

DETECTING KEY WORDS FOR FUTURE BUSINESS SUCCESS – CASE STUDY IN EDUCATION

Peter Harmath¹⁶³

DOI: <https://doi.org/10.31410/limen.2018.510>

Abstract: *Organizations seek to maintain different channels of communication. Those emerged channels – medias of information flow after a short period of functioning becomes obsolete and unclear. The expressed wishes are losing their scope. Such communication problem can be found in the case of educational organizations. The aim of the study is to map the coincidence between different messages issued and the real work scope. The selected organization in this case study is Faculty of Economics in Subotica. Listed accessible documents, issued reports, vision and mission statements from the web site of Faculty were considered as a set of key words describing activity or goals of organization. Input categories for text analyses are defined in three sets. Mission vision statements with introductory words of Educational Departments – first set, selected course descriptions – as second set, permitted Bachelor themes – as a third set. Searching for most frequently used terms – words form listed texts revealed some misfits between aims of institution and issued messages. The results are showing lot of missing terms as key words for promoting mission and work scope in future. Applied methodology of the text analysis has given very useful information as a base to compare the key words (the most commonly used terms) and to detect their possible connectedness with different sets of other documentation.*

Keywords: *Text analysis, management and business, key word selection, source document detection.*

1. INTRODUCTION

Organizations seek to maintain different channels of communication. Those emerged channels – medias of information flow after a short period of functioning becomes obsolete and unclear. Different message contains different directions, frequency circulate between the actors, the stakeholders of any organization. Many of those messages are losing track, e.g. vision and mission statements are created - made in every organization, but quite often they do not meet the needs. Furthermore, as part of formalization different documents, standards, roles, job descriptions are created. It is part of management efforts to support coordination, control as well as it becomes part of public relation (PR), promotion and other marketing communication element. From the bottom to the top level of general business strategy, messages are part of organizational culture. [1]

Information is one of the most important resources in the contemporary business environment. It's hard for any company to succeed without issued information about its products or services employees, and other key properties. Every day, companies issue unstructured and structured text to various sources such as web presentation, online services, reviews, emails, and other documents. These sources provide text, which is not easy to analyze without using the right text analysis tool. It's possible to perform text analytics manually, but the manual process is

¹⁶³ University of Novi Sad, Faculty of Economics in Subotica, Segedinski put 9-11, 24000 Subotica, Serbia

ineffective. In the study the so called “Traditional keywords system was applied. As a matter of fact, it is unable to read and understand language in web pages, and text documents. There was no need for that at this stage performed in Study. The text analytics software (Online-Utility.org) [2] for Online Operating System was applied to preliminary detect and count frequency of some common expressions or key words in large volumes of text data. (However, the used free online software supported different languages; the most of explored texts were written in English – [3]).

2. METHODOLOGY

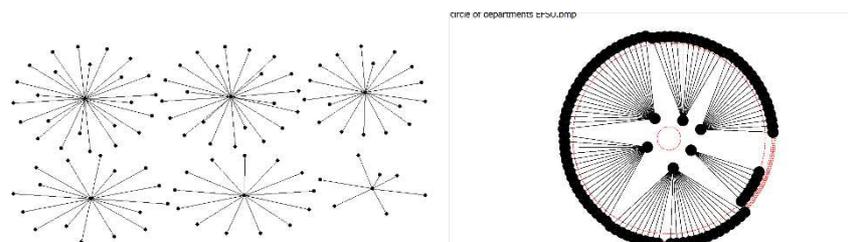
This approach is simplified to basic text analysis. It is not performed with more advanced text mining techniques as a process of exploring and analyzing large amounts of unstructured text data aided by software that can identify concepts, patterns, topics, keywords and other attributes in the data. The applied software belongs to free utility online service. [2] It collected expressions, and most often appearing key words in different documents.

The observed institution, Faculty of Economics in Subotica (In further text Faculty [4]) is organized into Departments and Consulting centers connected with scientific and educational activity. These parts are potential sources of general data, curricula data and can be listed as:

- Department for Agricultural Economics and Agribusiness
- Department of Business Informatics and Quantitative Methods
- Department of European Economics and Business
- Department of Finance, Banking, Accounting and Auditing
- Department of Management
- Department of Trade, Marketing and Logistics and Consulting Centers
- Centre for National and International Projects (CNIP)
- Educational Centre
- Foreign Language Centre
- Sports and Recreation Centre (SRCEF)
- The Information and Documentation center (INDOK)

Figure 1 shows simplified organizational structure with Departments (the sources of data) and teaching staff (somewhat corrected by author according to the considered - analyzed period). This period can be characterized with fairly intense fluctuation in staff. Basically, retirement of full professors and some younger colleges have been promoted to new positions (e.g. for department head, or entered other comities of Faculty, or was appointed to the University management level).

