

Determination of the Degree of Development and the Impact of the Information Environment on the Formation of A System of Social Control in Procurement Under the Russian Contract System (Method of Content Analysis of Information Resources on the Internet)

N.A. Mamedova¹ and A.N. Baykova²

The article is a result of research on the effect of environment on the development of information system of social control in procurement. Informatization process of social relations largely determines how certain trends of social activity will be popular and durable. To determine the degree of maturity of information content on the theme of social control in procurement in Russia, the study used the method of content analysis of information resources. To improve the quality of research results were analyzed according to the semiotic and conceptual and thematic units of content.

Introduction

Content analysis is a method of research aimed at quantitative analysis of texts and text arrays for subsequent meaningful interpretation of numerical patterns identified. The primary mechanism of the method is the identification and measurement of the frequency of the use of formal or substantive components were analyzed in total individual text or data array. The degree of frequency of use in this sense becomes the desired indicator. The method is used to visualize the dynamics of the intuitions of the text content (information file) on a repetitive manner, estimates, opinions and other forms of expression in order to be able to organize these intuitive feelings, give them meaningful and reasoned explanation of the subject. Using the method involves the development of targeted approaches to the collection of data, representing the contextual textual evidence on which the feeling of repetition and repetition frequency based. However, the potential of the method of content analysis may be

¹ Moscow State University of Economics, Statistics and Informatics (MESI), Moscow
Email: nmamedova@bk.ru (first author)

² Moscow State University of Economics, Statistics and Informatics (MESI), Moscow
Email: mia-94@mail.ru

110 Determination of the Degree of Development and the Impact of the
Information Environment on the Formation of a System of Social Control
in Procurement Under the Russian Contract System
(Method of Content Analysis of Information Resources on the Internet)

represented by a much wider. Using the method, the researcher can not only streamline their understanding of the text (information files), but also to justify their conclusions, to interpret the author's position on the nature of used them formal elements or structures, revealing even more than the author would like to put into words. In this regard, the content analysis method called the "scientific method of reading between the lines."

Thus, a systematic approach to the study of the context, the desire for objectification of the data analysis of the text or information set is non-exclusive characteristic of content analysis and observed with the use of other methods of word processing. However, the establishment of quality parameters of the text (information files) through a quantitative measurement of the formal elements and establish the relation between the detected quantitative indicators is a characteristic feature of content analysis as a method of scientific knowledge in the methods of analysis and processing of texts. Therefore, the need to clarify the definition of the method. Content analysis is a systematic quantitative analysis, evaluation and interpretation of the form and content of information sources. This type of analysis is primarily quantitative analysis reveals quantitative regularities of repetition components of text, while in the analysis may reveal structural patterns, qualitative laws by classification, ranking and establishing a causal relationship between the obtained results of content analysis.

Over other methods of analysis and processing of text content analysis method has several advantages. For example, advantage is its adaptability, which manifests itself in the possibility of using the method without limitation as to the minimum and the maximum volume of the analyzed information. Processability method also evident in its effective integration with many other methods, in relation to which the method of content analysis, and can act as a primary, and as collateral. This method is most suitable for the primary processing of large volumes of information and adjusting the flow of information on the stage of collecting and processing raw data for both theoretical and applied research. Also indisputable advantage is the possibility of formalization and computerization of the process and the results of the content analysis, due to its characteristics as a quantitative method of analysis and text processing.

Results and Discussion

As mentioned above, the mechanism of the method of content analysis is to calculate the frequency of occurrence of certain components in the analyzed text or data arrays, which is complemented by the identification of qualitative relationships (statistical methods) and structural relationships (through the analysis of structural relationships between the components). The result of the method is considered to justify the existence of some analytes in quantitative and qualitative characteristics. Obviously, the effectiveness of the method is due to the choice of components, i.e., the choice of units of analysis. Requirements for the unity of content analysis are obvious enough. Firstly, it should be easily identified in the text. Secondly, a unit of content analysis should enable the semantic interpretation, that is to be interesting and useful content in the scale of the study. The tasks of the content analysis conducted for the purpose of the study were a content analysis on the structural-semiotic units (keywords) and the conceptual and thematic units. To conduct content analysis were used open sources of information and communication on the Internet. To clarify the parameters of the sampling information it was decided to use the data set generated by the search engines. When choosing a search engine we used the following criteria: a high degree of relevance of search results; optimal set of advanced search functions; low percentage of references to duplicate content; absence (very low percentage) relevance of links aimed at commercial sites. With all of these criteria best search engine was determined search system Google.

To conduct content analysis on the structural-semiotic units and specific to the research conducted as keywords accepted words the phrase "social control in procurement."

