Analyzing Web Presence of Russian Universities in a Scientometric Context

Anastasiya Kuznetsova(✉), Stanislav Pozdniakov, and Ilya Musabirov

National Research University Higher School of Economics, Saint-Petersburg, Russia
adkuznetsova13@gmail.com, pozdniakov.stanislav@gmail.com,
ilya@musabirov.info

Abstract. In this paper, we analyse the strategies and stratification of Russian universities in the Northwestern region. By enriching traditional social network analysis scientometric tools, we developed web presence indicators focused on the contexts in which universities are linked with businesses and are mentioned in media. We treat resulting groups in terms of Gouldner's cosmopolitans versus locals theory, based on differences in their publication strategies, and embeddedness in business connections and media contexts.

Keywords: Russian universities · Northwestern region · Webometrics · Altmetrics · LDA · Topic modelling

1 Introduction and Related Work

One of the most important factors for university performance assessment is their publication activity. Universities in Russia are not exceptions from this trend. Because of this "publish or perish" mentality, a highly competitive environment appeared [1], and Russian universities are stratified in their publication activity. Researchers can target their publications either to an international audience of Web of Science (WoS) and Scopus-indexed journals; high-profile Russian journals, which are included in newly developed Russian Science Citation Index (RSCI), or less internationally oriented Russian journals. The majority of Russian journals are excluded from international science communication [3]. However, more and more Russian papers are indexed in the Web of Science Core Collection [8].

The aim of our research is to define strategies, that universities use to increase performance indicators. The focus is on the web data extracted from companies and news websites, which show the business and media context. In addition, we consider data about publications and universities domain-related data. We suggest that additional data from various sources is valuable for analysing and understanding universities agencies.

The difference in strategies allows universities to be considered within the locals and cosmopolitans theory. Gouldner applied this theory to academics, where locals are the group of individuals who interact within this group and use

D.A. Alexandrov et al. (Eds.): DTGS 2017, CCIS 745, pp. 113–119, 2017.
https://doi.org/10.1007/978-3-319-69784-0_9

interpersonal channels to spread the information. Oppositely, cosmopolitans are connected with several communities; consequently, they have weaker connections within these groups and use organizational and head channels for spreading information [5]. In Puttens study [10], universities faculties are considered to follow one of the roles of the introduced theory (cosmopolitan or local). Moreover, in [11] the importance of different orientations for scientists as locals and cosmopolitans is revealed. The concept of productivity is divided into scientific productivity, such as the sum of journal publications and presentations and organizational productivity: laboratory technical reports and memoranda. Cosmopolitan orientation is related primarily to the scientific productivity dimension, while locals orientation is positively associated with organizational productivity.

2 Data and Method

One of the common ways of university research assessment is scientometrics, which applies quantitative analysis to the following aspects of publications: citation indexes, authors, subjects and sources. More recent advances in the field have built on social network analysis methods, including analysing co-authorship and bibliographical coupling data from the network perspective, which is targeted to find structurally important authors, works, organizations and publication venues in their interconnections.

An alternative approach to assess scientific activity and impact is webometrics or cybermetrics [7].

Webometrics uses a toolkit of methods and approaches to judge a publications' web presence, including web impact factor (WIF) - the normalized number of unique in-links to the domain [6]. Researchers have demonstrated that the total number of links on a page of the website and the average number of links on a page could be indicators of the value and importance of the information on these pages. According to Thelwall [12], WIF correlates positively with the offline research activity of the university. The Webometrics Ranking of World Universities Project is the largest academic rating of higher education institutions and it is based on webometrics - the amount of web content (the number of pages and files) and the apparent impact of publications on the number of citations in social media.

However, the choice of ranking approach, while important for applications, limits the understanding of the structure and the strategies employed by the agents. Moreover, it influences the agents' desire to manipulate the system in various ways [4].

While web ranking methodologies pay considerable attention to universities' domains, we suggest that important information about universities can be gathered from their structural connections with other types of agents, including industrial companies and media.

We used the case of 51 universities from the Northwestern region of Russia, focusing on their research in Economics and Management. We focus on distinguishing agents' strategies based on the following metrics:

– The number of unique incoming links to the university's domain.
– Representation in major regional media – the frequency of mentioning a single university on media sites.
– Representation on the sites of business companies: the frequency of mentioning a university on the websites of Top-50 companies of the Northwestern Federal District according to the Expert magazine ranking.

To understand whether Northwestern universities prefer to be published in international or Russian journals, we extracted the number of RSCI and WoS publications in the fields of economic, business, management, and finance. We use hierarchical clustering with complete linkage method to reveal stable groups of universities characterized based on a number of WoS/RSCI indexed publications and on the number of unique external links to the university web domain.

