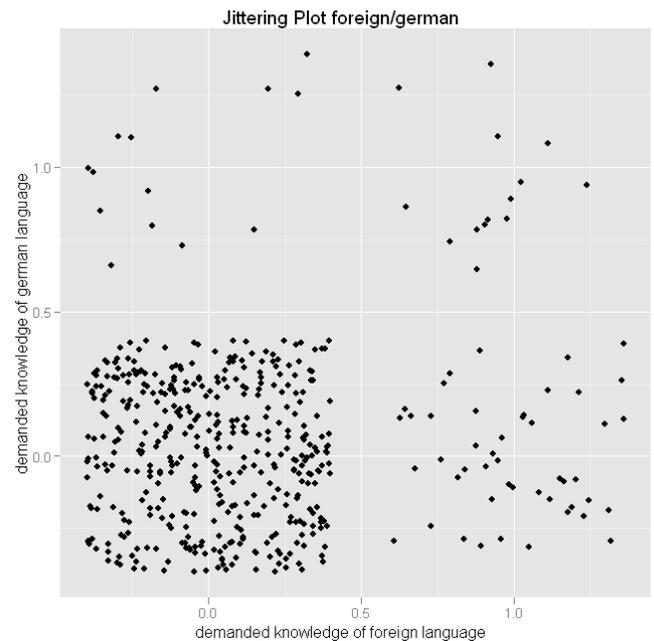

Fig.2. Jittering plot: foreign against German language

Thus, it is possible, that there exists a certain picture of a potential candidate. If a native German speaking candidate is expected by the writer of the job ad, he or she seems to implicitly assume an existing knowledge of the German language. The knowledge of the German language is pragmatically seen as a prerequisite for a successful graduation at the University of Business and Administration in Vienna due to the fact, that a majority of the courses are held in German.

In case of a shift of the picture of a potential candidate also the explicitly demanded skills seem to change. For example: a person who has to know an Eastern European language for a job in an Eastern European country could also be an inhabitant of this country. This inhabitant is with high probability not a native German speaking person. Thus, the demand for German speaking skills becomes an explicitly demanded skill for the potential candidate, because German now is a foreign language. He or she must have the ability to communicate with a partner enterprise in Austria, because these jobs are with high probability offered by companies with a relation to Austria.

It can also be seen, that other foreign languages have a very low frequency. These skills do not seem to be expected from a potential candidate.

In general, the language example leads to the assumption, that the explicitly demanded skills of job advertisements are not the entire range of skills which are in demand. There could exist a given picture of a graduate, but only a range of skills are explicitly stated, because some of the skills are implicitly assumed und others are not expected.

IV. METHODOLOGY

The study design was established under a stepwise approach. The phases are:

- Defining skill relevant words with high frequency
- Generating a binary skill related word/document matrix
- Finding a structure for content analysis
- Generating a binary skill/document matrix
- Statistical analysis

A. Defining skill relevant words with high frequency

With automatic text analysis high frequent words (frequency >=25) were stated. The skill/employability relevance of the words was checked. Analysis was done with the textmining package tm for the Cran R environment. [7]

B. Generating a binary skill related word/document matrix

A matrix in the following structure was automatically generated (n>1200; m>200).

TABLE II
SCHEMATIC TABLE OF THE DATA MATRIX

	skill related word 1	skill related word 2	…	skill related word m
job ad 1	0/1	0/1	0/1	0/1
job ad 2	0/1	0/1	0/1	0/1
…	0/1	0/1	0/1	0/1
job ad n	0/1	0/1	0/1	0/1

C. Finding a structure for content analysis

For content analysis a structure of demanded skills according to the skill relevant words, was established. This structure is according to [4] and [17]

The concept of skills divides them into cognitive skills (school disciplines like English, Mathematic…), technical skills (refer to specific skills needed in an occupation with potential reference to academic skills), soft skills (a construct of attitudes and behavior) and generic skills (broadly transferable skills like problem solving). [17]

Bennett [4] invented a framework for the development of generic skills, which gives a more detailed idea of the concept behind. Generic skills are subdivided into management of self, management of information, management of others and management of task. These dimensions were used for the textual concept analysis of the job advertisements.

D. Generating a binary skill/document matrix

According to the word/document matrix a skill/document matrix was established manually of the 484 finance related job advertisements via content analysis.