Public control of procurement is defined by the law mechanism for the rights of citizens, public associations, associations of legal entities to control the legitimacy of the state and municipal customers. At its core, social control in procurement is a form of interaction between civil society and the state to ensure the functioning of the legal order of the contract system in the procurement of goods, works and services for state and municipal needs. The objectives of social control in procurement include: a full submission of information, the development of civil society, reducing corruption, development of markets and trade

112 Determination of the Degree of Development and the Impact of the
Information Environment on the Formation of a System of Social Control
in Procurement Under the Russian Contract System
(Method of Content Analysis of Information Resources on the Internet)

relations. The implementation of these tasks is possible through the creation of an open information space in which interactions (functional level set), all participants of the contract system and society.

The legal basis for the functioning of social control in procurement in Russia is so basic pieces of legislation: Federal Law № 112-FL "On the basis of public control" and the Federal Law № 44-FL "On the contract system in the procurement of goods, works and services for state and municipal needs". Federal Law № 112-FL establishes the legal basis for the organization and implementation of public control over the activities of state authorities, local government, state and municipal organizations, other agencies and organizations engaged in accordance with federal laws separate public authority. In turn, the Federal Law № 44-FL regulates relations directed at providing state and municipal needs in order to improve efficiency, effectiveness of the procurement of goods, works and services, ensure openness and transparency of such purchases, the prevention of corruption and other abuses in the field of procurement. And as one of the types of control procurement establishes public oversight.

Thus, the legal framework of public control in procurement functions in Russia, however, the activity of the subjects of public control is insufficient. This study aims to assess the importance of social control in procurement by analyzing the information field and placed it open sources of information.

To evaluate the frequency of use of keywords in the information resources, the results of extended context search were distributed in accordance with the classification of information sources. In accordance with the above keyword structural-semiotic research source was defined thematic filter information resources (sites, portals). As classification criterion, it was decided to use the theme filter in combination with the degree of importance of the topic in the structure and content of web pages, the totality of which was formed by the results of the extended context search. Thus, information sources are classified as follows:

- site (portal), specializing in the topic;
- site (portal) wide profile information (main theme - procurement);

- site (portal) independent profile - specialization relating missing (news, panoramic, education);
- site (portal) government or municipal authority;
- site (portal) social organization, specializing on the subject of procurement;
- site (portal) social organization wide information profile;
- site (portal) reference and the legal system;
- links to educational materials, bills.

As the advanced search options in the search engine Google has taken the following mandatory parameters (Table 1).

Advanced search options on the structural-semiotic units

Name of the parameter	Basic / additional search terms	The value of the selected parameter
Select Keyword	Keywords	Public control procurement
Choice phrases	Phrases	Procurement for the needs
Search pages in the selected language	Language web page	In Russian
Search pages created in a particular country	Country of creating web pages	Russia
Search pages created or updated within the specified time	Date of creation / update	Any
Search by text, title or address of a page, as well as links to them	Location words	Anywhere on the page
Safe Search	Blocking inappropriate content	The function is activated
Search pages and files of a certain format	File Format	Either
Search for pages that are free to use, distribute and modify	Rights of use	With any license

Materials and methods

Content analysis on the structural-semiotic units.

To evaluate the results obtained by sampling information into account results in a certain time interval. Thus, the search results were classified according to the three stages of the legal regulation of the procurement activities in the Russian Federation: the information resources available and updated up to 2012; Information Resources posted and updated in the period 2012-2013; Information Resources posted and updated in 2014. Thus, the search covers a period of №94 federal law, the transitional period for the formation of the contract system (2012-2013) and the period of commencement of the contract system, regulated by federal law №44 (2014). The total number of information sources that match the search parameters, the source is 133. This number is the result of contextual search and does not include the hidden results (results are homogeneous (similar) presented the results of the search context).

Besides the classification of search results for the sources of information and limitation of accommodation (update) web page for evaluating advanced context sensitive searches used the following methods of ranking results on the basis of scoring. The first direction ranking - ranking Compliance common sense content of Internet pages (text) keywords, i.e. structural-semiotic unit conducted a content analysis. This approach allows taking into account the ranking parameter randomness use keywords. The second direction of ranking is ranking in order of importance of keywords in the content of the web page (text). Ranging in this direction allows considering setting the frequencies of keywords on a page (text). The order of distribution of GMAT results extended context of keyword search in accordance with a first direction of the ranking (in the parameter randomness use keywords) is defined the following:

- If the text is devoted to the topic of procurement, and the keywords used in the description of social control as the main or additional topics, and are not used in passing (in the list, listing, for specifying, when specifying links to legal documents, etc.) - assigned a score of three points;

- If the text is devoted to the topic of procurement, but the key words are used without description, detail that is casual (in the list, listing, for specification, when specifying links to legal documents, etc.) - assigned a score of two points;
- If no text is devoted to the topic of procurement, and key words are used without the description, detail that is casual (in the list, listing, for specification, when specifying links to legal documents, etc.) - assigned a score of one point.

