

# Analytical Method for Social Network User Profile Textual Content Monitoring Based on the Key Performance Indicators of the Web Page and Posts Analysis

Victoria Vysotska<sup>1,2</sup>

<sup>1</sup> Lviv Polytechnic National University, S. Bandera Street, 12, Lviv, 79013, Ukraine

<sup>2</sup> Osnabrück University, Friedrich-Janssen-Str. 1, Osnabrück, 49076, Germany

## Abstract

The article considers the development of methods and software for processing web pages and user messages in social networks, blogs, or forums. A method of textual content support based on analysing information about the comments from web pages of social network users is proposed. This method is built on the principles of web analytics and uses analytical data to evaluate web resources. The analytical method of text content support uses the analysis of key performance indicators to form many keywords to increase the potential audience of the blog or e-commerce sites. Software tools for technical support of textual content have been developed. A method of designing and implementing systems for monitoring textual content of Internet blogs and Internet forums, which reflect theoretical research results was proposed. From the standpoint of a systemic approach, propose the application of the principles of web resources processing for the implementation of the life cycle of textual content, which allowed to develop of a method of content support. The main problems of functional services for managing a web page or profile of users of social networks, blogs, and forums for their further promotion in search engines and attracting a potential/permanent audience are analyzed.

## Keywords 1

Content, text content, Web resource, business process, content management system, content life cycle, Internet blog, Internet forum, social network, page conversion, e-commerce

## 1. Introduction

The analysis of visitor's business content is important, but few people use it as a guide to action in the management of Web resources or Web pages [1]. However, people are beginning to realize that you can significantly increase the revenue – double or even triple – if you just find out what traffic is more likely to drive conversions, what visitors do (and what they do not) on a particular Web site, and how to measure the effectiveness of the changes they make on the site to increase traffic conversion.

Objectives and Key Results (OKR) are about understanding business goals [4]. You need to first understand these goals, and only then deal with specific performance indicators of the Web site. This should be agreed upon, and should begin with the setting of the OKR [5-10]:

1. Make a list of stakeholders;
2. Conduct a "brainstorming" with stakeholders;
3. Define an OKR list (include anything that can be considered a success for Web site):
  - a. Generate more conversions that lead to sales;
  - b. Download more directories in PDF format;
  - c. Encourage the customer to purchase several products / services at the same time (thereby increasing the average cost of the order);

---

COLINS-2022: 6th International Conference on Computational Linguistics and Intelligent Systems, May 12–13, 2022, Gliwice, Poland

EMAIL: victoria.a.vysotska@lpnu.ua (V. Vysotska)

ORCID: 0000-0001-6417-3689 (V. Vysotska)



© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

- d. Contribute to the creation of a more recognizable brand or product;
  - e. Increase traffic;
  - f. Provide customer service (reduce the number of calls to the call center);
  - g. Build relationships with visitors (for example, to increase blog comments, forum posts, etc.);
4. Highlight and articulate OKR (focus on the most important 5-10 OKR).

## 2. Related works

Google Analytics is a free data collection and reporting tool [11]. However, it is not able to improve the site. To analyze and interpret the data of the reports, and then take action, you need a clear algorithm of actions and a coordinated team [12].

Most business organizations around the world use Key Performance Indicators (KPIs) to measure performance. These are sometimes referred to as Key Success Indicators (KSIs) or Balanced Score Cards (BCS) [13]. KPIs are used in the analytical departments of companies to assess the situation in business. Once an organization has established an OKR, it needs a way to evaluate the success of its activities. This assessment allows obtaining key performance indicators [4].