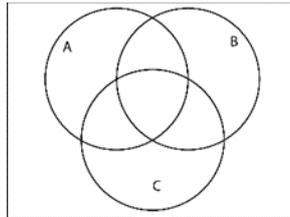
Figure 1: Authors compilation of organizational structure: departments with members – stuff.
Source [5], [6]



The study is based on filtering and comparing some of the key terms connected with three group of documents issued by Faculty of Economics in Subotica (Faculty). In order to narrow the scope, the chosen key terms are about mission and vision statements, curriculums, and themes of Bachelor graduate works. All this material is public and available on web site of the Faculty. Undergraduate academic studies last for four years, i.e. eight semesters, convey 240 ECTS points.

Input categories as different texts will be defined in three sets (set - A, - B, - C.) (see figure 2).

Figure 2: Three sets of documents as three categories to be analyzed



The first set is based on texts from Introductory statements of Faculty and general information on Departments (this set or category could cover additionally contains of any kind of promotional written reports, leaflets, etc.). As a matter of fact, it is representing the wide public information category.

The second set is based on texts from Accreditation documentation of courses and educational programs of Bachelor Studies (available in English selected by criteria of key words management, business, marketing, organization, entrepreneurship in the course title). This list is more professionally oriented. The following courses were taken into consideration: Farm Management, Management of Manufacturing Organic Foods, Marketing Management, Principles of Marketing, Services Marketing, Entrepreneurship, Human Resource Management, Organizational Behavior, Organizational Theory, Principles of Management, Project Management, Business Intelligence Systems, Electronic Business, International Business, Marketing and Management of Trade Companies.

The third set is based on, or includes, the key words extracted from titles of themes - graduation work themes of level – Bachelor degree. (Of course, for wider analysis, the list can be extended with titles of Master and Doctoral (PhD) works; Note: In this study the Master and PhD level is not analyzed). As a matter of fact, the third set is showing key words of what is “realized” the most, e.g. business, management, leadership problems, etc. It can be considered or it is part of the final “product” of Faculty. It is part of intellectual property of students gained through four-year study period.

To prepare the report review and for visualization the Venn diagram – (Gene List Venn Diagram) solver tool was applied. [7] By entering texts either the "text area" or "file upload" the sets can be assigned into the diagram.

3. RESULTS

The first set: Introductory text: Faculty mission and vision statements, Introduction of educational Departments. [4]

Table 1 shows size properties of text about introductory statements.

Table 1: Size properties of text about introductory statements [2]

Number of characters (including spaces):	8018
Number of characters (without spaces):	6714
Number of words:	1047
Lexical Density:	33.7154
Number of sentences:	40
Number of syllables:	2312

The texts in category of set A are filtered for phrases in length of 6 words. The following was found (see Table 2).

Table 2: Top phrases between Introductory statements of Faculty and Departments [2]

Some top phrases containing 6 words (without punctuation marks)	Occurrences (frequency)
students of this profile may find	6
of this profile may find employment	6
this profile may find employment in	4
curriculum provides students with specific fundamental	2
may find employment in a variety	2
profile may find employment in a	2
find employment in a variety of	2
specialized knowledge in the field of	2
this curriculum provides students with specific business	2
this curriculum enables students to	2
this profile may find employment as	2

The second set: Accreditation documents of courses in English (selected in this study)

The third set: Permitted bachelor themes.

Table 3 is showing text size properties:

Table 3. Size properties of text about permitted bachelor themes [2]

Number of characters (including spaces):	286496
Number of characters (without spaces):	201319
Number of words:	34300
Lexical Density:	11.1370
Number of sentences:	5412
Number of syllables:	77907

Table 3: Most popular themes and study programs [2]

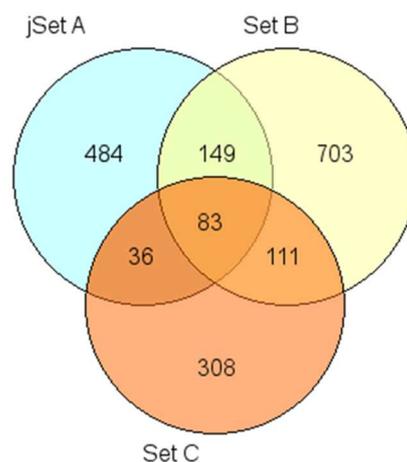
Some top phrases containing 6 words (without punctuation marks)	Occurrences (frequency)
finance program banking and insurance theme	235
study program finance banking and insurance	235
finance banking and insurance topics	232
banking and insurance theme of graduate thesis	232
study program marketing topic of graduate thesis	220

economics and business theme of graduate work	113
study program accounting and review topic	102
Accounting and Auditing Program	101
accounting and revision of the topic of graduate work	101
economics and agribusiness topic of graduate work	97
study program agrarian economy and agribusiness	96
program agrarian economy and agribusiness theme	96
agrarian economy and agribusiness theme of graduate	96
study program of European economy and business	81
European Economy and Business Theme Program	81
European economy and business theme of graduate	79
study program on the topic of graduate thesis	76
study program business information systems themes	76
study program for trade in the topic of graduate thesis	74
business information system programs	71
business information systems topics of graduate thesis	71

In order to simplify this rather big text file, words that appear at least two times have been considered in further analyses. According to the aim it was filtered manually with selected management and business specific terminology.

