The Application of Methods of Semantic and Correlation Analysis Through Studies of the Annual Reports of Russian Companies According to the Principle of Integrated Reporting "Reliability and Completeness"

Vera V. Plotnikova[1 [0000−0002−4661−0942]], Maria S. Shilovskaya[2], and Anton A. Dvorak[3]

[1] Saratov Socio-Economic Institute of Plekhanov Russian University of Economics, Saratov, Russia, http://www.seun.ru
[2 3] Yuri Gagarin State Technical University of Saratov, Saratov, Russia, Russia, http://en.sstu.ru

Abstract. The article contains an empirical study of the annual reports of the largest Russian companies for 2011–2015 according to the principle of integrated reporting "reliability and completeness". It is based on a computer model of the semantic analysis of the report texts and the correlation analysis of the values of financial ratios and the opinion of the company's management about their financial position and financial results. As part of study, the distribution algorithm of words by semantic groups is proposed, and software that allows semantic analysis of the information of annual reports according to the principles of integrated reporting is developed. The correlation analysis was carried out using standard Microsoft Excel tools. According to the analysis of the 30 largest Russian holding entities, the conclusion is that in 70% of the reports the opinion of the management of Russian companies does not have an affiliation with the actual values and the dynamics of financial ratios as well as it does not correspond to the principle of IR "reliability and completeness". In the end it was concluded that companies need to use a standardized system of financial indicators to generate reports that can provide an objective view of the company's current situation, and also to classify analytical interpretations of these indicators.

Keywords: integrated reporting, financial position, semantic analysis, correlation analysis

1 Introduction

The problem of reliability and quality of the information disclosed in the annual reports of Russian companies continues to be relevant for a long time. This is due to the uncertainty about the predicted value of such information and the objectivity of management's assessment of the company's financial position.

In other words, there is a situation when management tries to manipulate the opinion of stakeholders for its own benefit, hiding the real financial situation of the company. Subjective evaluations of the management of the company do not always reflect the real picture of doing business [11]. And this is the reason why over the last years, the attention to the quality of the descriptive information disclosed in the annual reports has increased significantly. The research is conducted on the basis of annual and integrated reports of the largest Russian companies. Russia participates in the pilot project on the transition to the international standards of integrated reporting (hereinafter IR) and the GRI standards. At the present point in time, Russian business is still not ready for open communication with all stakeholders, since often even financial statements fail to meet the requirements of reliability and completeness of information disclosed. The new model of IR is oriented toward a constructive dialogue between companies and users of its information. In this case, the requirements to the quality of financial and non-financial information disclosed, as well as to the management's objective overview of their financial position and the prospects for creating value for a long period, are raised. To a large extent, the form of the company management's presentation of information about financial situation in the content of annual reports depends on the conditions that affect the financial position of the company at the time of preparation of reporting. The way the management reacts to these conditions and accordingly offers its opinion about the current position of the company and the prospects for its development is the knowledgebase of our research, which allows us to draw a conclusion about the maturity of Russian business and also its availability to move to a new level of disclosed corporate information. One of the tools for analyzing the quality of information contained in the annual reports of Russian companies and the compliance of this information with the principles of integrated reporting is the semantic analysis combined with the correlation analysis. Carrying out that sort of research requires computer modeling of the process of word distribution by semantic groups according to Loughran-McDonald Word Lists by distinguishing positive, negative, uncertainty, and litigious, constraining, modal words that make it possible to analyze the entire volume of text information contained in the annual reports and through the use of correlation analysis to establish a correlation between the management's subjective opinion and the objective values of financial ratios. This approach allows determining the quality of financial and non-financial information for compliance with the principle of IR "reliability and completeness".

2 Literature Review

There are many studies related to determining the relationship between the quality of information disclosed in corporate reports of companies and their financial position. Some of them are dedicated to the study of the influence of the quality of information disclosed in annual reports of companies to the valuation of their shares traded on the securities market [5] Tennyson B. M. and