Similarly, in Web analytics, KPI is a Web metric that is essential to the success of an online organization. KPI requirements are [11-14]:

- In most cases, a KPI is a ratio, a percentage, or an average, rather than a processed number. This makes it possible to present data in context. Examples of raw data:
  - a. The website lost 15 orders yesterday because the server processing the online store order did not respond within  $M$  minutes;
  - b.  $N$  potential earnings were lost last week because the ordering system does not work for visitors who use Firefox;
  - c. Last month,  $L$  monetary units were spent on PPC keywords, which did not bring any conversion.
- The KPI must be time-bound. This emphasizes the changes and their speed.
- KPIs provide an incentive for action that is important to the business. Most parameters can be measured and evaluated, but this does not make them key to an organization's success.

Most of the complex work of preparing a KPI consists of defining the OKR. The key results used to set goals are KPIs. You just need to turn them into real Web-indicators that are available for social network, e-commerce, the circular economic etc. [15-18].

### KPI preparation algorithm

1. Determine the OKR. The KPI must meet the business objectives of the organization.
2. Convert OKR to KPI, i.e., determine specific Web-indicators that meet the business objectives of OKR (Table 1), such as server uptime, server response speed, notes on any offline companies that may affect the numbers, changes made to the site, information about the launch of new products or user reviews, etc. All this will help to better understand the data of the reports and, accordingly, increase their value.
3. Verify that KPIs are measurable and motivating indicators.
4. Create hierarchical KPI reports. You need to make sure that each recipient of the KPI report receives the data he needs. The more relevant the proposed information, the more attention and interest will be given to it.
5. Determine the partial KPIs. One of the most popular is to increase the conversion rate of the site. Usually, this indicator is easy to estimate, but at the same time, it is too contradictory - the visitor either carries out conversion, or not. For example, if the conversion is to upload a file, then the transition to the page may be a partial KPI. Similar partial KPIs include:
  - a. Go to the contacts page;
  - b. For a multi-page request form, fill in the first page;
  - c. Achieve the defining moment in the process of filling out the form;
  - d. Go to the promotion page;
  - e. Fill in the search box on the site.

6. Combine KPIs. After creating a list of required KPIs for each department representative, combine them and eliminate duplicates. The purpose of the KPI is to focus on important indicators for the business. If the KPI report presents all the key factors needed to assess success, then each KPI must be at least 10% of the total. If the importance of one KPI is much less than 10%, it must be removed or combined with another to get a more significant KPI.

**Table 1**

Example of converting OKR to KPI for social network, e-commerce, the circular economic etc. [19-27]

OKR department representatives	Proposed KPIs
Increase the number of visitors who come to the site from search engines; Sell more goods / services;	Increase the number of visitors who come to the site from search engines; The percentage of visits during which visitors add items to the cart; Percentage of visits during which visitors add items to the cart and place an order; The percentage of visits during which visitors interrupt the order after adding goods to the cart;
Increase the number of visitors who participate in the work of the site;	The percentage of visits during which visitors leave a comment in the block or upload a document; The percentage of visits during which visitors fill out a feedback form or click on a mail to link; Average time spent on the site per visit; The average number of pageviews per visit;
Sell more related products to customers;	Average cost of orders; The average number of goods per transaction;
Increase the positive customer experience of the site.	Percentage of visits during which only one page was viewed * bounce rate); Percentage of internal site searches that resulted in zero results; The percentage of visits that resulted in an application being submitted to support.

KPI reports are not something that does not change over time. They can and should change and evolve as departmental representatives learn to understand the performance of the Web site and take action to make changes. It is recommended that you review the KPI list at least quarterly. For example, an online marketer will obviously be interested in the difference between search engine visitors (SE) and non-search traffic, and how likely it is to convert that traffic (for example, a tour order may be a conversion). Example of KPI report for social network, e-commerce, web resources etc. [28-39]:

1. In month  $i + 1$  the profit of online orders decreased by  $x\%$  compared to month  $i$ .
2. Approximately  $y\%$  of all Web site visitors come from search engines.
3. For visitors from search engines the probability of entering the ordering system is almost  $k$  times higher than for visitors who did not come from search engines.
4. For visitors from PPC-systems the probability of entering the ordering system is  $a_1$ - $a_2\%$  higher than for visitors who found the site in the natural results of search results.
5. Since the ordering mechanism of the Web site does not support browsers other than *Name*, The website loses  $z_1$ - $z_2$  currency per month.