The order of distribution of GMAT results extended context of keyword search in accordance with the second direction ranking parameter frequency of the use of keywords on a page is defined as follows:

- If the text is devoted to the topic of social control in procurement (permanently connected and used keywords - frequency of use of keywords high) - assigned a score of three points;
- If the text is partly devoted to the theme of social control, this topic is the subject of this section in the overall structure of the text (the frequency of the use of keywords is high or low throughout the text, but always use a high frequency section (part of the text) - assigned a score of two points;
- If the text is devoted to the topic of the conjugate with the subject of public scrutiny in procurement (the theme of fighting corruption, the development of the contract system, the activities of public associations, public policy and other topics); If keywords are used without specification, detailing, that is casual (in the list, listing, for specification, when specifying links to legal documents, etc.) - assigned a score of one point.

As a result of the application of the methods of classification and ranking of the sample to the resulting information by keyword "social control in procurement" produced the following results of content analysis on the structural-semiotic units:

1. Score from the structure of sources (indicator - the frequency of the use of keywords by source type information). The most frequently keywords "social control", "social control in

116 Determination of the Degree of Development and the Impact of the Information Environment on the Formation of a System of Social Control in Procurement Under the Russian Contract System
(Method of Content Analysis of Information Resources on the Internet)

procurement" are used in the texts of analytical research papers on the subject of public scrutiny and news reports in periodicals. The analysis showed that in these materials the theme of social control is the main. Social control in the framework of a contractual change legislation considered in the comparative analysis of №44 and №94 of the federal law on the Ministry of the Russian Government, in reference and legal systems, blogs. Sites of private companies inform legislative change contract system, in particular, attention is paid to public control. Keywords are also used in textbooks on marketing and logistics, but the subject of public scrutiny in procurement is not disclosed. Thus, periodicals, reference and legal systems, the sites of governmental and nongovernmental organizations are the main sources of information on the topic of social control in procurement.

2. Evaluation of data on prescription accommodation (update) information (parameter - the frequency of use of keywords). The theme of social control has gained considerable resonance after the reform of the legislation on the contract system in procurement. Basically, all the texts that reveal the theme of social control in the procurement and use of words containing contextual search, dated 2013, 2014 year. Most of the studied sources that use keywords at least three times, contain outdated information, the texts published before 2012.
3. Score from the dynamics of important topics (index - frequency of use of keywords in the data stream). In consideration of cumulative sources of social control is the theme of an independent object, when keywords are used more than 10 times. Public control is an additional theme when considering purchasing system where keywords are used from 1 to 10 times (the amount of text were taken into account). The use of keywords can also be random when the overall context of the text does not match the subject in question. Context analysis sources revealed that the subject of public scrutiny in procurement basically not an independent object of consumption (56.39% of the total number of search results). In the context of the theme of the contract system in the procurement of the theme of social control in the area of procurement used in 34.58 of the total number of search results. As an independent object use theme of social control is 9.02% of the total number of search results (according to the table 2).

Results of content analysis on the structural-semiotic units

Number of occurrences of keywords	Number of sources	The proportion of the total number of sources, %
0	75	56,39
1-10	46	34,58
over 10	12	9,02
Total:	133	100

4. Evaluation of changes in the flow of information (index - the quality of information). The theme of social control in procurement is the main source of the 28 examined 133 sources. Partially information devoted to public control in 45 sources. While noting the coincidence of the main content of the text in four cases (for example, the presentation of the text of the federal law №44. Remains a source of no significance for the research topics of social control and, in most cases, provided the information they are not devoted to the topic of procurement.

Content analysis on the conceptual and thematic units.

As for the previous direction of content analysis of the direction characterized by the use of keywords relating to ongoing investigations. At this stage, phase content analysis Keywords are the words of the phrase "social control." A narrower definition of the form of keywords defined in order to expand the search capabilities and take into account when analyzing the large number of results.

Conceptual-thematic units used in the content analysis of the words of phrases that are directly or indirectly associated with the keyword. They may have a direct association as disclosed subject of public scrutiny or accompany the implementation of mechanisms of social control in procurement practice. Indirect association occurs when the phenomena associated with the theme of social control in the area of procurement, are more extensive independent phenomenon. As a result, the sample

118 Determination of the Degree of Development and the Impact of the
Information Environment on the Formation of a System of Social Control
in Procurement Under the Russian Contract System
(Method of Content Analysis of Information Resources on the Internet)

conceptual and thematic units for conducting content analysis were formed by the following list:

- anti-corruption expertise;
- public examination;
- public discussion;
- public Policy;
- the fight against corruption;
- legal literacy;
- citizenship;
- civic engagement;
- public initiative;
- transparency of procedures;
- transparency.

