WHY EMPLOYEES LEAVE RUSSIAN COMPANIES? ANALYZING ONLINE JOB REVIEWS USING TEXT MINING

D. N. SOKOLOV, L. V. SELIVANOVSKIKH, E. K. ZAVYALOVA, M. O. LATUKHA
Graduate School of Management, Saint Petersburg University, Russia

In this study we analyze topics and sentiments of online job reviews for 989 organizations operating across 12 different knowledge-intensive industries in Russia. Using text mining techniques, such as topic modeling and sentiment analysis, we identify factors of job satisfaction and examine how they differ for former and current employees of Russian organization. The analysis reveals that (1) working arrangements and schedule, (2) working conditions, (3) job content, (4) salary/wage, (5) career development, (6) psychological climate and interpersonal relations with co-workers are the six key topics discussed by employees online in relation to job satisfaction, with the latter — psychological climate and interpersonal relations — being the most widely discussed topic, especially for current employees. Overall, our study suggests that in their decision to leave the company, employees are more likely to tolerate economic factors of job satisfaction (such as salary, career development and working arrangements) rather than socioemotional factors (such as poor relationships with their co-workers and content of work).

Keywords: employee turnover, job satisfaction, text mining, topic modeling, sentiment analysis, Russia.

JEL: M51.

Increased international competition, intensified globalization forces, and advancements in information technology, as the main drivers forming the contemporary business landscape and labour market, lead to significant changes in the psychological contract between the employees and the employers. The former avoid seeking life-long employment and job security in return for commitment, dedication, and loyalty [Al Ariss, Sidani, 2016]. Employers in turn continue strengthening work requirements and encouraging their workers to "go the extra mile" for improved performance. Increased job demands as well as the workers' belief that their employer has very high expectations about

their performance substantially enhance overall employee turnover rates in modern companies [Jaramillo, Mulki, Boles, 2013].

Depending on the managerial practices implemented by middle/top management to reach certain corporate objectives as well as the general employee attitudes towards these practices, the employee turnover rate can be skewed either toward the left (i. e. a relatively high percentage of employees leave an organization after a certain period of time) or toward the right (i. e. a relatively small percentage of employees leave as, evidently, the positive aspects of a job overcome the negative ones). Here, most scholars would agree that higher turnover, ceteris paribus, is worse than lower turnover due its deleterious effects on organizational results [Shaw et al., 2005]. Among the negative consequences are the reduced morale among the remaining organizational members, high replacement costs, reduced effectiveness and efficiency in serving existing prospects, critical financial performance consequences, and loss of potential competitive advantage [Katsikea, Theodosiou, Morgan, 2015; Sager, Varadarajan, Futrell, 1988]. A small group of authors, however, believes that up to a certain point employee turnover may provide positive benefits to an organization. For instance, [Meier, Hicklin, 2007] test the provocative proposition that moderate levels of turnover may positively affect organizational performance using data from several hundred public organizations over a nine-year period. Their findings, in particular, indicate a nonlinear relationship for a secondary output characterized by greater task difficulty. So, exploring different aspects of employee turnover and, specifically, understanding the potential antecedents of the intentions to quit in a specific setting are highly relevant research objectives for scholars and practitioners.

The key issue, however, is that pertinent literature devotes limited amount of attention to the investigation and analysis of the factors that influence job-related attitudes and actual turnover behavior. In particular, literature struggles to provide practical recommendations on how job satisfaction, a complex multidimensional concept comprising salary, supervision, promotion opportunities, psychological climate and other job-related dimensions, and subsequent turnover can be managed. Job satisfaction, being the strongest predictor of employee withdrawal decisions and the key answer to the question "why employees leave?" provided by extant research [Griffeth, Hom, Gaertner, 2000; Hom, Griffeth, 1995; Li et al., 2016], is often measured as one single, relatively general construct. As a result, scholars rarely distinguish between different factors of job satisfaction and their role in employee turnover. Prior research also has a number of important methodological limitations. Specifically, the dominant approach to measuring such constructs as employee turnover and job satisfaction is based on surveys/questionnaires as key data collection instruments. However, as the decision regarding a worker's intention to stay at a particular company or to leave for another tends to be personal, employees/respondents may reluctantly share information with third parties. Moreover, the responses are often affected (if not determined) by the research design set by the researcher. Consequently, when filling out the questionnaire form (i. e. choosing from a pre-determined list of answers) the employees may oversimplify their stories and hide important details regarding the job as well as their real attitudes towards the company.

In this paper, we try to overcome the aforementioned limitations by applying powerful text mining techniques to unique empirical data of user-generated online job reviews. The textual job reviews contain valuable but highly unstructured information regarding different aspects of work in certain organization. In our analysis we particularly aim at identifying the key topics (factors of (dis)satisfaction) discussed in online job reviews and examining how they differ for former and current employees (i. e. whether they led to a withdrawal decision). By doing so, we find a way to listen to and

comprehensively analyze the real voice of employees without pre-determined questions and answers.