Actions to be taken by departmental representatives based on this KPI report:

- Check whether the reason for the decline in online profits is a seasonal factor that is characteristic of the entire industry or only for the online channel.
- At first glance,  $y\%$  of visitors who come to the site from search engines – can be a remarkably high figure. But is it the result of an effective search engine marketing strategy or are other channels just not working very effectively?

- At high  $a_1$ - $a_2$ % most of the budget of PPC-companies they work perfectly. However, perhaps good results here are caused by shortcomings in search engine optimization. Therefore, this question needs to be investigated in more detail. Although in the short term, it makes sense to increase the budget of PPC-companies.
- Create a better ordering mechanism that other browsers will support.

### 3. Methods and materials

As a result of segmentation of KPI-reports usually receive too much detailed information which cannot be used to give instructions to workers. For the developer of marketing strategy such information, on the contrary, is simply necessary. For example, of KPI-reports are built using the principle of hierarchy:

1. The retail director's marketing director needs hierarchical KPIs, for example:
  - a. Average conversion rate;
  - b. The average cost of the order;
  - c. The cost of attracting visitors.
2. The developer of the marketing strategy needs the same information, segmented by means (paid or in-kind search results, email marketing or banner display, etc.).

Segmentation using the in-depth data method is a great way to quickly understand the behavior of different segments of visitors. By identifying key segments of visitors who come to the site, you can create special profiles for them to make individual reports. Such separate segmented reports allow more detailed, fast and effective research of visitor behavior. Segmentation in most cases includes the type of visitor, the source of the transition or the geographical location of the visitor, for example:

- Examples of segments by visitor type;
- New and returning visitors;
- Customers and non-customers;
- New customers who are customers (or returning visitors who are customers);
- Examples of segments by source of visitors;
- Visitors who came as a result of the search (or not as a result of the search);
- Visitors who came for affiliate programs (or not for affiliate programs);
- Visitors who came only for paid search results;
- Visitors who came only for natural search results;
- Visitors who came only by e-mail;
- Examples of segments by geographical location;
- Only visitors from Lviv, only from Ukraine, Europe, etc.;
- Only regional visitors (Europe, Asia, Africa, Oceania, etc.).
- Ukrainian-speaking visitors (or visitors who use all other languages of the world).

When performing segmentation, it is necessary to find a balance between the clarity of information about visitor behavior and large amounts of data.

KPIs vary significantly depend on the e-business sector, such as retail, tourism, technology, B2B, finance, and so on. Even within the sectors, there are large differences, such as the sale of vouchers, tours or airline tickets, and the retail sale of souvenirs or clothing. Even when compared to competitors with the same goals, estimates can only be very approximate. The exact path that visitors must take to reach the goal, and their impressions of the process, will be different for each Web site. The biggest changes in these areas can have a significant impact on conversion rates. For example, retail managers want to distinguish between visiting existing customers and visiting non-customers. Therefore, the use of a standard industry-wide conversion rate can be misleading.

Conversion rates for e-commerce can be calculated in different ways.

$$K_{cv} = \frac{N_{cv}}{N_{vt}}, K_{cv} = \frac{N_{cv}}{N_{vr}}, K_{cv} = \frac{N_{cv}}{N_{vtb}}, K_{cv} = \frac{N_{cv}}{N_{vrb}}; \quad (1)$$

where  $K_{cv}$  is the conversion factor,  $N_{cv}$  is number of conversions,  $N_{vt}$  is the total number of visits to the Web-site,  $N_{vr}$  is the total number of visitors to the Web-site,  $N_{vtb}$  is the total number of visits to the

Web-site, during which the product was added to the cart,  $N_{vrb}$  is the total number of visitors to the Web-site, during which the product was added to the cart.