This list of conceptual-thematic units was formed by interviewing through interviewing participants, simultaneously satisfying the following requirements: non-professionally with procurement; not members of associations and unions of legal entities; are not state or municipal employees. In the process of interviewing was formed by a set of associative phrases, while in the sample for the content analysis was selected phrases that have the highest number of repetitions of the participants.

Analyzing the list of conceptual and thematic units, it should be noted that the sample included units that characterize the mechanisms of social control - public debate, public examination, anti-corruption expertise, as well as units that characterize the necessary conditions for the exercise of social control - information openness, transparency procedures, and public initiative. Also in the sample units were related to social control, as a general private, ie units, which depends on the content and implementation of the function of social control - civil, activity, citizenship, legal literacy, public policy, the fight against corruption.

Thus, the final sample phrases for search are presented by keywords and the conceptual and thematic units. Aware of the fact that the number of results advanced context sensitive searches can vary over time, despite the installed options, you must record the results on a specific

date. This date for the study is conducted August 1, 2014. The causes of variation of search results are the following: adding information resources in the Internet; update pages of existing information resources, change the contents of the page; deleting pages of existing information resources.

Implementation of a content analysis on the conceptual and thematic units aims at the following objectives:

- definition of search results for keywords "social control" and each of the selected conceptual and thematic units (simple search - search for one phrase);
- definition of search results when combining keywords with each of the selected conceptual and thematic units (advanced search - search by two phrases);
- defining relations of the results of keyword search and the search results on the conceptual and thematic units (appraisal phrases popularity to organize a search query);
- the definition of (deviation) of the results of simple and complex search (evaluation of search results for the pair relevance).

In assessing the results of the extended context based search described above, the search parameters were used the following approach. The first approach is to compare the result of a keyword search, and search results on the conceptual and thematic units. Thus, each the result of associative combinations compared with the results of keyword search that allowed determining the ratio of the frequency of use of key words and concepts and thematic units. The second approach is the comparison of the results found for the conceptual and thematic units (the results of a simple search) with the results of a complex search (search for the pair relevance). Thus, each search result on each conceptual-thematic units compared with the corresponding result of a complex search (when the search query is generated using combinations of keywords phrases conceptual-thematic units). It is possible to determine the link between individual conceptual and thematic units and keywords.

Calculation of the ratio of the search result by keyword and conceptual-thematic units held by calculating the difference between the number of search results for keywords and the number of results for each

120 Determination of the Degree of Development and the Impact of the
Information Environment on the Formation of a System of Social Control
in Procurement Under the Russian Contract System
(Method of Content Analysis of Information Resources on the Internet)

conceptual-thematic unity. For the absolute value was taken the number of results based on keywords. At the date of August 10, 2014 the number of results extended context of keyword search "social control" was 1359 results. Calculation of the ratio of the results of simple and complex search is carried out by calculating the percentage of the result of a simple search to a result of a complex search. 100% was made the number of results found for each simple conceptual and thematic unity. Using the extended context search in the Google search engine, taking into account the above parameters search yielded the following results (Table 3).

The results of calculating the ratio of the search result by keyword and conceptual-thematic units can be interpreted as follows. The greater the deviation of the search result by the conceptual and thematic unity of the result of a keyword search, the more popular, more semantic isolation has the phrase conceptual-thematic units as compared to the keyword. That is, the frequency of use in the context of a search query phrases such high popularity due to the conceptual and thematic units, a wide scope of use. Analysis of results of the extended context of search (Table 3) showed that high semantic isolation have the following conceptual and thematic units "transparent procedures", "public policy", "public discussion", "information transparency", "public initiative". Average values of semantic isolation showed the following conceptual and thematic units, "citizenship", "civic engagement", "public examination". Low values of semantic isolation showed the following conceptual and thematic units, "anti-corruption expertise", "anti-corruption", "legal literacy". These results also show the degree of popularity of the conceptual and thematic units (within the parameters of the search) in information and communication on the Internet, compared with the popular keywords, and also among themselves.