others related the bankruptcies of some American companies to the quality of managerial descriptive information disclosed in annual reports [13]. One study showed a positive relation between the quality evaluation of information disclosed of the financial condition of Italian companies and the forecast precision of financial analysts [10]. In all these studies the methods of economic and mathematical modeling were used as a methodological tool for analysis. According to the world practice, lately the interest to evaluation and analysis of the quality of corporate information disclosed with the use of text or semantic analysis has increased significantly. Semantic analysis helps to study the linguistic structure of the text, determine the tone and mood of corporate reports that helps to assess their impact on the ability to identify situations when there are attempts to avoid management from disclosing the real financial position of the company. Text mining analysis works as an additional tool for forecasting the bankruptcy of the company, also by using this technique it is possible to distinguish management's unintentional signals about the current financial position of the company and its further development which are hidden in the report texts [8]. As a part of the text mining analysis of company reports, it is possible to determine whether certain categories of words are prevalent and to what extent they are used by management in individual cases. Some authors explicitly specify the presence of fraudulent application in the annual reports and conduct an appropriate research on how the linguistic structure of the reported text helps to identify indicators of true or false statements through textual analysis of emotions [6]. In the works of Caserio C., Panaro D., Trucco S. [2] there is a study that was conducted on a sample of US listed financial companies, in which, through the use of semantic and regression analysis, a relation between the content of MD&A (management's analysis of the financial condition and results of operations) and the financial position of the company was established. Also, it was revealed how much the content of this report affects the quality of forecasting the financial condition of the company in the future. Among the Russian studies, the most notable are the works of Efimova O.V., which details the relation between long-term sustainable development of the company and the influence of all stakeholders, and puts forward the thesis on the importance of qualitative of non-financial information disclosed to improve the sustainability of the company's development [4]. In the work of Malinovskaya N.V. (2015), a general analysis was carried out to ensure that the corporate reporting of the largest Russian companies was consistent with the leading principles for the preparation and presentation of integrated reporting, and concluded that the quality and information transparency of the reports were gradually improving [9]. It is important to pay attention to the insufficient study of the chosen research topic in the context of the activities of Russian companies, meaning to identify the compliance of company reports with the principles of integrated reporting using semantic analysis tools, since earlier such studies were conducted only on the basis of foreign companies (in particular, Italian [10] and American [2]). This places emphasis on our research in the context of the adaptation of the considered text analysis techniques to the linguistic features of the Russian language.

3 Research methodology

The basis of this study is the analysis of the quality of information disclosed by Russian companies on the financial position in accordance to the principle of IR "reliability and completeness" and the review of financial and non-financial information for material errors in the reported texts, as well as the inclusion of all material information in the report (both positive and negative). Based on this, we proposed two hypotheses of the study: I0 – the quality of information disclosed in annual reports of Russian companies is consistent with the principle of IR "reliability and completeness"; I1 – the quality of information disclosed in annual reports of Russian companies does not meet the principle of IR "reliability and completeness". Currently, the analysis of the semantic orientation of company reports is characterized by the presence of a significant variety of text analysis tools. Among them, a specific place is hold by Fog-index, Naive Bayes Classifier, Loughran-McDonald Word Lists, and others. At the same time, it is important to pay special attention to the fact that any of these methods of analysis are based on linguistic features and textual structure in the native language of authors, then there is the English language, this fact is emphasized by scientists. In our opinion, the most interesting is the method of carrying out text analysis on Loughran-McDonald Word Lists, according to which there are several semantic fields or groups that determine the tonality of a word. These include positive, negative, uncertainty, litigious, constraining, modal strong and modal weak words [7]. Considering the fact that the native language for the authors of this research is the Russian language which is characterized by certain expressiveness, instead of the group modal strong and modal weak words the combined group of modal words is proposed. Modality in the Russian language is expressed by special forms of inclination, intonation and lexical means. In this group, in addition to the modal verbs expressing not the action itself, but the relation to it, it is preferable to include modal particles and words (of course, unfortunately, probably, etc.). This group also includes words that give a strong emotional coloring to the events described. For example, impressively, colossal, cardinal, etc. This category of words is rarely found in annual reports, but by their number one can estimate the degree of expression of opinion of the authors in relation to the events described [12]. Examples of words in accordance with the author's classification of semantic groups are given in Table 1.