In the list below, you can also replace the conversion word with a transaction word. In other words, the visitor can make a purchase and, if he really liked the site and the organization of the purchase process, he returns to make an additional purchase during the same session. Depending on which Web analytics tool is used and what consents are accepted in the organization, this can be defined as one conversion with two transactions or two conversions with two transactions. For example, Google Analytics will show one conversion and two transactions since the visitor became a buyer, and this can only happen once during a visit. Other intra-site factors that significantly affect, and therefore may complicate, comparison is the following:

- Visibility of the Web site in search engines (natural and paid search results);
- Usability (convenience and ease of use) of the site (ease of navigation on the site, intuitive navigation system);
- Adequacy of displaying the Web site in all major browsers;
- The need for pre-registration / authorization for the purchase;
- Response time and page load;
- Quality of text and graphic content of the page;
- Use of trust factors such as shopping security logos, privacy policy, warranty, use of encryption for payment pages, customer recommendations, etc.;
- The presence of broken links or the absence of some images;
- Fast and accurate search on the site.

Examples of KPIs by roles in the organization

To illustrate the example, the online tourism website or online blogs/forum about tourism was chosen. Its business goals are twofold: to sell tours and try to get applications for professional services. To do this, the Web site has several main sections:

1. Online store section. Purpose is to sell tourist tours, the price of which is relatively high compared to the prices in most online stores in this area.
2. Section for generating requests for services. Purpose: to try to get visitors to apply for professional services (excursions, trainings, tips for exclusive tours with guides, consultations for individual tours). These are also expensive services.
3. Brand promotion section. It includes writing articles for blogs, which give positive practical advice on organizing recreation based on services and tours provided on the company's website.

In terms of roles on the Web site  $S_{tm}$ , there is an online store manager  $M_{is}$ , marketing manager  $M_{mr}$ , copywriter  $M_{cp}$  (employee who writes quality and effective unique content, i.e. author, journalist, copywriter of the content Web-site) and Web-master  $M_{vm}$ . For each calculate their own KRI:

$$S_{tm} = \langle M_{is}, M_{mr}, M_{cp}, M_{vm} \rangle. \quad (2)$$

The site with the online store probably has the most KPIs to choose from, as the main purpose (purchase) is easy to assess. The purpose of the site (to encourage visitors to add the product to the cart) is defined quite clearly. Google Analytics has a whole section dedicated to e-commerce reports. But most KPIs are better taken from other sections. The online store manager, includes additionally the number of visitors  $N_{vr}$  offered by the KPI:

$$M_{is} = \langle N_{vr}, S_{cc}, S_{co}, S_{cv}, S_{ro}, P_{nv}, I_{nv} \rangle, \quad (3)$$

where  $S_{cc}$  is average conversion rate,  $S_{co}$  is average cost of orders,  $S_{cv}$  is the average cost per visit or the average usefulness of the visit,  $S_{ro}$  is average ROI or average return on investment,  $P_{nv}$  is percentage of profits from new visitors,  $I_{nv}$  is index of new customers at the first visit - a new defined KPI.

$S_{cc}$  and  $S_{co}$  Google Analytics calculates in the e-commerce section. By default, it also calculates two types of values  $S_{cv}$  – the usefulness of the purpose of the visit (based on the usefulness of the goals) and the usefulness of the visit (based on the data of e-commerce transactions).

The formula for calculating the return on investment in Google Analytics:

$$ROI = \frac{Income - Expenses}{Expenses}, \quad (4)$$

where  $Income$  is profit,  $Expenses$  is costs.

A negative ROI indicates that you are losing money: the cost of attracting visitors is greater than the cost. Of course, when launching a new company, the ROI is likely to be negative until the number of returning visitors increases and the brand becomes recognizable, leading to an increase in conversions.