The results of the extended context of keyword search and the conceptual and thematic units

№	Option search (keyword and conceptual-thematic units)	Search result (considering the received settings) (simple search)	Result paired relevance to the keyword (advanced search)	Evaluation of deviation (the ratio of the search result by keyword and conceptual-thematic units)	Evaluation of deviation (the ratio of the simple search result with the result of a complex search (search for the pair relevance))
1	2	3	4	5	6
1	public control	1350	—	0	—
2	anti-corruption expertise	652	153	- 698	23,5 %
3	public examination	2300	117	+ 950	5,1 %
4	public discussion	17100	402	+ 15750	2,4 %
5	public policy	22100	452	+ 20750	2,0 %
6	the fight against corruption	679	75	- 671	11,0 %
7	legal literacy	547	24	- 803	4,4 %
8	civil position	4100	74	+ 2750	1,8 %
9	civic activity	3660	85	+ 2310	2,3 %
10	public initiative	13400	178	+ 12050	1,3 %
11	transparency of procedures	30200	221	+ 28850	0,7 %
12	information transparency	13700	114	+ 12350	0,8 %

122 Determination of the Degree of Development and the Impact of the
Information Environment on the Formation of a System of Social Control
in Procurement Under the Russian Contract System
(Method of Content Analysis of Information Resources on the Internet)

The results of calculating the ratio of the results of simple and complex searches can be interpreted as follows. Application of a complex search query, comprising several phrases are always significantly narrows the search area and the number of search results, respectively, since the number of search results inversely proportional to the search terms. For example, the number of results simple search on the word "fairy tale" gives 23,900,000 results. Using complex search for "tales of Pushkin" is 1 300 000 results. The deviation in this case is 5.4%, despite the fact that a simple search for "Pushkin" shows the result in 13.9 million hits. The ratio of the result of complex and simple search can be very different, vary from 100% to close to zero, but the main value of establishing a relationship is that it gives definition to diagnose the connection between the search queries.

To conduct content analysis, we define the following relation between the result of complex and simple search. The smaller the amount of deviation of the results of complex search (pair relevance - two phrases) on the number of results on the appropriate conceptual and thematic unity (ie, the larger percentage), the more stable relationship is observed between the keywords "social control" and a separate conceptual and thematic unit. In this case the deviation shows that the number of results the search query in two phrases (keywords and conceptual-thematic unity) slightly corrects (narrows) the number of results a search query on a specific conceptual and thematic unity. Thus, a large number of coincidences between the results of simple and complex research suggests that in the information space, there is a significant number of pages in the content of which there are complex search phrases desired interconnected common sense.

Conversely, the greater the amount of deviation of the results of a complex search (steam relevance - two phrases) on the number of results on the appropriate conceptual and thematic unity (i.e., the smaller percentage), the less stable relationship is observed between the keywords "social control" and separate conceptual and thematic unity. This dependence shows that the result of a complex search in large communication distorts the results of a simple search on the conceptual and thematic unity. For the purpose of this content analysis of this result indicates that the number of pages of information and communication on the Internet, which are the same and have the common sense of the

phrase keywords "social control" and the phrase on a separate conceptual and thematic unity, slightly. This result demonstrates a low degree of correlation between complex search phrases.

Because the number of conceptual-thematic units is defined as the absolute number of the final sample, the analysis of the ratio of the complex and the simple search is carried out exclusively within the sample without relations with other possible search queries. Analyzing the results of the extended context search in the table 3, it should be noted that the total number of results regarding the high degree of dependence is observed for the conceptual and thematic units' anti-corruption expertise "(23.5%)," Fighting Corruption "(11.0% deviation) and "public examination" (5.1%). All other results show a deviation of 5% or less. Indicators with an average degree of dependence (from 5% to 2%) should be attributed to the following indicators conceptual-thematic units "legal literacy" (4.4%), "public discussion" (2.4%), "civic engagement" (2.3%), "public policy" (2.0%). Low degree of dependence (below 2%) was observed for the conceptual and thematic units "citizenship" (1.8%), "public initiative" (1.3%), "information transparency" (0.8%), "the transparency of procedures "(0.7%).

If we compare the results obtained by the definition of semantic isolation conceptual-thematic units of keywords and the results to determine the quality of the connection between the conceptual and thematic units and keywords, you can estimate the direction of the relation between the results (Table 4)

124 Determination of the Degree of Development and the Impact of the Information Environment on the Formation of a System of Social Control in Procurement Under the Russian Contract System (Method of Content Analysis of Information Resources on the Internet)

Correlation of the results found for the extended context of semantic degree of isolation and connection quality

The degree of semantic isolation results simple search	The results of calculating the ratio of the results of simple search	The results of calculating the ratio of the results of simple and complex search	The quality of communication between the results of simple and complex search
The high degree of sense of separateness	"Transparent procedures", "public policy", "public discussion", "information transparency", "public initiative"	"Anti-corruption expertise", "anti-corruption", "public examination"	High-quality communication
The average degree of sense of separateness	"Citizenship", "civic engagement", "public examination"	"Legal literacy", "public discussion", "civic engagement", "public policy"	Average quality communication
Low degree of sense of separateness	"Anti-corruption expertise", "anti-corruption", "legal literacy"	"Citizenship", "public initiative", "information transparency", "transparent procedures"	Low quality communication

Analysis of the results tables 4 to determine the presence inversely proportional relationship between the degree of isolation of conceptual-semantic unit's thematic content analysis and quality of communication between the key words and concepts and thematic units. One cannot speak of the absolute values of the inverse proportional relationship, but a definite trend correlation results advanced context sensitive searches traced. Errors occur due to errors of the order of distribution of the search results between levels (error application techniques comparison purposes), and also due to the error operation of search when determining the relevance of search results steam.