As a rule, positive words are the most common in the texts of reports. Depending on the context, an inversion of the value of positive and negative words can occur. For example, when you use the word "growth" in the phrase "revenue growth," it is treated as an unambiguously positive, and in the phrase "costs growth" – as a negative one. The group of legal words includes legal terms and words that mark legally binding actions. Also it includes words that do not necessarily indicate a legal basis, but indicate an attitude toward the judicial sphere. The word selection for the formation of the dictionary database for each semantic group was carried out on the basis of the study of typical phrases and word combinations found in the texts of annual reports, taking into account the specifics of the financial and economic sphere regarding to the accounting,

Table 1. Examples of words by semantic groups

Semantic group	**Words**
Negative	costs, crisis, exacerbate, inefficient, lose, recession, struggle
Positive	profit, achievement, prosperity, benefit, leader, effective
Uncertainty	risk, forecast, about, almost, guess
Litigious	contestation, offering, consolidation, contract, own, become operative
Modal	can, to be able to, to want, of course, unfortunately, probably, it must be, impressively
Constraining	impose, determining, stick, require, allowing passage, limit

analysis and audit. In the course of analysis it is important to establish whether there is a relation between specific groups of words, or rather their share in the total number of words in the reports, and financial indicators of companies. This goal can be achieved through a correlation analysis which establishes not only the dependence by itself, but determines its strength and character.

4 Data Collection and Pre-processing

As the test object, there were selected 30 companies which are included in the final rating of the quality of information disclosed in the integrated reporting by Russian companies (http://transparency2015.downstream.ru/#/ru/1410), conducted by the Russian Regional Network on integrated reporting. According to this rating, the participating companies are divided into 5 groups by transparency of information: I level – "Information disclosed at the level of the best international practices"; II Level – "Information disclosed on international requirements"; III Level – "Information disclosed in excess of Russian legislative requirements"; IV level – "Information disclosed in accordance with Russian legislative requirements"; V (zero) transparency level – "Non-transparent level". The criterion for selection of companies for analysis was their rating in the overall assessment. The priority was given to companies with the highest score, included in the I, II, III level of transparency. Financial companies and organizations whose accounting is developed in a foreign language and companies that are the part of the State Atomic Energy Corporation "Rosatom" and the holding of Rosseti PJSC have been excluded from this list. In this case it makes economic sense to use only the reporting of the corporations themselves, since they also participate in the rating. The texts of the financial performance reviews for 2011–2015 were extracted (150 texts). The limitation of the analysis period is connected with the fact that only in 2011 Russian companies whose securities were admitted to trading at stock exchanges and (or) other organizers of trade on the securities market have a responsibility to provide consolidated financial statements in accordance with International Financial Reporting Standards (IFRS) [1]. The calculation of the total number of words in each text was initially carried out using the text editor MS Word, in which this function is

performed automatically. But the use of this editor assumes the implementation of the semantic analysis by the researcher manually, that is, by reading the text, searching and distributing the necessary words by groups. In connection with the great labour intensity of this work, we developed a special program "Semanomics – TextAnalyser" which performs semantic analysis automatically with the output of the number of words by groups and their frequency in the text on the screen. Further, the company's financial situation was determined by calculating financial ratios: Return on sales, ROS; Return on assets, ROA; Return on investments, ROI; Current ratio, CR; Financial stability, FS; Turnover of current assets, TR; Own funds ration, OFR. The reporting data from Consolidated Statement of Financial Position and Consolidated Statement of Income of the analyzed companies for the period 2011–2015 is used.

5 Implementation algorithm of the analysis

The basis of the algorithm for implementing semantic analysis is a combination of several ways of content search. The purpose of the program is to determine the number and frequency of occurrence of the words of each selected semantic group in the analyzed text. For this, the words and phrases of each semantic group are divided into 3 subgroups, and then 2 methods of finding matches in the text are implemented. Words that do not carry semantic meaning (conjunctions, prepositions, particles, etc.) are singled out. The scheme of specific cases of possible interactions between subgroups of words and variants of the search inside the implementation algorithm is shown in Fig. 1.