Of course, ROI is an indicator of efficiency for total gross profit. It does not consider what profit you get from the sale. It also does not take into account the number of transactions or visitors. For example, the ROI for a company (too centralized) may be high and the profit may be small. And with a lower ROI of a less specialized company, the profit can be quite large due to the large number of visitors. The formula for the rate of return is as follows:

$$RR = \frac{Income - Expenses}{Income} \quad (5)$$

Buying expensive goods, including tours, usually takes more time to think than buying cheap goods, such as souvenirs. This is usually equivalent to the number of visits required to persuade to make a purchase. Such KPI as  $P_{iv}$ , will allow to learn, whether it is characteristic of a site from Internet tourism. The value of the probability that new visitors will become new customers on the first visit will help to calculate  $I_{nv}$ .

$$I_{nv} = \frac{P_{tv}}{P_{nv}}, \quad (6)$$

where  $P_{tv}$  is the percentage of transactions from new visitors,  $P_{nv}$  is percentage of new site visitors.

Value  $I_{nv}=1$  suggests that new and returned visitors will become customers with equal probability. A value less than 1 means that a new visitor is less likely to become a customer than a returned one. And a value greater than 1 means that a new visitor will be more likely than a returning visitor.

Attracting good visitors to the Web site (those who generate sales or requests for services) is one of the main tasks of marketing. Online marketing includes the following sources: search engine optimization (free search engine rankings), PPC advertising (paid search results), banner advertising, affiliate networks, blog marketing, links from other sites, and email marketing.

To determine the best traffic, you need to analyze the conversion rate  $K_{cv}$ , company costs, earnings, and ROI. Therefore, the KPI for the marketer significantly intersects with the KPI for the online tourism manager. An important difference is that marketers pay attention not only to the conversion rate for purchases, but also to the conversion of goals, as this speaks to building relationships with visitors who are likely to eventually go to the purchase. If you omit such indicators as the total number of site visitors, the KPI for the marketer are as follows:

$$M_{mr} = \langle P_{vz}, K_{cz}, P_{vk}, P_{ck}, I_{ck}, S_{rk}, P_{vk}, P_{ob}, K_{vb}, I_{yk} \rangle, \quad (7)$$

where  $P_{vz}$  is the percentage of visits by type of tool through AdWords;  $K_{cz}$  is target conversion rate (as a percentage) by AdWords asset type;  $P_{vk}$  is the percentage of visits by business type through AdWords;  $P_{ck}$  is the percentage of conversion of business-type goals through AdWords;  $I_{ck}$  is goal conversion index by company type;  $S_{rk}$  is average ROI by company type;  $P_{vk}$  is the percentage of new and returning visitors;  $P_{ap}$  is percentage of new and returned buyers;  $K_{vb}$  is brand recognition ratio;  $I_{yk}$  is company quality index.

Company quality index  $I_{yk}$ . The new KPI is about assessing how well-targeted a company is, that is, how effective they are at attracting targeted traffic to an online tourism site.

$$I_{yk}(x) = \frac{P_{cv}(x)}{P_{vk}(x)}, \quad (7)$$

where  $I_{yk}(x)$  is company quality index function (for company  $x$ ),  $P_{cv}(x)$  is the function of determining the selection of conversion prices for evaluation from the company  $x$ ,  $P_{vk}(x)$  is the function of determining the percentage of evaluation from the company  $x$ .

Let us say, for example, that  $P_{vk} = 50\%$  of AdWords visitor ratings, but this company source is only responsible for  $P_{cv}=20\%$  of conversions. This company works inefficiently, because if two companies are equally targeted and each generates 50% of traffic, then two should give 50% of conversions. If one time will be effective for another, generating more of its own conversion rate, then, by definition, this company is the best targeted.