E. Statistical analysis

The skill/document matrix was analysed by total frequency, relative frequency by profession and several clustering methods to show different patterns of skills. As clustering methods hierarchical clustering (hclust) and a clustering approach due to Neural Networks Martinez [10] and

Martinez/Schulten [11] were used under the framework of Mazanec [12] and his software application.
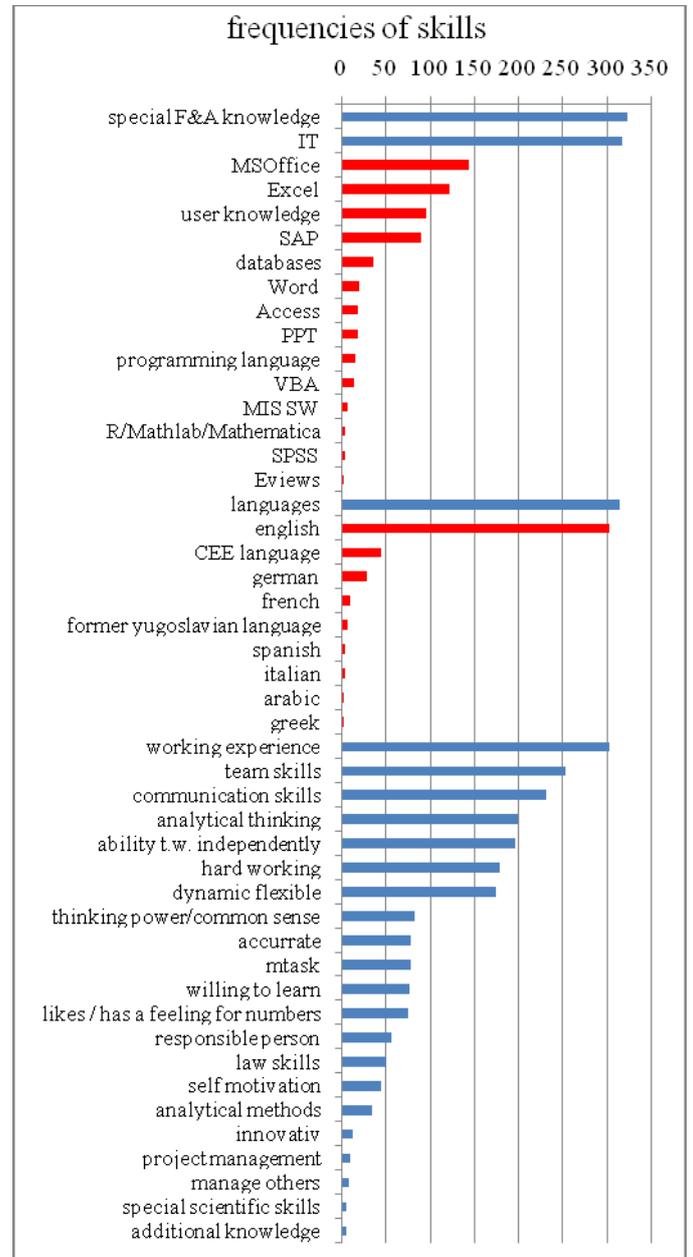
V. OUTCOME

A. Frequencies

Fig.3. frequencies of skills

It can be seen that special technical knowledge in finance and accounting is the item with the highest frequency, followed by IT skills and languages. As a consequence these factors seem to be most important for the decision process. Also a set of other factors seem to have a high impact on the recruiting decision. Working experience seems to be an important factor although it is a sample of job ads for young graduates. According to languages English is very dominant and the lingua franca phenomena seem to be true.

Nevertheless Eastern European languages (CEE languages) also play an important role in the Austrian job market.

B. Cluster Effects

Hierarchical clustering shows, that there seems to be the evidence that skills occur in certain clusters. The complete linkage method was used with correlation distances between the skill objects [1]. For graphical interpretation the cluster "cutting line" was drawn by 1-cor(skill[x];skill[y])=1,1. The following clusters occurred: a scientific skill cluster, an international management skill cluster, an IT skill cluster and a general management skill cluster. Thus, jobs in finance related fields seem to go into one of these directions.

Next to an hclust approach a partitioning clustering approach was used [12]. Learning vector quantization [10] & [11] and TRN32 (the software application behind [12]) was used. This method optimizes a predetermined number of initial data points to make them "good" prototypes of the cases in their respective clusters. [12], p. 220. The outcome of this approach under four clusters shows that in general the hierarchical cluster outcomes seem to be related to the hclust solutions. There seems to be one cluster with high values in languages (30.58%). Another cluster with languages, MSOffice (general IT) and high values of different work attitudes (for example: hard working, work independently, team skills) (19.83%). A third with high analytic skill values, calculating IT software and high communication and team values (14.88%). The forth has medium characteristics in several skill dimensions and seems to be a cluster for more general management jobs (34.71%).