Theoretical underpinnings

Employee turnover and job satisfaction

According to [Katsikea, Theodosiou, Morgan 2015, p. 368], "...it is incongruous to suggest that still little insight is available into why people quit their jobs [Lee et al., 2004] ...given the ubiquity of empirical studies devoted to employee turnover in the literature". In particular, for over one hundred years, scholars, mostly in the fields of personnel and organizational psychology, organizational behavior and human relations, have been investigating various issues connected to the antecedents and correlates [Cotton, Tuttle, 1986; Griffeth, Hom, Gaertner, 2000] and consequences of employee turnover [Hancock et al., 2013]. [Mobley, 1977], for instance, investigates the satisfaction-turnover relationship and suggests a heuristic model of the employee turnover decision process, which includes the following psychological and behavioral stages: (a) evaluation of existing job; (b) experienced job (dis)satisfaction, leading to alternative forms of withdrawal (e. g., absenteeism); (c) thinking of quitting; (d) evaluation of expected utility of search and cost of quitting; (e) intention of search for alternatives; (f) search for alternatives; (g) evaluation of alternatives; (h) comparison of alternatives vs. present job; (i) intention to quit/stay; (j) quit/stay. Meanwhile, [Morrow, McElroy, 2007; Shaw, 2011] examine the turnover-performance relationship, with the former focusing on a multi-unit commercial bank setting operating in the United States and the latter gathering substantial evidence indicating turnover rates have negative implications for different dimensions of organizational performance, like safety, productivity and monetary. Similarly, [Ton, Huckman, 2008] examine the influence of employee turnover on operating performance in settings that require high levels of knowl-

edge exploitation, specifically they focus on U.S. retail stores.

According to the seminal paper by [March, Simon, 1958] the two main drivers of employee turnover are desirability and ease of movement. This perspective determined two different approaches in pertinent literature: the economic approach, which focuses on factors that reflect market opportunities, job alternatives, labor demand and pay levels, and the psychological approach, which concentrates on "...the psychological processes, antecedents of voluntary turnover, individuals' perceptions of their jobs and work environment and the impact of these on their motivation to leave..." [Balabanova et al., 2016, p. 23]. Under the latter approach, literature can be divided into process and content research. The first one focusses on how people quit, whereas, the second addresses why they do this. In an attempt to answer the latter question — why employees voluntarily quit their jobs? — scholars focus of studying the predictors of employee turnover, such as organizational commitment dimensions, job satisfaction dimensions, and perceived alternative employment opportunities, consequently exploring the various socioemotional, economic, external reasons and motives for quitting [Woo, Maertz, 2012].

In particular, special attention is paid to job satisfaction (as well as its antecedents), or the degree to which individuals like their jobs, which is seen as a central measure of how well a company satisfies the needs of employees and fulfills their expectations [Balabanova et al., 2016; Hausknecht, Rodda, Howard, 2009]. For instance, [Hom, Griffeth, 1995; Griffeth, Hom, Gaertner, 2000] in their comprehensive meta-analysis of turnover antecedents specify that the job satisfaction factors, or main work environment factors, are compensation (pay, pay satisfaction, distributive justice), leadership (supervisory satisfaction and leader-member exchange), co-worker (work group cohesion, co-worker satisfaction), stress (role clarity, role overload, role conflict, and overall stress), and others (promotional chances, participation,

and instrumental communication). They additionally identify the key external environmental factors, such as job content (job content, routinization, work satisfaction, and job involvement) and external environment (alternative job opportunities and comparison of alternatives with present job) to be relevant turnover predictors. Furthermore, [Katsikea, Theodosiou, Morgan, 2015; Wright, Cropanzano, 1998] identify the immediate precursors of export salespeople's intentions to leave — role ambiguity and role conflict — by determining what the main antecedent of job satisfaction is — job stress. Therefore, one can conclude that job satisfaction is a key mediator in the relationship between different exogenous and endogenous variables and an employee's inclination to leave / to stay with an employer [Hom, Kinicki, 2001; Hausknecht, Rodda, Howard, 2009; Steel, Lounsbury, 2009; Direnzo, Greenhaus, 2011], though such turnover predictors as age, tenure, pay, performance, organizational commitment and intention to quit are deemed important as well depending on the contextual specificity.

Country-specific focus in employee turnover studies

Despite the relevance and popularity of the topic, the majority of prior studies implement a context-free approach to the investigation of the predictive power of turnover-related factors, focusing on employee turnover in — ironically — Western developed countries [Holtom et al., 2008]. Relatively few studies, nonetheless, demonstrate that the cultural context is in fact important for the understanding of what leads to job dissatisfaction and, as a result, to employee turnover (e. g. the works of [Chen, Francesco, 2000] on China, [Balabanova et al., 2016; Dusek et al., 2016] on Russia, and [Carraher, 2011] on Estonia, Latvia, and Lithuania). For instance, [Chen, Francesco, 2000] explores, inter alia, the moderating effect of gender and education on the relationship between organizational commitment and employee turnover intentions and finds that,

under the influence of personalism and guanxi, Chinese employees behave differently from their western counterparts. Comparatively, [Carraher, 2011] examines the efficacy of predicting turnover for workers from the Baltic region using attitudes towards benefits, compensation, gender and age across a four year time frame and reveals that attitudes towards benefits are generally significant predictors of turnover for all types of workers, whereas satisfaction with pay is typically significant for non-entrepreneurs. Lastly, [Balabanova et al., 2016] considers those turnover antecedents that are expected to be the most relevant in explaining employees' intentions to leave in the context of Russia: wage satisfaction, core job-related characteristics satisfaction, and interpersonal relations satisfaction. They base their model on Herzberg's (1966) two-factor theory of motivation, which commonly used to differentiate between intrinsic — hygienic — rewards that include transactional and relational aspects of work, and extrinsic — motivating — rewards, such as such as interesting and meaningful work, personal development, empowerment, etc. [Herzberg, 1966]. The authors further build their logic from the premise that turnover factors may be culture or context-dependent; in order to account for the idiosyncrasy of contemporary Russia, they also include "blat" — the possessed personal connections and networks — as a mediator in the relationship between different forms of job satisfaction and intentions to quit the Russian organizations.