An index value of  $I_{yk} = 1.0$  means that a visitor from this company will convert with the same probability as a visitor from any other company. A value of  $I_{yk} < 1.0$  means, respectively, that a visitor from this company is less likely to convert than a visitor from any other company. And the value  $I_{yk} > 1.0$ , respectively - the visitor will convert more likely than the seller from any other company (Table 2). The company Forum is a very well targeted company. Company Yahoo! Organic is also well

targeted, but the number of conversions is quite low, so you do not need to pay attention to it until more data is collected. You can waste a lot of time and effort finding the reason why visiting with Yahoo! Organic is almost three times more efficient than Google organic. Although, the statistical sample is too small.

**Table 2**

Convert for different company

Company	$P_{vk}$ , % visits	$P_{cv}$ , % conversions	$P_{ak}$ , % all conversions ( $P_{cv}/N_{cv}$ )	$I_{yk}$ ( $P_{ak}/P_{vk}$ )
Forum	2,02	51,00	19,32	9,56
Google cpc	4,90	11,00	4,17	0,85
Google organic	40,84	83,00	31,44	0,77
YSM ppc	3,62	3,00	1,14	0,31
Yahoo! Organic	0,56	3,00	1,14	2,04
Referral	29,59	73,00	27,65	0,93
Direct	16,22	39,00	14,77	0,91
Other	2,25	1,00	0,38	0,17
Total N	$N_{vk}=100,00$	$N_{cv}=264,00$		

Brand recognition ratio:

$$K_{vb} = \frac{N_{bq} + N_{tv}}{N_{aq} + N_{tv}}, \quad (8)$$

where  $N_{bq}$  is the number of search queries with the brand name;  $N_{tv}$  is number of direct visits to the Website;  $N_{aq}$  is total number of search queries.

Note that in the formula, search queries are keywords that are entered into search engines. Direct visits are included because they are made by people who know the address of the Web site, which means that employees exclude the brand from the reports of visits to the Web site. For site content developers, the main goal is to maximize audience engagement. How much time people spend reading the content of the site and how much of it they read - these are the key indicators of assessing audience engagement.

There are three categories of content in Web sites:

- Product and organization information from corporate information sites; product reviews sites, blogs, technical support sites, online training sites, etc.
- Advertising content from sites with free content that earn revenue from the sale of advertising (banners or text ads), which are placed alongside other content on the site.
- Subscribed content coming along with advertising revenue can offer subscription, i.e. the user pays a subscription to receive materials (possibly more complete versions of articles).

Regardless of the business model of the content site, increasing the interest of visitors is a key factor for success. That is why content managers are always looking for ways to include additional topics in each article or page to increase that interest. Accordingly, for content sites, the number of visits per day, week or month is an important KPI.

$$M_{cp} = \langle S_{tp}, S_{pv}, P_{vv}, S_{nc}, P_{zv}, P_{av}, K_{vb}, P_{ht}, P_{st}, P_{lt} \rangle, \quad (9)$$

where  $S_{tp}$  is average length of stay on the site through AdWords;  $S_{pv}$  is the average number of pageviews per visit through Google Analytics;  $P_{vv}$  is bounce rate (as a percentage), for example for one page as;  $S_{nc}$  is average number of ad clicks per  $N_{vr}$  visit;  $P_{zv}$  is the percentage of interest of visitors;  $P_{av}$  is the percentage of new visitors and those who returned through Google Analytics;  $K_{vb}$  is brand recognition ratio;  $P_{ht}$  is the percentage of repeat user visits that occurred since the previous visit in less than days  $t_1$  (calculated according to Google Analytics);  $P_{st}$  is the percentage of repeat user visits that occurred since the previous visit between  $t_1$  and  $t_2$  days at  $t_1 < t_2$  (calculated according to Google Analytics);  $P_{lt}$  is the percentage of repeat user visits that occurred from a previous visit of more than  $t_2$  days (calculated according to Google Analytics).