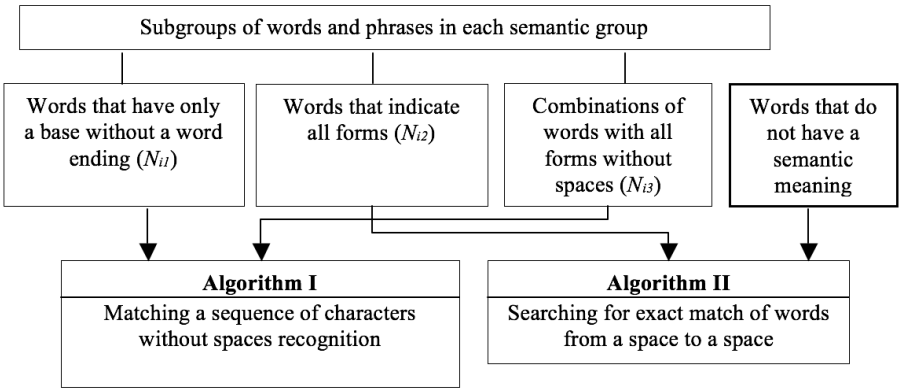

Fig. 1. Implementation algorithm of the semantic analysis

Words and phrases from the first and second groups are not included in other, longer words, so algorithm I is suitable for its search and it is simpler to implement. Words from the third group, as well as words that do not carry

semantic meaning ("and", "but", "not") can be part of other words, therefore a more sophisticated search algorithm II is implemented for them. The total number of words in the text is initially determined by the number of spaces (1):

$$N_w = N_{sp} + 1, \tag{1}$$

where N_w – total number of words; N_{sp} – the number of spaces. And then the number of words without semantic meaning is subtracted from it, and recalculation is carried out taking into account the word combinations which are considered as one word (Russian Federation, Ministry of Energy RF, public company "Company name", etc.). As a result, only the number of words bearing the semantic meaning is obtained: $N_{ws} \leq N_w$ (where N_{ws} is the number of words with semantic meaning, "semantic" words). The number of words and phrases of each group Ni consists of searching by subgroups (2):

$$N_i = N_{i1} + N_{i2} + N_{i3}, \tag{2}$$

where N_i is the number of words and phrases of each semantic group (positive, negative, uncertainty, litigious, constraining, modal);N_{i1} – the number of words included in 1 subgroup (a base without a word ending);N_{i2} – the number of words in the 2 subgroup (all forms of the word);N_{i3} – the number of words included in the 3 subgroup (Combinations of words with all forms without spaces). Their frequencies f_i are defined as the ratio of the found quantity to the total number of "semantic" words (3):

$$f_i = \frac{N_i}{N_{ws}}. \tag{3}$$

Then, the program displays the results of calculating the word number and frequency of each semantic group in the analyzed text. The final stage of the analysis is the calculation of pair correlation coefficient (Pearson) between the financial ratios and the frequency of words of each semantic group in the report for each company separately by formula (4):

$$r_{xy} = \frac{\bar{xy} - \bar{x} * \bar{y}}{\sigma_x \sigma_y}, \tag{4}$$

where x – the independent variable which characterizes financial ratios (ROA, ROS, ROI, CR, FS, TR, OFR);y – the dependent variable which characterizes the frequency of words of each semantic group (positive, negative, uncertainty, constraining, litigious, modal);σ_x – standard error of the mean of the independent variable x;σ_x – standard error of the mean of the dependent variable y [3]. This coefficient makes it possible to estimate the strength of the connection between the two analyzed parameters. Its value varies from -1 to 1, where 1 means the exact direct relation, and -1 – the exact inverse relation. Accordingly, if the ratio value is 0, then there is no relation between the parameters [3]. The connection is considered strong if the value of the coefficient falls within the interval (0.7;1.0) or (-1.0;-0.7). The value in the intervals (0;0.3) or (-0.3;0) indicates a weak connection between the parameters.

The calculation of the pair correlation coefficient can be performed using the statistical analysis package STATISTICA using the Correlation Matrices function. There is another way – the "Data Analysis" package in MS Excel. We used a simpler method – selecting the function "CORREL" from the list of the function wizard, as shown in Fig. 2.