Considering the contextual focus of our research, we argue that the political transformations that occurred after the Global Financial Crisis not only shaped the economic situation and international relations of Russia, but also affected the beliefs and attitudes of Russian employees toward their jobs. In particular, the results of different sociological surveys and opinion polls demonstrate that their perception of responsibility, trust in governmental institutions, national pride, etc., have changed significantly in the last decade (all according to

Levada-center data). Specifically, locals value their work a lot less, regardless of the pay, and consider their job primarily as a way to earn a livelihood [Levada-center, 2016; 2017]. In general, the non-Western collectivist approach to human resource management as well as the inherited Soviet "traditions" in conducting business [Vaiman, Holden, 2011; Latukha, 2015] are continuing to prosper in the aforementioned conditions: Russian employees, on average, are not treated as valuable resources in comparison to their Western counterparts, development of necessary skills and competences is often neglected, and motivation and rewards' systems are not properly utilized; there are, most importantly, limited career perspectives and work monotonicity [Walker, 2015], which negatively affect the working behavior of the locals and result in extraordinarily high turnover rates [Balabanova et al., 2016]. Having this said, the present study aims to expand the pertinent literature on employees' turnover to the context of companies operating in knowledge-intensive industries (i.e. industries of the tertiary and quaternary sectors of the economy) in the period of 2016–2018 in Russia. Such industries are more vulnerable to employee turnover than those from other sectors of the economy as valuable knowledge regarding environmental conditions, business practices, and relationship norms that prevail in foreign markets accumulates after many years of personal experience, which means finding a competent replacement to fill a vacant position in a specific market may be extremely difficult [Katsikea, Theodosiou, Morgan, 2015].

Methodology

The most common research approaches and methods employed by scholars investigating different aspects of employee turnover (like antecedents and consequences) are questionnaires for data collection, and structural modelling and regression analysis for data analysis. The main drawback of these methods is they are mostly one-sided, i.e., exploring specific employees' roles — people who left/are leaving (reluctant leavers vs reluctant stayers; enthusiastic leavers vs enthusiastic stayers) or general indicators unrelated to turnover behavior (such as job satisfaction and commitment) (e.g. [Li et al., 2016]). They are also subject to many types of biases caused, for instance, by common rater, item characteristic, item context and measurement context effects [Podsakoff et al., 2003] — particular data collection instruments may often assume (determine) particular answers. Meanwhile, despite the accumulated massive of heterogeneous electronic data becoming an important information source for understanding different management-related phenomena [Platanou et al., 2018], there are few research articles in the different management domains that apply research methods capable of processing big data. In terms of our research objective, gathering information left by thousands of employees anonymously (online) for further analysis via text mining techniques is a good solution to obtaining valid and reliable results.

Data collection

This study is based on open text data generated by users of otrude.net, one of the largest databases of reviews of employers, operating in Russia. The data were extracted from the website via a web-scraping script written in the Python programming language extended by the Beautiful Soup module. The website constitution allowed collecting data for the period of April 2016 — April 2018, earlier entries were encrypted, making the required data more difficult and time-consuming to retrieve for further analysis.

As mentioned above industries of the tertiary and quaternary sectors are more vulnerable to employee turnover as valuable knowledge accumulates after many years of practice and personal experience (employees have more bargaining power and make more substantive job-related decisions); therefore, we focused primarily on those reviews left about companies operating in industries with
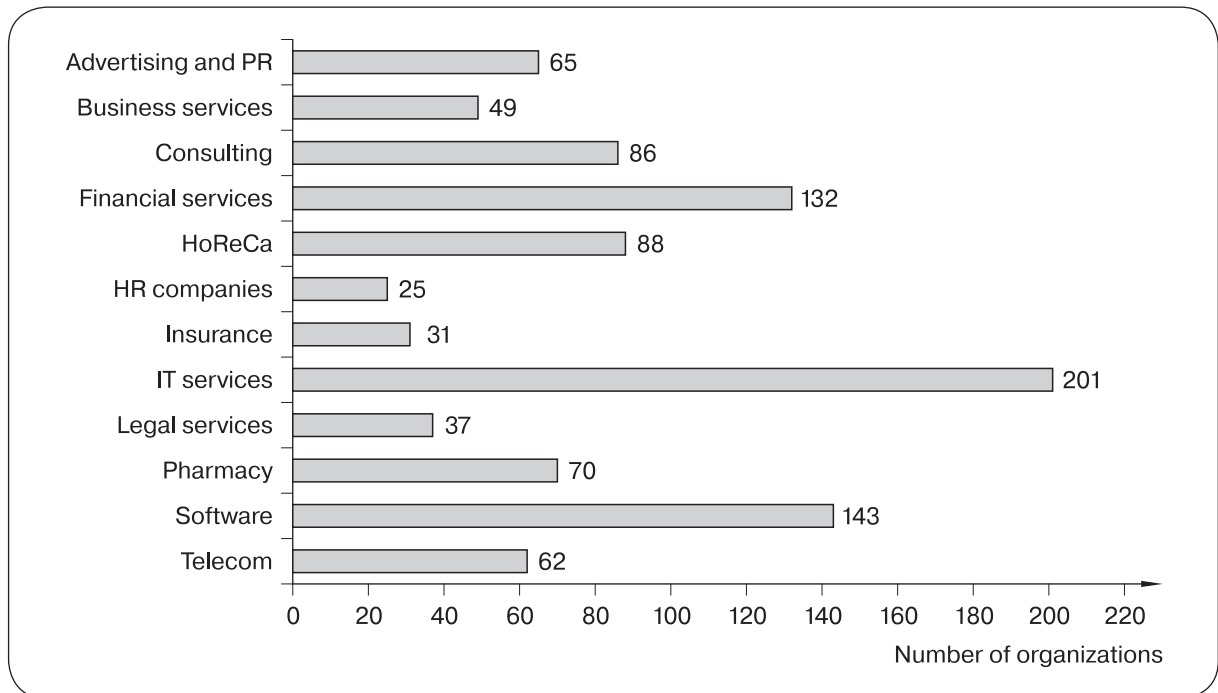
Fig. 1. Distribution of the sample across industries