The average bounce rate (as a percentage)  $P_{vv}$  can be found through Google Analytics. From the point of view of the author of the content, high value  $P_{vv}$  means low interest of visitors, i.e. weak interest

in the site. Segmentation in this case is the most important condition for making information decisions. One-page bounce rate:

$$P_{vvp} = \frac{N_{vnp}}{N_{inp}}, \quad (10)$$

where  $N_{vnp}$  is the number of one-page visits for this page via Google Analytics;  $N_{inp}$  is the number of times users visit this page directly through Google Analytics.

Average ad clicks per  $N_{vr}$  visit:

$$S_{nc} = \frac{N_{cr}}{N_{av}} \cdot N_{vr}, \quad (11)$$

where  $N_{cr}$  is the average number of clicks on AdWords advertising;  $N_{av}$  is total number of visits via Google Analytics;  $N_{vr}$  is number of visits for analysis.

Visitor interest rate:

$$K_{zv} = \frac{N_{ad}}{N_{av}}, \quad (12)$$

where  $N_{ad}$  is the total number of actions on the site through AdWords;  $N_{av}$  is total visits through Google Analytics.

Percentage of visitors interested:

$$P_{zv} = \frac{N_{zv}}{N_{vk}}, \quad (13)$$

where  $N_{zv}$  is the total number of interested visitors through AdWords;  $N_{vk}$  is total number of visitors through Google Analytics.

To analyze revisits related to repeat visits, you need to choose the ideal time intervals for a particular e-business model  $t_1 < t_2$ . With a successful e-business:

$$P_{ht} \gg P_{st} \gg P_{lt}. \quad (14)$$

This is usually impossible to achieve. But in the periodic study of these indicators, can deduce patterns for adjusting the content, which in turn improves the ratio at least as  $P_{ht} \geq P_{st} \geq P_{lt}$ .

Webmasters are responsible for the efficient and continuous operation of the website. Therefore, they need to know what the load on the servers will be, i.e. how many visitors can expect to see the server. You also need to know which browsers and language settings users use most often.

$$M_{vm} = \langle K_{du}, P_{ul}, P_{ub}, P_{us}, P_{ur}, P_{up}, P_{um}, P_{ep}, K_{is} \rangle, \quad (15)$$

where  $K_{du}$  is indicator of the number of visitors, visits and page views;  $P_{ul}$  is the percentage of visitors who support English / Ukrainian;  $P_{ub}$  is the percentage of visitors who use a particular browser through Google Analytics;  $P_{us}$  is an proportion of visitors who use a specific operating system;  $P_{ur}$  is the percentage of visitors who have a high, medium / low screen resolution;  $P_{up}$  is the percentage of visitors who have a high-speed Internet connection;  $P_{um}$  is the percentage of visitors who have a high-speed Internet connection;  $P_{ep}$  is the percentage of pages published with an error;  $K_{is}$  is internal search indicator.

The metric  $N_{du}$  is basic for Webmasters and is determined through Google Analytics:

$$N_{du} = \langle S_{nvt}, S_{nut}, S_{npt}, S_{npv} \rangle, \quad (16)$$

where  $S_{nvt}$  is average number of visits for a certain period of time;  $S_{nut}$  is the average number of unique visitors over a period of time;  $S_{npt}$  is the average number of page views for a given period of time;  $S_{npv}$  is the average number of pageviews per visit.

Percentage of error pages that should be minimized:

$$P_{ep} = \frac{N_{ep}}{N_{pp}}, \quad (17)$$

where  $N_{ep}$  is the total number of pages published with an error;  $N_{pp}$  is the total number of pages viewed.