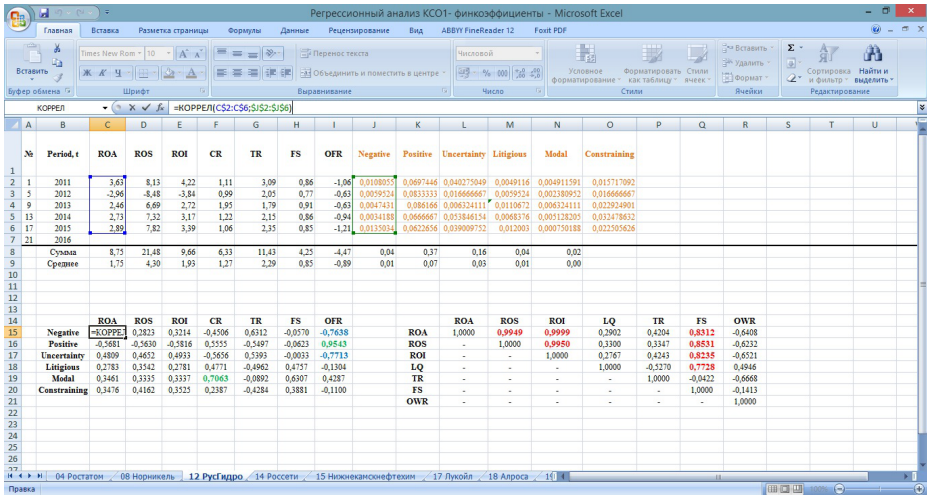

Fig. 2. Calculation of the pairwise Pearson's correlationcoefficient in MS Excel

For example, in the selected cell C15, the result of calculating the pair correlation coefficient between the return on assets (ROA) and the share of negative words in the financial survey from the integrated reports of RusHydro PJSC is displayed.

6 Empirical Findings

A general analysis of the entire sample showed that the financial position and the frequency of words in different semantic groups are interrelated weakly or moderately in most companies. Significant influence on this result was provided by insufficient data due to a thin volume of analysis (only 5 years). In order to show the results of the correlation analysis visually with the calculation of the number of companies for which a certain type of correlation was detected (direct or reverse, strong or moderate), Table 2 is formed. Here the following notation for the types of dependence is used: Dr.Md. – direct moderate; Dr.St. – direct strong; Rv.Md. – reverse moderate; Rv.St. – reverse strong. The numbers indicate the number of companies in the digest that have a dependency of the corresponding type. In the analysis, we took into account only those types of dependence, in which the total number of companies with a moderate and strong relation in total exceeds the value of 8 out of 30 or 25% of the total sample.

Table 2. Results of the correlation analysis of the dataset of firms

	ROA	ROS	ROI	CR	TR	FS	OFR
Negative	Dr.Md.1	Dr.Md.1	Dr.Md.3	Dr.Md.3	Dr.Md.2	Dr.Md.1	Dr.Md.0
	Dr.St.2	Dr.St.2	Dr.St.0	Dr.St.0	Dr.St.2	Dr.St.4	Dr.St.2
	Rv.Md.5	Rv.Md.5	Rv.Md.5	Rv.Md.3	Rv.Md.4	Rv.Md.3	Rv.Md.6
	Rv.St.8	Rv.St.7	Rv.St.7	Rv.St.6	Rv.St.5	Rv.St.4	Rv.St.8
Positive	Dr.Md.6	Dr.Md.5	Dr.Md.6	Dr.Md.3	Dr.Md.3	Dr.Md.5	Dr.Md.6
	Dr.St.10	Dr.St.10	Dr.St.10	Dr.St.6	Dr.St.7	Dr.St.6	Dr.St.8
	Rv.Md.2	Rv.Md.1	Rv.Md.2	Rv.Md.2	Rv.Md.2	Rv.Md.3	Rv.Md.3
	Rv.St.2	Rv.St.3	Rv.St.2	Rv.St.1	Rv.St.4	Rv.St.2	Rv.St.0
Uncer-	Dr.Md.2	Dr.Md.4	Dr.Md.2	Dr.Md.4	Dr.Md.4	Dr.Md.3	Dr.Md.6
tainly	Dr.St.7	Dr.St.5	Dr.St.6	Dr.St.3	Dr.St.4	Dr.St.5	Dr.St.6
	Rv.Md.3	Rv.Md.5	Rv.Md.5	Rv.Md.3	Rv.Md.1	Rv.Md.3	Rv.Md.5
	Rv.St.7	Rv.St.6	Rv.St.5	Rv.St.4	Rv.St.5	Rv.St.5	Rv.St.3
Litigious	Dr.Md.6	Dr.Md.5	Dr.Md.5	Dr.Md.6	Dr.Md.5	Dr.Md.3	Dr.Md.4
	Dr.St.1	Dr.St.2	Dr.St.2	Dr.St.2	Dr.St.2	Dr.St.4	Dr.St.4
	Rv.Md.2	Rv.Md.2	Rv.Md.2	Rv.Md.5	Rv.Md.4	Rv.Md.2	Rv.Md.2
	Rv.St.4	Rv.St.3	Rv.St.4	Rv.St.4	Rv.St.2	Rv.St.6	Rv.St.5
Modal	Dr.Md.0	Dr.Md.0	Dr.Md.0	Dr.Md.2	Dr.Md.3	Dr.Md.2	Dr.Md.3
	Dr.St.4	Dr.St.4	Dr.St.4	Dr.St.4	Dr.St.4	Dr.St.4	Dr.St.2
	Rv.Md.3	Rv.Md.4	Rv.Md.3	Rv.Md.5	Rv.Md.1	Rv.Md.4	Rv.Md.2
	Rv.St.5	Rv.St.4	Rv.St.5	Rv.St.3	Rv.St.5	Rv.St.5	Rv.St.6
Const-	Dr.Md.3	Dr.Md.4	Dr.Md.3	Dr.Md.1	Dr.Md.3	Dr.Md.4	Dr.Md.2
raining	Dr.St.2	Dr.St.1	Dr.St.1	Dr.St.3	Dr.St.3	Dr.St.2	Dr.St.2
	Rv.Md.1	Rv.Md.2	Rv.Md.2	Rv.Md.4	Rv.Md.1	Rv.Md.6	Rv.Md.2
	Rv.St.8	Rv.St.7	Rv.St.8	Rv.St.7	Rv.St.5	Rv.St.4	Rv.St.8