high knowledge intensity, specifically software, IT services, advertising and PR, insurance, telecommunications, HoReCa, pharmaceuticals, legal services, financial services, business services, consulting and HR services industries (see Fig. 1). For the sake of validity and reliability, we considered only those companies that had more than one "useful" review, meaning that the difference between upvotes and downvotes was nonnegative. As a first step, we obtained a list of 989 organizations and their corresponding web-pages for further review retrieval of review texts. We then ran a data collection script with minor delays in between requests to ensure there was no excessive server load.

This data collection procedure yielded a sample of 10 935 recordings, each of which contained the following information: date of review, company location, job position (which the author of the review occupies/occupied/applied for), his/her work experience, review text containing the advantages and disadvantages (referred to on the website as "areas for improvement") of a par-

ticular job, the author's rating of the company (5-point "star" scale), and the review usefulness rating (upvotes and downvotes) left by other users. After data collection, we screened and cleaned the data by omitting duplicated reviews, entries left by job applicants and clients, and reviews with low user ratings (the number of downvotes > the number of upvotes). In most cases, being part of the latter group served as an indication of either a comment with highly irrelevant and/or untruthful information, or a fake review purposefully left by employers for the sake of increasing the company's standing. As a result of this data removal process, we yielded a sample of 6145 observations.

Data pre-processing

To prepare the text data for further analysis a number of pre-processing procedures have been launched. First, all of the words in the collected reviews were lemmatized, i. e. transformed in their basic form: nominative case (due to language-specific declension of nouns,

e. g. "of organization" → "organization"), singular ("colleagues" → "colleague"), masculine (most parts of speech are gender-specific in Russian) and infinitive forms (e. g. "worked" → "work"), while the constituting letters were switched from proper case to all lowercase. Lemmatization was performed using myStem package for Python (developed and distributed by Yandex) [MyStem, 2018]. Second, with the nltk package corpus we excluded the most common stop words (including prepositions, conjunctions and pronouns) from further analysis. Third, all punctuation characters were removed, while words with semantically important punctuation marks were manually located/identified and transformed into punctuation-free forms. Finally, we divided the review texts into two groups — job advantages and job disadvantages — and calculated a binary indicator — "positive" or "negative" — for each review. We labeled a particular review as positive only in those cases when the overall author's job rating amounted to at least three stars.

Data analysis: Topic modelling and sentiment analysis

In this study, we combine two powerful text mining techniques: topic modelling and sentiment analysis. Topic modelling is a process of semantic clustering of big text data based on co-occurrences of words. We use the Latent Dirichlet Allocation (LDA) algorithm implemented in the BigARTM library for Python [Vorontsov et al., 2015]. According to the LDA data-generation process, every word in a document is generated as follows: in the first step, a topic is randomly selected from the multinomial distribution of topics; in step two, a word is randomly picked from the selected topic, which consists of a multinomial distribution of words [Hong, Davison, 2010]. Both, the topic and word distributions, have a Dirichlet prior. "Latent" in this case means that the topics are hidden in the data sets and can only be inferred from observable data. For instance, if some words frequently co-occur in many documents (reviews) the LDA algorithm classifies

them in a single topic for further interpretation by the researcher. The LDA model requires the researcher to specify the amount of topics to be discovered [Grün, Hornik, 2011]. When the exact number of topics is unknown, in order to get interpretable results trial and error is the way to go.

Sentiment analysis, a subfield in natural language processing, a machine-learning technique that used to automatically classify texts by valence [Liu, 2012; Pang, Lee, 2008] and identify the authors' views on specific entities [Feldman, 2013]. Based on exogenous input of user sentiments associated with text data the sentiment analysis algorithm learns to predict user emotion to a given text (document, sentence, word or combination of words). In this study, the classification task is to label the reviews as negative or positive and then apply sentiments to topic modelling results. This aims to reveal how users' sentiments vary for different topics and how these variations are different for two groups of users: former employees and current employees.

Results

Job-related topics users discuss online

During the topic modelling procedure, we specified different numbers of topics and experimented with other model parameters. The most interpretable results were found for a 7-topic model. The topics and top words associated with them are presented in Table 1.