“The visualization is designed to be useful for the researcher to determine the number of publications that fall into each subcategory of research. The usefulness is dependent on how the searcher utilizes it.” [8]. For visualizing the results of key word selection figure 3 is showing the three Venn diagrams with chosen terms. For further analysis the common words between categories are presented (except intersection between set A and B).

Figure 3: Three sets with selected words from three source text categories Source: GeneVenn diagram [7]



Terms in intersections of sets A, B, C [7] - the vocabulary of real key words

According to the common elements (words) the next intersections were analyzed:

Found in gene lists one and three (Set A and set C) = 36

Some of the terms (selected by author) Genes' names:

European, analytical, audits, banks, bookkeeping, co-operation, compliance, component, corporate, direct, economies, employment, exchange, financial, firms, globalization, growth, intelligence, logistics, method, methodology, money, partnerships, prices, public, regional, relationships, rural, sale, selected, statistical, stock, sustainable, tax, time.

This intersection represents terms that are “mentioned” in mission and vision statements of Faculty. It is the list of common word that are directed to future students. For sure there are missing some terms that fulfil modern expectations of future managers, entrepreneurs and leaders - e.g. we suggest mentioning as part of set A (the introductory statements for future students, a vision and mission statements) of: leadership, networking, start up, entrepreneurship.

Found in gene lists two and three (Set B and set C) = 111

Some of the terms (selected by author) Genes' names:

adoption, agribusiness, agriculture, assets, assortment, attitudes, behavior, budget, case, causes, challenges, changes, channels, characteristics, commercial, communications, company, comparative, components, compulsory, conflicts, consequences, consumer, consumers, count, course, creation, credit, culture. decision, demand, digital, diversification, dynamics, economy, electronic, elements, employees, enterprises, entrepreneurship, ethics, example, external, farm, financing, general, genesis, global, goods, group, health, improvement, incentives, internal, internationalization, interpretation, investment, legislation, levels, line, location, means, measures, mix, motivation, motives, organic, organization, payment, personality, plan, platforms, policy, politics, potential, power, practice, practices, problems, project, promotion, protection, purchase, requires, responsibility, results, review, revision, risks, satisfaction, security, selection, service, social, society, strategy, stress, structure, support, tasks, technology, tourism, tracking, transport, using, values, word.

This list is much more diverse and richer list in sense of necessary terms needed to be mentioned in modern communication channels. It is recommended to be adapted for widening the set A (Introductory statements of Mission, Vision and Departments).

Found in all three gene lists = 83

Some of the terms (selected by author) Genes' names:

accounting, activities, agricultural, analysis, approach, banking, basic, between, business, communication, companies, competitive, concept, contemporary, design, designing, development, distribution, economic, education, environment, finance, food, human, implementation, industry, information, innovation, insurance, integration, international, introduction, leadership, making, management, market, marketing, markets, offers, organizations, orientation, performance, planning, principles, problem, process, production, products, projects, promotion, quality, research, resource, role, sales, services, skills, software, solutions, solving, staffing, standards, strategic, strategies, system, team, techniques, theory, trade, training, use, work.

This list is the summary of terms found in lists. However, it is result of selection it represents objectively the “common” area.

4. DISCUSSION

Based on the findings, in first set of data we have determined that key words are oriented to encouragement of future employment of our students. It is worth to consider more in future to emphasize future potential of self-employment, encouraging entrepreneurship so not to be just employee but employer creating and maintaining work. Terms like leadership, start up, accelerate, network, seems to be missing from these introductory messages.

The second set is based on terminology connected with narrower expert knowledge of selected subjects. Because of thematic selection of terms in this study some key words may be lost. But the scope of exploration is basically focused on connection between issuing actor, department and the used “vocabulary” so it can be expected to be found in narrowed list. That is the main reason why the intersection is not presented in details. It can be considered as part of internal communication in organization (Faculty), not showing the effects on student’s choice directly. The interpretation of this set should include more sophisticated selection of key words before visualization. It was not completed for this section.

Third set represents a really huge collection of terms. It somewhat referred the “diffusion” “penetration” of subjects – courses to the student’s way of thinking and areas of interests. It is very interesting to read the common terms on the level of intersection of all three sets and finding the key term as a holder of subject area. Also, the list of the unique terms that are not common for, or in other words, not promoted in vision and mission statements nor in accreditation of courses but used by students, is part of misconnected area of interest? Or just the influence of charismatic performer - mentor?