To analyse the context, in which universities activity was mentioned on media and companies websites, we used the latent Dirichlet allocation algorithm (LDA) [2] with a 50-topic model.

During the data gathering process, it became clear, that some institutions are frequently mentioned on media purely in the negative context, so thematic context was considered in our analysis. Using Pearsons residuals we chose several prominent topics strongly associated with the universities in our sample.

3 Analysis and Results

Figure 1 shows that Northwestern region universities are differentiated in publication activity strategies, since they have different proportions of publications in Web of Science Core and RSCI journals.

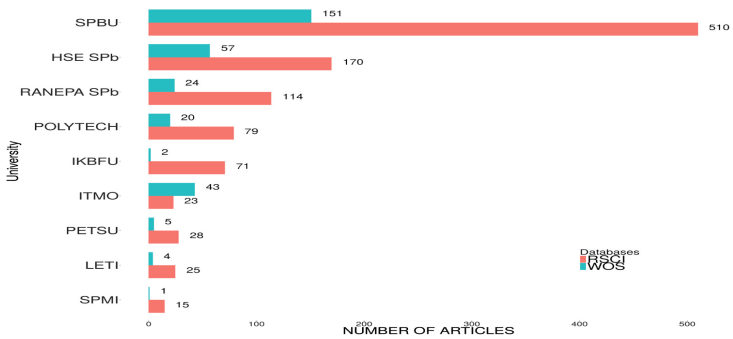

Fig. 1. The number of economics and management publications in WoS and RSCI, Russian Northwestern Universities

Hierarchical clustering of webometrics indicators reveals three groups of universities (Fig. 2). For RANEPA and HSE campuses in Saint-Petersburg, indicators were multiplied by the percentage of researchers in these campuses (0.294

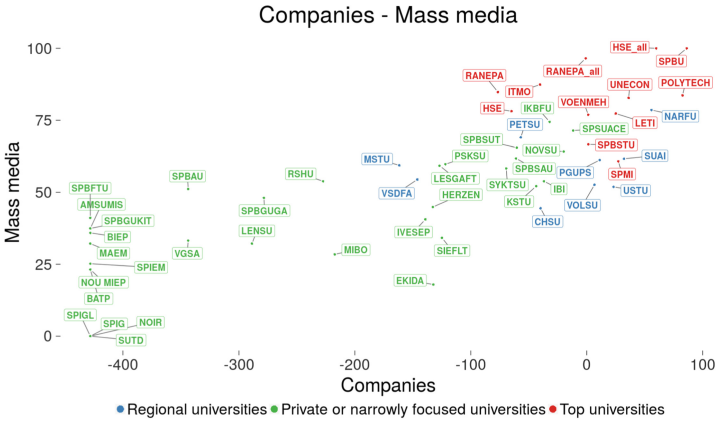

Fig. 2. Map of Russian Northwestern universities based on normalized frequency of mentioning on media and companies websites (Color figure online)

and 0.133 respectively). In addition, non-corrected indicators for these institutions as a whole are marked "HSE_all", and "RANEPA_all" on Fig. 2.

Some universities appear in news more often than others, simultaneously, being out of the of top regional companies attention at all (green cluster). This could be a sign of a low level of business connections and a lack of collaboration with these universities, for example, they could have a lack of collaborative educational programmes and internships.

The red cluster consists of both technical and economic universities occupying top positions in the region. The most popular and presented in media universities are well-known, highly reputed ones e.g. HSE University, St. Petersburg State University, The Russian Presidential Academy of National Economy and Public Administration (RANEPA), St. Petersburg State University of Economics (UNECON), ITMO University.

The blue cluster includes narrowly focused regional universities, which are connected with local topics. They are also more visible in RSCI than in Scopus and WoS.

The green cluster is occupied by narrowly specialized small universities, which are not highly represented in scientometric databases and they strongly differentiate on Media-Company coordinates. Generally, they are less frequently discussed in media than universities in red and blue clusters; however, some of them (Saint-Petersburg State University of Architecture and Civil Engineering, Immanuel Kant Baltic Federal University, and Novgorod State University - Fig. 2, right side) indicate a high level of embeddedness within the industrial sector. Meanwhile, another group of institutions (the Baltic University of Ecology, Politics, and Law; Murmansk Academy of Economics and Management; and St. Petersburg Institute of Economics and management - Fig. 2, left side) demonstrates no visible web traces of collaborations with top companies.

Prominent topics associated with universities from red and blue clusters are shown in Table 1.

Universities from the red cluster with specialization in social science and economics are related to the topics about industry, innovations, science, research, business executives, and education. They are treated as makers and producers of scientific and innovative knowledge [9]. They are also mentioned in the context of business, finance, investments and entrepreneurship. This indicates that they have close ties with the business sector through consulting, business incubators and other collaborative initiatives.