We could interpret and label six out of seven of the identified topics. Overall, we found that in job reviews users discuss working arrangements and schedule, working conditions, job content, salary/wage, career development and psychological climate. The most frequently occurring topic was psychological climate: on average, a discussion of this issue appeared in 20% of each review text. A typical comment within this topic was "small and amicable team, everybody is always ready to help" (translated by authors). Another major topic was connected to salary/

Topics of job reviews

Topic	Topic interpretation	Top 10 words (translated from Russian to English)	Share, %
Topic 1	Working arrangements, schedule	Day, hour, ability, shift, person, director, place/position, every/each, time, to know	12
Topic 2	Working conditions	Office, colleague, corporate, payment, medical insurance, labor/effort, center, day off, white, to live	11
Topic 3	Job content	Department, case, question, client, plan, sale, project, to perform, training, problem/issue	13
Topic 4	Salary/wage	Month, day, money, plan, to receive, salary/wage, salary, day off, minus, bonus	17
Topic 5	Career development	Leadership/Management, growth, year, career, good, opportunity, project, experience, high, level	11
Topic 6	Psychological climate, interpersonal relations with co-workers	Team, good, salary/wage, person, office, interesting, excellent, friendly, to help/assist, to arrange	20
Topic 7	Not interpreted	Person, day, money, month, to say, nothing, to do, to want, to talk, just	15

wage. On average 17% of the review text raised different compensation issues. For instance, many reviewers identified "salary/wage paid without delays" as one of the key advantages of their job. Meanwhile, the remaining topics were ranked by popularity as follows: job content (13%; e. g. "a lot of reporting work"), working arrangements and schedule (12%; e. g. "frequent overtime work"), working conditions (11%; e. g. "the dinner place could be better"), career development (11%; e. g. "there can't be any career growth without close connections to boss").

Rankings of employees' sentiments for the identified job-related topics

Table 2 provides the summary of sentiments distribution by topics. Although the overall share of positive reviews is rather small, the largest share of positive reviews is associated with the psychological climate topic (4.65%), indicating that employees (both former and current) tend to discuss the relations with their fellow co-workers in a more positive way than they discuss other job-related topics and factors. Working conditions is second most positive topic (2.36%). The most negative topic is working arrangements and schedule: only 1.28% of reviews in average contain positive messages on

working time. A comparatively moderate share of positive reviews belongs to the job content (1.66%), career development (1.96%) and salary/wage (2.1%) topics.

Differences in rankings of sentiments of former employees and current employees

Analyzing the differences in sentiment rankings between former employees and current employees, the most notable observation regards the psychological climate. Current employees assess the work environment to be approximately 5 times more positive than employees who left the company (9.71% for current and 1.79% for former employees). A significant difference was observed for the job content topic (3.05% and 0.87% for current and former employees respectively). Meanwhile, for the salary/wage, working arrangements and schedule and career development topics our findings revealed a relatively small difference between these two employee groups.

Discussion and conclusion

While previous studies show that job satisfaction is one of the main predictors of employee turnover, they rarely set the objective to understand its many dimensions in rela-

Shares of positive reviews by topics

Topic	Former employees, % ($n = 3930$)	Current employees, % ($n = 2215$)	Total, %
Overall	11.2	29.53	17.77
Topic 1. Working arrangements and schedule	0.84	2.05	1.28
Topic 2. Working conditions	1.42	4.01	2.36
Topic 3. Job content	0.87	3.05	1.66
Topic 4. Salary/wage	1.52	3.11	2.1
Topic 5. Career development	1.28	3.17	1.96
Topic 6. Psychological climate, interpersonal relations with co-workers	1.79	9.71	4.65
Topic 7. Not interpreted	1.7	2.01	1.82

tion to employee withdrawal decisions. In this study, we present a framework that automatically derives factors of job satisfaction and classifies the sentiments associated with them for online job reviews of two groups of employees: those who still work for a particular company and those who left. This framework effectively employs LDA topic modeling and sentiment analysis, which are two highly acclaimed analytical methods that are specifically designed to tackle the challenge of big data in textual formats. Conceptualizing the discussed topics as factors of job satisfaction, we used more than 6000 reviews from a popular Russian job review web-site (http://otrude.ru) to explore what job-related topics Russian employees discuss online, what the rankings of employees' sentiments are for the identified job-related topics, and how the rankings of sentiments of former employees differ from those of current employees.

Our study contributes to turnover and job satisfaction literature in at least three ways. First, while prior studies focused on job satisfaction as single integrative concept, we decomposed it and identified six major factors of job satisfaction (topics) discussed in online job reviews. They included working arrangements and schedule, working conditions, job content, salary/wage, career development, psychological climate and interpersonal relations with co-workers. The most widely discussed topics were psychological climate and interpersonal relations with co-workers and salary/wage, while least mentioned were the topics of job content, working arrangements and schedule, working conditions, career development. These results are consistent with job satisfaction and employee turnover literature that, on the one hand, identifies similar dimensions of job satisfaction (e. g. [Griffeth, Hom, Gaertner, 2000; Zimmerman, Swider, Boswell, 2018; Hom et al., 2017]), but, on the other hand, rarely makes explicit attempts to rank or compare them.