According to Table 2, it can be seen that 50% of companies in the digest have a clear positive relation between all financial ratios and the frequency of positive words. This situation is considered to be natural for companies, as in the case of improving financial condition, it is obvious that when reflecting their opinion in the report, management will use positive words more often. As previously mentioned, some companies use positive words to relief negative events, which can lead to the identification of feedback between these indicators. There is another obvious correlation between financial ratios and the frequency of negative words. Approximately 40% of companies have a reverse and strong or moderate relation between profitability, their own funds ratio and the proportion of negative words in the reports. For all other financial ratios, a similar relation is observed only in 30% of companies. This suggests that the better the company's financial position, the fewer negative words are used in the reports. This is a controversial issue, since this dependence is not found by all companies in the digest. The relation between the frequency of uncertain words and financial indicators is of particular interest. This dependence is direct strong or moderate for almost all financial ratios for one third of the companies, and for the other third of the

companies it is inverse. For the remaining 30% of companies this dependence is not established or is weak. This may indicate that the leadership of the first third of companies is about expressing their opinion with reserve and making forecasts using more words indicating uncertainty or uncertainty in the events described, despite the improving financial condition. It will also be the matter-of-course that, if there is weakening of financial standing of the company, the number of insecure words will increase, and vice versa, as is the case for another third of the companies. The relation between the financial position and the frequency of restrictive words is reversed and is observed in 35% of companies. This confirms our view that the increase in the number of verbal constructions indicating the increased pressure or the emergence of new obligations and the worsening of the financial situation are interrelated [12]. The frequency of modal words in the financial results reviews for most of the companies does not have a close connection with the financial position of the company. But, nevertheless, as the data in Table 2 show, for those companies where it is found, the direction of this connection is predominantly the reverse. That is, about 25–30% of companies with a deteriorating financial situation often use modal words, and vice versa. In general, the direction of this connection is irrelevant, since both the improvement and deterioration of the financial situation may influence the increase or decrease in the number of modal words with the same degree. In this case, the value should have only the quantity, since the more of them, the better reflected the attitude of the management to the information provided [12]. Depending on the semantic group, in 40–50% of the companies the profitability ratios (ROA, ROI, ROS), as well as OFR have a closer connection with virtually all groups of words. This suggests that the company's financial result is estimated by the company's management better than the financial position. Management pays more attention to the description of financial results due to the greatest popularity of information on them from stakeholders, since often the success of the company is primarily measured by profitability and efficiency. The conclusions reached allow to reject a hypothesis of the investigation of H0 and to accept the hypothesis H1. The quality of information disclosed in the annual reports of Russian companies does not correspond to the principle of IR "reliability and completeness". In 70% of the reports, the opinion of the management of Russian companies does not have a clear relation with the actual values and the dynamics of financial ratios. These results confirm the thesis that Russian business is not ready for open interaction with all stakeholders.