Table 1. Results of LDA applied to media texts

Group of universities	Topics	Context
Top universities with specialization in social sciences and economics red cluster	monitoring_education, directors_researchers, research (social sciences), economics, prices, finances	"this data was presented in the report of the Director of the Institute of Social Analysis and Forecasting (RANEPA) and professor of HSE Tatiana Maleva "Man in the solidarity pension system" "issues under monitoring openness of Russian universities for applicants" "The second lecture in the framework of the HSE and Jaguar Land Rover was devoted to the future of the banking sector"
St. Petersburg universities with specialization in STEM disciplines red and blue clusters	industry, innovations, research2 (science)	"Chemists from ITMO University have developed a colorless ink for color printing" "boosting the competitiveness of innovative industries by establishing strong out-sectoral and in-sectoral linkages and the general infrastructure"
Regional universities or narrowly focused universities blue cluster	local, arctics, transport	"The core of this long-standing debates is whether the butter "Vologda" is the unique product of Vologda region..." "Research Expedition "Floating University" started from Arkhangelsk and is going to the Arctic on the board of the research vessel"

The universities with STEM specialization are unsurprisingly related to research in science, industry and innovations. These universities are related to the topics about technological research and developments in this area.

Regional or narrowly specialized universities from the blue cluster are connected with local topics. For example, Arctic University (NARFU) was mentioned in the context of the Arctic Research Programme.

We found no prominent topics for universities from the green cluster.

4 Conclusion

In this paper, we enriched the scientometric analysis of university research activity with the set of webometric indicators. In addition, we investigated at the context in which universities are mentioned in media. Northwestern universities are stratified across different dimensions. Based on the connections with companies and publication strategies we revealed several groups of universities.

The cluster of top universities invests in high-profile publication ties development. They are linked in business and media as expert centers, business education providers and innovation leaders, with a saliency in LDA topics on research, innovations, business and industry development.

The second group is characterized by more narrowly focused universities with an advanced level of publication activity and performance in the region. They demonstrate the same exposure in the media and collaborations with top companies.

The third cluster consists of universities with low publication activity in economic and management. These universities may be oriented to satisfy the needs of regions and smaller regional companies, and they tend to be out of the attention of the media. Instead, they focus on building strong connections with particular industrial partners.

Acknowledgements. The article was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2017 2018 (grant No. 17-05-0024) and by the Russian Academic Excellence Project "5-100".

We are grateful to Viktor Karepin for his help with this research.

References

1. Altbach, P.: Anarchy and exploitation in scientific communication - university world news, March 2017. http://www.universityworldnews.com/article.php?story=20170328140116938
2. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. **3**(Jan), 993–1022 (2003)
3. Dyachenko, E., Fursov, K.: Russian scholarly journals in science communication. High. Educ. Russia Beyond **11**, 7–9 (2017)
4. Espeland, W.N., Sauder, M.: Rankings and reactivity: how public measures recreate social worlds 1. Am. J. Sociol. **113**(1), 1–40 (2007)
5. Gouldner, A.W.: Cosmopolitans and locals: toward an analysis of latent social roles. I. Adm. Sci. Q. **2**, 281–306 (1957)
6. Ingwersen, P.: The calculation of web impact factors. J. Doc. **54**(2), 236–243 (1998)
7. Ingwersen, P., Björneborn, L.: Methodological issues of webometric studies. In: Moed, H.F., Glänzel, W., Schmoch, U. (eds.) Handbook of Quantitative Science and Technology Research, pp. 339–369. Springer, Dordrecht (2004). doi:10.1007/1-4020-2755-9_16
8. Kasyanov, P.: Russian academic publishing landscape. High. Educ. Russia Beyond **11**, 9–11 (2017)

9. Kim, E.H., Zhu, M.: Universities as firms: the case of us overseas programs. In: American Universities in a Global Market, pp. 163–201. University of Chicago Press (2010)
10. Putten, J.V.: Cosmopolitan and local faculty orientations: a reanalysis of historical data. Paper presented in the Open Track at the EAIR 36th Annual Forum in Essen, Germany (2014)
11. Stahl, M.J., McNichols, C.W., Manley, T.R.: Cosmopolitan-local orientations as predictors of scientific productivity, organizational productivity, and job satisfaction for scientists and engineers. IEEE Trans. Eng. Manag. **2**, 39–43 (1979)
12. Thelwall, M., Harries, G.: The connection between the research of a university and counts of links to its web pages: an investigation based upon a classification of the relationships of pages to the research of the host university. J. Assoc. Inf. Sci. Technol. **54**(7), 594–602 (2003)