Establish the presence or absence of communication will allow conducting correlation and regression analysis. To do this, we need to calculate the correlation, linear regression to build and test the hypothesis according to two related quantities. In our case, the associated values are the results of a simple search and complex search results. Recall that the condition is a simple search advanced contextual search on the set parameters in one phrase (keyword "social control" and the conceptual and thematic units). In turn, the complex search condition

is advanced contextual search on the set parameters on the phrase pair, which consists of a keyword and a separate conceptual and thematic units. For correlation and regression analysis is not required to use the results of simple keyword search, analysis is carried out using only the results of a simple search on the conceptual and thematic units. Thus, there is a related sample of 11 pairs of values - the results of a simple search (X) and the results of complex search (Y). Required:

- calculate the correlation coefficient;
- test the hypothesis according to the random variables X and Y, with a confidence level of 98-99% (significance level - 0,02-0,01 respectively);
- to calculate the coefficients of the linear regression;
- build a scatter plot (correlation field) and the graph of the regression line.

Calculation of the correlation coefficient.

The correlation coefficient is a measure of the probability of mutual influence of two random variables. The correlation coefficient can take specific values (-1 to +1). If the absolute value is closer to 1, it indicates a strong connection between the values, and if closer to 0 - what it says about the weak coupling or its absence. If the absolute value of the correlation coefficient takes a value of 1, then we can talk about the functional connection between random variables, that is, when one value can be expressed by other means of a mathematical function. Compute the correlation coefficient can be as follows:

$$R_{xy} = \frac{M_{xy} - M_x M_y}{S_x S_y}$$

- (1.1) To calculate the correlation coefficient is necessary to make a table of values x_k^2 , y_k^2 and $x_k y_k$ (Table 5).

126 Determination of the Degree of Development and the Impact of the Information Environment on the Formation of a System of Social Control in Procurement Under the Russian Contract System (Method of Content Analysis of Information Resources on the Internet)

The values for the calculation of the correlation coefficient

Number of search options (k)	Context search options conceptual and thematic units	The result of a simple search (X)	The result of a complex search by relevance (Y)	X ²	Y ²	XY
1	2	3	4	5	6	7
1	anti-corruption expertise	652	153	425104	23409	99756
2	public examination	2300	117	5290000	13689	269100
3	public discussion	17100	402	292410000	161604	6874200
4	public policy	22100	452	488410000	204304	9989200
5	the fight against corruption	679	75	461041	5625	50925
6	legal literacy	547	24	299209	576	13128
7	civil position	4100	74	16810000	5476	303400
8	civic activity	3660	85	13395600	7225	311100
9	public initiative	13400	178	179560000	31684	2385200
10	transparency of procedures	30200	221	912040000	48841	6674200
11	information transparency	13700	114	187690000	12996	1561800
	THE SUM OF THE RESULTS	108438	1895	2096790954	515429	28532009

It is necessary to carry out calculations using formulas 1.2. and 1.3. The calculations are shown in Table 6.

$$M_x = \frac{1}{n} \sum_{k=1}^n x_k, M_y = \frac{1}{n} \sum_{k=1}^n y_k, M_{xy} = \frac{1}{n} \sum_{k=1}^n x_k y_k \quad (1.2)$$

$$S_x^2 = \frac{1}{n} \sum_{k=1}^n x_k^2 - A_x^2, S_y^2 = \frac{1}{n} \sum_{k=1}^n y_k^2 - A_y^2 \quad (1.3)$$

Calculation of quantities by the formulas 1.2. and 1.3.

№	The procedure for calculating	M_x	M_y	S_x^2	S_y^2	M_{xy}	$S_x^2 * S_y^2$	$\sqrt{S_x^2} * S_y^2$	R_{xy}
1	Calculation M_x ; M_y ; M_{xy} ; (dividing the sum of the results on the number of search options)	9036,5	157,9166666667			2377667,41666667			
2	Calculation M_x ; M_y ; M_{xy} ; (calculation of a square value M_x ; M_y ; M_{xy})	81658332,25	24937,673611111 2			5653302344278,36			
3	Calculation S_x^2 ; S_y^2 ; (dividing the sum of the square of the results on the number of search options and subtracting the square sum of the results)			93074247,25	18014,7430555556				

Thus, by calculating the correlation coefficient determined the random variables, which amounted to 0.734. The resulting value represents the presence of a stable connection between random variables. In our case diagnosed stable relationship between the results of a simple search and complex search results. However, since the estimate is made of the correlation coefficient for the final sample of random values and may therefore deviate from its general significance beyond the final sample of random variables, it is necessary to check the significance of the correlation coefficient. Thus, it is necessary to test the hypothesis of random variables depending on the final sample.