So it can be concluded, that skills appear in certain clusters and curricula or learning strategies could be adapted to these combinations of skills.

VI. CONCLUSION

The paper has several aims. Firstly, the limitations and strengths of job advertisements as a data source are shown. Secondly, the analytical methodology for the purpose of analyzing job advertisements is illustrated. Thirdly, the outcome related to the Austrian data will be presented. The cluster analysis shows that a certain set of differently combined skills seems to be an indicator for the "fitness for purpose" in the job market. So curriculum planer and students should not just invent courses for certain skills but also focus on a balance of different demanded skills.

REFERENCES

[1] Anerberg, Michael R. (1973). Cluster Analysis for applications, Academic Press, 1973.
[2] Arcodia, C.V., Barker, T. (2003). The Employability Prospects of Graduates in Event Management: Using Data from Job Advertisements, *R. W. Braithwaite and R. L. Braithwaite Riding the Wave of Tourism and Leisure Management*, Coffs Harbour, 5-8 February, 2003.
[3] Bennett R. (2002): Employers´ Demand for Personal Transferable Skills in Graduates: a content analysis of 1000 job advertisements and an associated empirical study, Journal of Vocational Education and Training, Volume 54, Number 4; 2002; pp. 457-475.
[4] Bennett Neville et. Al. (1999): Patterns of core and generic skill provision in higher education, Higher Education; 1999; 37 pp. 71-93.
[5] Bourdieu, Pierre (1986): The Forms of Capital, J.G. Richardson's Handbook for Theory and Research for the Sociology of Education, (1986), pp.241-258.
[6] Clyde, Laurel A. (2002): An instructional role for librarians: an overview and content analysis of job advertisements, Australian Academic & Research Libraries; 2002; Volume 33 No. 3.
[7] Feinerer; Hornik and Meyer (2008): Text Mining Infrastructure in R, Journal of Statistical Software, March 2008, Volume 25, Issue 5.
[8] Fugate M.; Kinicki A.J. and Ashforth B.E. (2004): Employability: A psycho-social construct, its dimensions, and applications, Journal of Vocational Behaviour 65; 2004; pp. 14-38.
[9] Harvey, L. and Green D. (1993): Defining quality, Assesment and Evaluation in Higher Education; 1993; Vol. 18 No.1, pp. 9-34.
[10] Martinetz et. al. (1993): "Neural Gas" Network for Vector Quantization and its Application to Time-Series Prediction, IEEE Transactions on Neural Networks; 1993; 4(4) pp. 558-569.
[11] Martinetz Th. and Schulten K. (1994): Topology Representing Networks, Neural Networks; 1994; 7(5) pp. 507-522.
[12] Mazanec J. A. (1999): Simultaenous positioning and segmentation analysis with topologically ordered feature maps: a tour operator example, Journal of Retailing and Consumer Services; 1999; 6 pp. 219-235.
[13] Pember M. (2003): Content analysis of recordkeeping job advertisements in Western Australia: knowledge and skills required by employers, Australian Academic & Research Libraries; 2003; Volume 34, No.3.
[14] Rolfe, H. (2002): Students´demands and expectations in an age of reduced financial support: the perspectives of lecturers in four English universities, Journal of Higher Education Policy and Management; 2002; Vol. 24 No. 2, pp. 171-82
[15] Sirvanci M. (1996): Are Students the True Customers of Higher Education?, Quality Progress; October 1996; pp. 99-102.
[16] Srikanthan G. and Dalrymple J. (2003): Developing alternative perspectives for quality in higher education, The International Journal of Educational Management; 2003; 17,2/3, pp. 126-136.
[17] Stasz C. (2001): Assessing skills for work: two perspectives, Oxford Economic Papers; 2001; 3, pp. 385-405.
[18] Eagle, Lynne and Brennan, Ross (2007): Are students customers?, Quality assurance in education; 2007; Vol. 15, No.1, pp. 44-60.
[19] Todd, P.A.; McKeen, J.D. and Gallupe R.B. (1995) The Evolution of IS Job Skills: A Content Analysis of IS Job Advertisements from 1970 to 1990, MIS Quarterly;1995; Vol.19, No. 1, pp. 1-27.
[20] Brown, P. and Hesketh, A. (2004) The Mismanagement of Talent: Employability and Jobs in the Knowledge Economy, Oxford: Oxford University Press.