Second, we found that former employees and current employees significantly differ in their sentiments attributed to different topics. While leavers as expected demonstrated a larger overall share of negative sentiments, these shares were different for certain topics. These findings could indicate that some factors of job satisfaction were more important for employee withdrawal decisions than others. Furthermore, we identified that psychological climate and job content were the key topics that differentiate former employees' and current employees' comments. There were fewer differences in sentiments for salary/wage, working arrangements and schedule and career development topics. This result gives us an understanding that "the boiling point" forcing the employees to make the decision to

quit is to greater extent associated with interpersonal relations than with formal work arrangements usually provided by company's HR practices. This suggests that in their decision to leave a company employees are more likely to tolerate salary/wage, career development and working arrangements dissatisfaction than bad relationships with co-workers and the lack of interest in the job itself [Woo, Maertz, 2012]. Additionally, the distribution of sentiments across topics illustrates that employees tend to negatively perceive the job design and work itself, while social and reward issues considered as the positive aspects of work. This is more or less in line with the studies exploring the drivers of job satisfaction and dissatisfaction by incorporating the ideas of extrinsic and intrinsic motivation from Herzberg's motivation-hygiene theory [Porter, Steers, 1973; Sachau, 2007].

Third, since we focused on reviews on Russian companies, our study also contributes to the understanding of employee behavior in Russia. The shares of topics discussed and sentiments associated with them suggest that for Russian employees socioemotional factors of job satisfaction are of greater importance than economic ones (i.e. salary and promotion opportunities). This finding reevaluates the results obtained in earlier context-specific studies, such as [Carraher, 2011; Balabanova et al., 2016] that focus on CIS countries and determine satisfaction with pay as a key factor of withdrawal decisions for local employees. The latter study, for instance, found wage satisfaction to be the strongest predictor of employees' intentions to leave compared to core job-related and interpersonal relations satisfaction, whereas as our research assumed both would go hand in hand. Our findings may also partially explain or be a consequence of the quite peculiar model of labor market in Russia. According to [Gimpleson, Kapeliushnikov, 2015], one of the main features of the Russian labour market is the domination of price adjustment to shocks over quantity adjustment (i.e. the

combination of high wage flexibility and high employment rigidity). Because of that Russian employees tend to tolerate decreases in wage more than employees from developed and other emerging markets.

Our study also provides insights for managers seeking to motivate and retain employees. Companies with high turnover rates should not consider economic factors as the main and/or only reason for employees leaving the organization and they should carefully inspect possible problems associated with the socioemotional factors. In many cases the prevention of employee outflow would imply resolving intra-team conflicts (psychological climate) and redesigning key working practices and routines (work content). Additionally, companies should pay more attention to the non-economic effects of different managerial interventions, including the implementation of context-sensitive HR practices. For instance, while pay-for-performance salary schemes increase extrinsic motivation and shot-term productivity, in the long run they might affect psychological climate, reduce job satisfaction and cause unwanted employee turnover.

Limitations and further research directions

Due to the novelty of the data collection and analysis techniques, our study should be treated as an exploratory one, which is at its early stage development. Although we endeavored to maximize the quality of our work with appropriate procedures and validation techniques, the study suffers from a number of limitations. Firstly, we referred to one data source, which, in fact, provided us with rather noisy data. In the future, it would be necessary to gather data from multiple resources, including other job review databases and websites in other languages. Here, a comparative study within and across different national contexts and industries seems to be a perspective research direction. In addition, by considering the massive of rich and heterogeneous electronic data accumulated by companies as well (e.g., an-

nual reports, corporate websites and newsletters, etc.) scholars can investigate whether there are any differences in the job-related topics determined by employees and organizations (i.e. two different perspectives). This would allow, for instance, determining potential psychological contract breaches, which, depending on the severity, could lead to diminishing the employment relationship characterized by organizational citizenship behavior [Guest, Clinton, 2005]. According to [Seopa, Wöcke, Leeds, 2015], "[t]he breach of the psychological contract [which consists of the employees' individual beliefs regarding the terms and conditions of the exchange agreement between themselves and their organizations]… could trigger the intention to leave… or result in the actual exit from the organization" (p. 722).

Secondly, due to the nature of our data, it cannot be considered completely unbiased in terms of randomness. It is still a subject of debates whether people who write online reviews have peculiar personalities and social characteristics and, thus, whether they can be considered representative of the general population [Ljepava et al., 2013]. Lastly, both LDA and sentiment analysis are in need of more effective and accurate means of validation for their findings. As a caution, no matter how big "big data" is, the information and knowledge derived from it may still be rather misleading [Lazer et al., 2014; Lin, Lucas, Shmueli, 2013]. Thus, an important task for researchers is to elaborate standards of validity for application of text mining techniques in management and related social science disciplines.

LITERATURE

Al Ariss A., Sidani Y. 2016. Comparative international human resource management: Future research directions. *Human Resource Management Review* **26** (4): 352–358.

Balabanova E., Efendiev A., Ehrnrooth M., Koveshnikov A. 2016. Job satisfaction, blat and intentions to leave among blue-collar employees in contemporary Russia. *Baltic Journal of Management* **11** (1): 21–43.

Carraher S. M. 2011. Turnover prediction using attitudes towards benefits, pay, and pay satisfaction among employees and entrepreneurs in Estonia, Latvia, and Lithuania. *Baltic Journal of Management* **6** (1): 25–52.

Chen Z. X., Francesco A. M. 2000. Employee demography, organizational commitment, and turnover intentions in China: do cultural differences matter? *Human Relations* **53** (6): 869–887.