Testing the hypothesis of random variables depending on the final sample using the Student's t test of significance (t-test).

The random variable t follows t-distribution, so the table t-distribution should find the critical value of the criterion ($t_{kr,\alpha}$) for a given level of significance α (in our case $\alpha = 0,01$ or $0,02$). Calculate the random variable t-criterion can be defined as:

$$t = \frac{R_{xy} \sqrt{n-2}}{\sqrt{1-R_{xy}^2}} \quad (1.4)$$

If calculated according to formula 1.4. value t (in absolute value) will be less than the critical value of the criterion (found on the table of the t-distribution), the dependence between random variables X and Y is not. If on the contrary, we must conclude that the experimental data of the calculation of the correlation coefficient does not contradict the hypothesis about the dependence of random variables.

Testing the significance of the correlation coefficient (to test the hypothesis according to random variables) is shown in Table 7.

Verification made under the formula:

$$t = (R_{xy} * \sqrt{n-2}) / \sqrt{1-R_{xy}^2} \quad (1.5)$$

$$R_{xy} = 0,713474472 \quad n = 11$$

130 Determination of the Degree of Development and the Impact of the Information Environment on the Formation of a System of Social Control in Procurement Under the Russian Contract System (Method of Content Analysis of Information Resources on the Internet)

Test the significance of the correlation coefficient

Indicator	Correlation factor (R_{xy})	R_{xy}^2	Number of each options (n)	Number of degrees of freedom of Student's (n-2)	$\sqrt{n-2}$	Calculation $\sqrt{1-R_{xy}^2}$	Calculation t criterion
value of the index	0,7341645836	0,5389976358	11	9	3	0,6789715489	3,2438675147

Analyzing the results of testing the hypothesis according to random variables should be noted that with high probability the experimental data for the calculation of the correlation coefficient does not contradict this hypothesis. This means that the result of the calculation of the correlation coefficient of a stable correlation between random variables finite sample can be extended to a general correlation coefficient outside the final sample of random variables. To further study the nature of the relationship between random variables necessary to calculate the coefficient of the linear regression equation.

Calculation of the coefficient of the linear regression equation.

Linear regression equation is an equation of the line, about describing the dependence between random variables X and Y. If we consider that the value of X is free, and Y - a function of x, the regression equation should be written as follows:

$$Y = a + b * M_x \tag{1.6}$$

In this case, the value of X (the simple search results) is indeed free value, in turn, the magnitude of Y (complex search results) is completely dependent on the value of X as a basis of the search results is the complex combination of keywords and a notional focal units.

To calculate the coefficients of the linear regression - permanent regression coefficient (a) and variable regression coefficient (b), it is necessary to make calculations using formulas 1.7 and 1.8.

$$b = R_{xy} \frac{\delta_y}{\delta_x} = R_{xy} \frac{S_y}{S_x} \tag{1.7}$$

$$a = M_y - b * M_x \tag{1.8}$$

The procedure for calculating the coefficients of the linear regression is shown in Table 8.

Calculation of the linear regression equation: $Y = a + b * X$

Procedure for calculating the linear regression equation.

M_x	M_y	R_{xy}	S_x^2	S_y^2	S_x^2/S_y^2	$\sqrt{S_x^2/S_y^2}$	coefficient t a	coefficient b
903 6,5	157,91666 66667	0,73416 45836	930742 47,25	18014,7430 555556	0,00019 35524	0,01391 23108	0,01021 39258	65,61852 58786

Linear regression equation has the form: $Y = 65,6185258768 + 0,0102139258 * X$

To evaluate the prediction error is also necessary to make calculations, the procedure for calculating prediction error Y for a given value of x in accordance formulated 1.9 and 1.10 shown in the table.

$$\delta_{y/x} = \delta_y \sqrt{1 - R_{xy}^2} = S_y \sqrt{1 - R_{xy}^2} \quad (1.9)$$

$$\delta_{y/x} = \frac{\delta_{y/x}}{M_y} 100\% \quad (1.10)$$

An error estimate for the linear regression equation.