According to this initialized study, there are many further questions to be explored - e.g. What are the hidden connections that make certain differences between statements and real achievements? What are the sources of weighted influence on results? Solving these questions may need including methodology of Social Network Analysis (SNA) [9], or more narrowed option - the methodology of Organizational Network Analysis (ONA) [9]. The question is connected with different measures such as centrality and prestige, betweenness and brokerage, cauterization, centrality, etc.

5. CONCLUSIONS

In each institution regardless of the ownership form, the founder, the field of education or research the mission, vision, introductory, etc. - statements are created and promoted, launched in all means, obtainable for communication inside and outside of organizations.

Searching for key words and terms with preselected lists may cause data to be distorted, especially if there are many questions of similarity. Some of the key terms used in professional sense became common in everyday conversations. When creating new messages for future economists and entrepreneurs (?) it is becoming crucial question. The influence of this circumstance is day to day more and more perceptible. Considering the results, the most presented key words, as a student’s selection for bachelor themes, there is a certain influence of attractiveness of most popular subjects and mentors. That can cause mass interest and orientation toward relatively limited number of options. According to studied key words of themes there are only few really popular areas for students. If we look for key words associated with those favorite fields between promotional texts we can hardly find enough as encouragement. It is warning sign for future to think over introductory statements with aim to

develop more “fashionable” or more attractive invitational messages. For sure, this question is not just about selecting the most popular terms. It is connected with scientific potential of Faculty staff too.

Regardless of potentials in different software support [10], thorough knowledge and creativity is needed to adequately define criteria for selection of key words as well as to reveal the meaning of professional terminology from different fields. Presented analysis is simplified according to the limited list of entered documents to text analyzer tool. It is considered to be enough to present some applicable methods for quick and effective analyses. Efficiency depends on deliberately selected terms for analyses or their highlighted order.

Missing but have to be stressed

The issued messages do not involve for every department of the Faculty adequate presentation, meantime the emphasized and wished messages are losing meaning. Obviously, the introductory statements of vision and mission do not meet up to date expectations and professional work in Faculty. Modestly represents the real professional work (according to much wider connections between e.g. professional messages – accreditation documentation). The introductory statements and other supporting documents of promotion of Faculty should be rearranged with added new terms of renewed study areas.

The extension of this study is planned to be added in the near future. It is the analysis of connection between mentors and members of graduation committee to map the collaboration between course holders inside and between different departments – understood as different study areas internal interdepartmental coworking. This kind of study may give on information about the members of collective that are in position of networking, in positions called brokers or influencers.

The result obtained from text analysis tools are really applicable, but only for experts – involved stuff of the examined field. The usability of results may be depending on subject, range and differences. But for sure, finding out some key expressions key words offers a fast glimpse into hidden relations and potentials in organizational communication and formalization. If applicable, there are potentials for comparison for narrowed research.

REFERENCES

- [1] Ahmetagic E, Harmath P., (2009), Projektovanje organizacije, Ekonomski fakultet u Subotici, Subotica
- [2] Online-Utility.org - Utilities for Online text analysis tool from free online access, <https://www.online-utility.org/text/analyzer.jsp>, accessed 25.10.2018.
- [3] Web page of Faculty of Economics in Subotica, <http://www.ef.uns.ac.rs/english/courses-in-english.htm>, accessed 27.10.2018.
- [4] Faculty, homepage of Faculty of Economics in Subotica, www.ef.uns.ac.rs accessed 27.10.2018.
- [5] Pajek 2.04 Vladimir Batagelj and Andrej Mrvar, Software Free for noncommercial use, <http://pajek.imfm.si/>, accessed 10.10.2018.
- [6] SocNetV 2.3, (2017), by Dimitris V. Kalamaras, <http://socnetv.org> accessed 10.10.2018. SocNetV 2.3, Manual, (2017), by Dimitris V. Kalamaras, <http://socnetv.org>
- [7] Core Graphic Module by Vijayaraj Nagarajan, and Web implementation by Mehdi Pirooznia. October 2006, usm.edu <http://genevenn.sourceforge.net/index.htm> 27.10.2018

- [8] Burris John W, (2015) Visualizing Keyword Searches Using the Venn Diagram, Journal of Advances in Information Technology Vol. 6, No. 2, May
- [9] Cross R., Parker A., (2004), The Hidden Power of Social Networks: Understanding How Work Really Gets Done in Organizations. Harvard Business Review Press; 4.1
- [10] Nguyen T.N., Zhang, J. (2006)., "A novel visualization model for web search results," IEEE Transactions on Visualization and Computer Graphics, vol. 12, no, 5, pp. 981,988,