Cotton J. L., Tuttle J. M. 1986. Employee turnover: A meta-analysis and review with implications for research. *Academy of Management Review* **11** (1): 55–70.

Direnzo M. S., Greenhaus J. H. 2011. Job search and voluntary turnover in a boundaryless world: A control theory perspective. *Academy of Management Review* **36** (3): 567–589.

Dusek G. A., Clarke R., Yurova Y., Ruppel C. P. 2016. Employee turnover in international brand hotels in Russia: A comparison of nationals and foreign nationals. *Journal of East-West Business* **22** (1): 51–75.

Feldman R. 2013. Techniques and Applications for Sentiment Analysis. *Communications of the ACM* **56** (4): 82–89.

Gimpleson V. E., Kapeliushnikov R. I. 2015. Russian model of the labor market: The test of the crisis. *Journal of the New Economic Association* (2): 249–254.

Griffeth R. W., Hom P. W., Gaertner S. 2000. A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management* **26** (3): 463–488.

Grün B., Hornik K. 2011. Topicmodels: An R package for fitting topic models. *Journal of Statistical Software* **40** (13): 1–30.

Guest D., Clinton M. 2005. *Contracting in the UK. Current Research Evidence on the Impact of Flexible Employment on and the Nature of the Psychological Contract*. Kings College, London.

Hancock J. I., Allen D. G., Bosco F. A., McDaniel K. R., Pierce C. A. 2013. Meta-analytic review of employee turnover as a predictor of firm performance. *Journal of Management* **39** (3): 573–603.

Hausknecht J. P., Rodda J., Howard M. J. 2009. Targeted employee retention: Performance-based and job-related differences in reported reasons for staying. *Human Resource Management* **48** (2): 269–288.

Herzberg F. 1966. *Work and the Nature of Man.* Corwell Co: N. Y.

Holtom B. C., Mitchell T. R., Lee T. W., Eberly M. B. 2008. Turnover and retention research: A glance at the past, a closer review of the present, and a venture into the future. *Academy of Management Annals* **2** (1): 231–274.

Hom P. W., Griffeth R. W. 1995. *Employee Turnover.* South-Western: Cincinnati, OH.

Hom P. W., Kinicki A. J. 2001. Toward a greater understanding of how dissatisfaction drives employee turnover. *Academy of Management Journal* **44** (5): 975–987.

Hom P. W., Lee T. W., Shaw J. D., Hausknecht J. P. 2017. One hundred years of employee turnover theory and research. *Journal of Applied Psychology* **102** (3): 530–545.

Hong L., Davison B. D. 2010. Empirical study of topic modeling in Twitter. In: Melville P., Leskovec J., Provost F. (eds). *Proceedings of the First Workshop on Social Media Analytics*. Association for Computing Machinery: N. Y.; 80–88.

Jaramillo F., Mulki J. P., Boles J. S. 2013. Bringing meaning to the sales job: The effect of ethical climate and customer demandingness. *Journal of Business Research* **66** (11): 2301–2307.

Katsikea E., Theodosiou M., Morgan R. E. 2015. Why people quit: Explaining employee turnover intentions among export sales managers. *International Business Review* **24** (3): 367–379.

Latukha M. 2015. Talent management in Russian companies: Domestic challenges and international experience. *The International Journal of Human Resource Management* **26** (8): 1051–1075.

Lazer D., Kennedy K., King G., Vespignani A. 2014. The parable of Google flu: Traps in big data analysis. *Science* **343** (March): 1203–1205.

Lee T. W., Mitchell T. R., Sablynski C. J., Burton J. P., Holtom B. C. 2004. The effects of job embeddedness on organizational citizenship, job performance, volitional absences, and voluntary turnover. *Academy of Management Journal* **47** (5): 711–722.

Levada-center. 2016. *Russian Public Opinion 2013–2015*. [Electronic resource]. Available at: http://www.levada.ru/cp/wp-content/uploads/2016/01/2013-2015-Eng1.pdf (accessed 10.05.2018).

Levada-center. 2017. *Labor Orientations*. Press-release. [Electronic resource].Available at: https://www.levada.ru/en/2017/ 02/28/labor-orientations/ (accessed 10.05.2018).

Li J. J., Lee T. W., Mitchell T. R., Hom P. W., Griffeth R. W. 2016. The effects of proximal withdrawal states on job attitudes, job searching, intent to leave, and employee turnover. *Journal of Applied Psychology* **101** (10): 1436–1456.

Lin M., Lucas Jr. H. C., Shmueli G. 2013. Research commentary — Too big to fail: Large samples and the p-value problem. *Information Systems Research* **24** (4): 906–917.

Liu B. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies* **5** (1): 1–167.

Ljepava N. R., Orr R., Locke S., Ross C. 2013. Personality and social characteristics of Facebook non-users and frequent users. *Computers in Human Behavior* **29** (4): 1602–1607.

March J. G., Simon H. E. 1958. *Organizations*. Wiley: N. Y.

Meier K. J., Hicklin A. 2007. Employee turnover and organizational performance: Testing a hypothesis from classical public administration. *Journal of Public Administration Research and Theory* **18** (4): 573–590.

Mobley W. H. 1977. Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of Applied Psychology* **62** (2): 237–240.

Morrow P., McElroy J. 2007. Efficiency as a mediator in turnover — Organizational performance relations. *Human Relations* **60** (6): 827–849.