7 Conclusions

The originality of our research is the proposed empiric treatment to the analysis of financial information disclosed according to the principle of integrated reporting "reliability and completeness" in integrated or annual reports of Russian companies, combining elements of semantic and correlation analysis. Based on the analysis of 30 largest Russian holding entities, the conclusion is that in 70% of the reports the opinion of the management of Russian companies does

not have a significant relation with the actual values and the dynamics of financial ratios and does not comply with the principle of IR "reliability and completeness". We found for 30% of companies strong or moderate dependence of financial position with all groups of words, except legal or judicial, which justifies the conclusion that the number of words of this group is not significant for the purposes of forecasting the financial position. For 50% of companies, a strong or moderate correlation between the number of positive words and financial position is direct, and the number of negative words (40% of companies) is the reverse. Correlation with restrictive words is two-way, since it can be direct and inverse in equal proportion for different companies. It was also concluded that performance measurements are more closely related to all groups of words than financial indicators. This suggests that management pays more attention to describing financial results than describing the financial situation. In the future, this study will be deepened by a comparative analysis of two groups of companies with low and high scale of bankruptcy. The method of semantic analysis used in the research by distributing words from the texts of reports on certain groups that characterize the tonality of the text is not new. But, the possibility of adapting this methodology for the analysis of annual reports, which in the end will allow us to assess their compliance with the principles of IR, determines our contribution to the development of the prospects for open interaction of Russian business with all interested parties. In our opinion, companies need to use a single system of financial indicators to generate reports that can provide an objective view of the current situation of the company, as well as to unify the analytical interpretations of these indicators. This will eliminate the possibility of manipulating the opinion of stakeholders by the company management.

References

1. About the consolidated financial statements. Federal Low Russian Federation 27.07.2010 no. 208-OC
2. Caserio, C., Panaro, D., Trucco, S.: Management discussion and analysis in the US financial companies: A data mining analysis. In: Mancini, D., Dameri, R.P., Bonollo, E. (eds.) Strengthening Information and Control Systems: The Synergy Between Information Technology and Accounting Models, pp. 43–57. Springer International Publishing, Cham (2016)
3. Dougherty, C.: Introduction to Econometrics: the fourth edition. Oxford, New York (2011)
4. Efimova, I.: Analysis of sustainable development of the company: stakeholder approach. Economic Analysis: Theory and Practice 45, 41–52 (2013)
5. Frazier, K.B., Ingram, R.W., Tennyson, B.M.: A methodology for the analysis of narrative accounting disclosures. Journal of Accounting Research 22(1), 318–331 (1984)
6. Goel, S., Uzuner, O.: Do sentiments matter in fraud detection? Estimating semantic orientation of annual reports. Intelligent Systems in Accounting, Finance and Management 23(3), 215–239 (2016)
7. Loughran, T., McDonald, B.: When is a liability not a liability? textual analysis, dictionaries, and 10Ks. Journal of Finance 66(1), 35–65 (2011)

8. Loughran, T., Mcdonald, B.: Textual analysis in accounting and finance: A survey. Journal of Accounting Research 54(4), 1187–1230 (2016)
9. Malinovskaya, N.: Analysis of corporate reporting of Russian companies for compliance with the integrated reporting principles. Economic Analysis: Theory and Practice 45, 36–48 (2015)
10. Pisano, S., Alvino, F.: New european union's requirements and IFRS practicestatement management commentary: Does MD&A disclosure quality affect analysts forecasts? Journal of Modern Accounting and Auditing 11(6), 283–301 (2015)
11. Plotnikova, V.: The importance of the report on the opinion of executives on financial position and financial performance of the company for stakeholders. In: Proc. Int. Sci. Conf. Accounting, Management and Finance: Prospects for Development Under Economic Instability. pp. 138–142. Saratov Socio-Economic Institute of Plekhanov Russian University of Economics (2016)
12. Plotnikova, V., Shilovskaya, M.: Analysis of information disclosure in corporate reporting in terms of compliance with integrated reporting principle strategic focus and future orientation. Economic Analysis: Theory and Practice 16(5), 919–934 (2017)
13. Tennison, B.M., Ingram, R.W., Dugan, M.T.: Assessing the information content of narrative disclosures in explaining bankruptcy. Journal of Business Finance and Accounting 17, 391–410 (1990)