$$\text{Absolute error: } \sigma_{x/y} = \sqrt{S_x^2} * \sqrt{1 - R_{xy}^2}$$

$$\text{Relative error: } \delta_{y/x} = \left(\frac{\sigma_{xy}}{M_y} \right) * 100\%$$

The calculation of prediction error for a given value Y X

S_y	R_{xy}^2	$\sigma_{x/y}$	$\delta_{y/x}$
134,219011528	0,5389976358	91,1308901462	57,71%

Error equation:

$$\sigma_{x/y} = 91,1308901462$$

Further investigation of the dependence of random variables involves the construction of scatter plots (correlation of the field) and the graph of the regression line. The scattering diagram - a graphical representation of the corresponding pairs of random variables X and Y in the form of points in

132 Determination of the Degree of Development and the Impact of the Information Environment on the Formation of a System of Social Control in Procurement Under the Russian Contract System
(Method of Content Analysis of Information Resources on the Internet)

a plane with coordinates axes X and Y. In the same coordinate system and construct a graph of the regression line. A special feature is the choice of construction diagrams and scale of initial points on the axes, provides a visual diagram. The procedure and results of the calculation of the regression line points: point A with coordinates $(X_{\min}; Y_{\min})$ and point B with coordinates $(X_{\max}; Y_{\max})$ is shown in Table 9.

Verifying plotting the regression line is performed by applying a coordinate system point average values of the random variables X and Y coordinates $(M_x; M_y)$.

Calculation of the regression line and the points on the linear regression equation.

$$Y = 65,6185258768 + 0,0102139258 * X$$

Compute point on the regression line equation of the linear regression

X_{\min}	Y_{\min}	X_{\max}	Y_{\max}
547	71,2055433074	30200	374,0790860334

Coordinates of the regression line:

A (547;71) ; B (30200;374)

Scatter plot and graph the regression line shown in Figure 1.

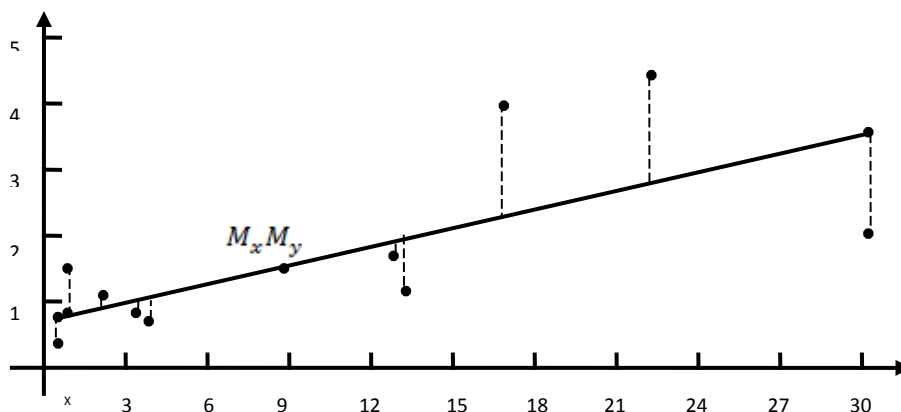


Figure 1. Scatter diagram (correlation field) and the graph of the regression line

The distribution of points in pairs of random variables shows that they correspond to the direction of the graph of the regression line, however, the degree of approximation to the scattering points of the regression line is not sufficient, which reduces the quality of the linear regression equation. However, validation plotting a linear regression on the coordinate system point average values of the random variables X and Y coordinates ($M_x; M_y$) showed perfect agreement with the coordinates of the point ($M_x; M_y$) generated from the linear regression line. Thus, the lack of proximity of the scattering points (coordinates of points on the results of a pair of simple and complex search) is due to an insufficient number of random variables finite sample.

Findings

The results of correlation and regression analysis confirmed the presence of a stable connection between the results of advanced context sensitive searches as part of the content analysis showed an association between the results of keyword search and conceptual and thematic units. It speaks to the uniformity of information flow, which brings together thematic pages on the Internet for all investigated units of content analysis. Thus, the purpose of the content analysis by determining the degree of development and the impact of the information environment on the formation and implementation of methods of social control in procurement under the Russian contract system achieved. The results of the content analysis can be used to further study the dynamics of changes in the interim results of content analysis to determine the areas of development, structure and content of the flow of information on topics related to information support of public control in procurement and contract system development.

134 Determination of the Degree of Development and the Impact of the
Information Environment on the Formation of a System of Social Control
in Procurement Under the Russian Contract System
(Method of Content Analysis of Information Resources on the Internet)

References

The Federal Law of 21.07.2014 № 212-FL "On the basis of public control in the Russian Federation"

The Federal Law of 05.04.2013 № 44-FL (ed. By 04.06.2015) "On the contract system in the procurement of goods, works and services for state and municipal needs"