MyStem. 2018. [Electronic resource]. Available at: https://tech.yandex.ru/mystem (accessed 10.05.2018).

Pang B., Lee L. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* **2** (1–2): 1–135.

Platanou K., Mäkelä K., Beletskiy A., Colicev A. 2018. Using online data and network-based text analysis in HRM research. *Journal of Organizational Effectiveness: People and Performance* **5** (1): 81–97.

Podsakoff P. M., MacKenzie S. B., Lee J. Y., Podsakoff N. P. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology* **88** (5): 879–903.

Porter L. W., Steers R. M. 1973. Organizational, work, and personal factors in employee turnover and absenteeism. *Psychological Bulletin* **80** (2): 151–176.

Sachau D. A. 2007. Resurrecting the motivation-hygiene theory: Herzberg and the positive psychology movement. *Human Resource Development Review* **6** (4): 377–393.

Sager J. K., Varadarajan P. R., Futrell C. M. 1988. Understanding salesperson turnover: A partial evaluation of Mobley's turnover process model. *Journal of Personal Selling & Sales Management* **8** (1): 21–36.

Seopa N., Wöcke A., Leeds C. 2015. The impact on the psychological contract of differentiating employees into talent pools. *Career Development International* **20** (7): 717–732.

Shaw J. D. 2011. Turnover rates and organizational performance: Review, critique, and research agenda. *Organizational Psychology Review* **1**(3): 187–213.

Shaw J. D., Duffy M. K., Johnson J. L., Lockhart D. E. 2005. Turnover, social capital losses, and performance. *Academy of Management Journal* **48** (4): 594-606.

Steel R. P., Lounsbury J. W. 2009. Turnover process models: Review and synthesis of a conceptual literature. *Human Resource Management Review* **19** (4): 271–282.

Ton Z., Huckman R. S. 2008. Managing the impact of employee turnover on performance: The role of process conformance. *Organization Science* **19** (1): 56–68.

Vaiman V., Holden N. 2011. Talent management's perplexing landscape in Central and Eastern Europe. In: Scullion H., Collings D. (eds). *Global Talent Management*. Routledge: London; 178–193.

Vorontsov K., Frei O., Apishev M., Romov P., Dudarenko M. 2015. Bigartm: Open source library for regularized multimodal topic modeling of large collections. *Communications in Computer and Information Science* **542**: 370–381.

Walker C. 2015. 'I don't really like tedious, monotonous work': Working-class young women, service sector employment and social mobility in contemporary Russia. *Sociology* **49** (1): 106–122.

Woo S. E., Maertz C. P., Jr. 2012. Assessment of voluntary turnover in organizations: Answering the questions of why, who, and how much. In: Schmitt N. (eds). *Oxford Library of Psychology. The Oxford Handbook of Personnel Assessment and Selection*. Oxford University Press: N. Y.; 570–594.

Wright T. A., Cropanzano R. 1998, Emotional exhaustion as a predictor of job performance and voluntary turnover. *Journal of Applied Psychology* **83** (3): 486–493.

Zimmerman R. D., Swider B. W., Boswell W. R. 2018. Synthesizing content models of employee turnover. *Human Resource Management*. https://doi.org/10.1002/hrm.21938

Почему сотрудники покидают российские компании? Интеллектуальный анализ текстов онлайн-отзывов о работе

Д. Н. Соколов
Аспирант, Высшая школа менеджмента Санкт-Петербургского государственного университета, Россия
E-mail: d.sokolov@gsom.pu.ru

Л. В. Селивановских
Аспирант, Высшая школа менеджмента Санкт-Петербургского государственного университета, Россия
E-mail: louisa.selivanovskikh@gmail.com

Е. К. Завьялова
Профессор, Высшая школа менеджмента Санкт-Петербургского государственного университета, Россия
E-mail: zavyalova@gsom.spbu.ru

М. О. Латуха
Доцент, Высшая школа менеджмента Санкт-Петербургского государственного университета, Россия
E-mail: marina.latuha@gsom.spbu.ru

В данной статье анализируются темы и связанные с ними сантименты в онлайн-отзывах о работе в 989 российских интеллектуально-емких компаниях. Пользуясь такими методами интеллектуального анализа текстов, как тематическое моделирование и анализ тональности текста, авторы предпринимают попытку раскрыть ключевые факторы удовлетворенности трудом и выявить разницу в отношении к ним текущих и уже уволившихся сотрудников. Результаты анализа показывают, что (1) график работы, (2) физические условия труда, (3) содержание работы, (4) заработная плата, (5) карьерное развитие, а также (6) психологический климат и межличностные отношения являются ключевыми темами, обсуждаемыми в исследуемых онлайн-отзывах. При этом последняя — психологический климат и межличностные отношения — оказалась самой часто обсуждаемой темой, в особенности среди продолжающих работать в компании сотрудников. Полученные данные позволили сделать вывод о том, что при решении об уходе из компании сотрудники более склонны выдерживать негативные факторы экономического характера (связанные, например, с заработной платой, карьерными перспективами или графиком работы), нежели негативные факторы социально-эмоционального характера (например, отношения с коллегами или содержание работы).

Ключевые слова: текучесть персонала, удовлетворенность трудом, интеллектуальный анализ текста, тематическое моделирование, анализ тональности текста, Россия.

JEL: M51.