Internal search metrics are also determined through Google Analytics:

$$K_{is} = \langle P_{uis}, S_{vrs}, P_{uos}, P_{uns}, T_{svs}, S_{nup}, P_{uur}, P_{unr}, P_{ucs}, P_{pop}, P_{cus}, P_{bus}, P_{ksp}, P_{uts}, N_{nns}, P_{nrp} \rangle, \quad (18)$$

where  $P_{uis}$  is the percentage of visits that use site search;  $S_{vrs}$  is the average number of search results viewed per search;  $P_{uos}$  is the percentage of visitors who left the site after viewing the search results;  $P_{uns}$  is the percentage of visitors who conduct multiple searches on the site during the visit (excluding multiple searches for the same keyword);  $T_{svs}$  is the average time spent on the site to visit after the search;  $S_{nup}$  is the average number of pages viewed by visitors after the search;  $P_{uur}$  is the percentage of visitors who use the site search;  $P_{unr}$  is the percentage of visitors who do not use site search;  $P_{ucs}$  is



the percentage of conversions from visitors using site search;  $P_{pop}$  is the percentage of failures after visiting one page as a search result;  $P_{cus}$  is the percentage of buyers among visitors who use site search;  $P_{bus}$  is the percentage of purchases made among visitors using site search;  $P_{ksp}$  is the percentage of visitors that view more than  $k$  pages after a search;  $P_{uts}$  is the percentage of visitors who spent more than  $t$  time on the site after the search;  $N_{nns}$  is the number of zero search results on the site;  $P_{nrp}$  is the percentage of zero search results on the site pages, in particular,

$$P_{nrp} = \frac{N_{nps}}{N_{vps}}, \quad (19)$$

where  $N_{nps}$  is the total number of zero-page search results;  $N_{vps}$  is the total number of search pages viewed.

To calculate for further analysis  $K_{ps}$ - the rate of use of site search:

$$K_{ps} = \frac{N_{sv}}{N_{ns}}, \quad (20)$$

where  $N_{sv}$  is the number of visits with site search;  $N_{ns}$  is number of visits without site search.

As the number of e-commerce sites that support RIA technology (Rich Internet application) increases, so does the need to define KPIs for them.

Rich web application (rich Internet application RIA or installable Internet application) is a Web application that has many features of traditional software. The concept is closely related to a one-page program and can allow the user to interact with features such as dragging, background menu, WYSIWYG editing, and more. HTML5 is the modern standard for developing advanced web applications that are supported by all major browsers. Typically, a RIA system:

- Runs locally in a security environment - "sandbox" - a mechanism for safe execution of programs
- Runs in the browser and does not require additional software installation;
- Passes the required part of the user interface to the Web client, leaving most (program resources, data, etc.) on the server.

Analysts should not think in terms of  $K_{du}$  pageviews, but in terms of  $K_{as}$  actions and events that indicate user / visitor / customer interactions with the site. That is, according to the results of the analysis of indicators, it is necessary to redefine  $K_{as}$  the set of actions to be performed by visitors, so that it is considered an interaction with the site, in particular,

$$K_{as} = \langle P_{nv}, P_{uv}, S_{nv}, S_{tv} \rangle, \quad (21)$$

where  $P_{nv}$  is the percentage of new site visitors;  $P_{uv}$  is the percentage of unique visitors;  $S_{nv}$  is average number of views per visit;  $S_{tv}$  is average duration of visit;  $S_{cc}$  is the average conversion rate.

Combining the KPIs of interaction with the  $K_{du}$  site and  $K_{as}$  event tracking, you can define the following KPIs:

$$K_{usa} = \alpha(K_{du}, K_{as}) = \langle P_{vcu}, P_{sau}, P_{siu} \rangle, \quad (22)$$

where  $P_{vcu}$  is the percentage of visitors interacting with different types of content presentation, such as zooming, panning, viewing the next communication;  $P_{sau}$  is the percentage of visitors who run various kinds of the events, such as lose, pause, next, rating, click on the ad;  $P_{siu}$  is the percentage of interaction with the site, i.e. the execution of certain actions, such as subscribe, register, comment, determine the rating, add to favorites.

Defining different KPIs allows to focus on those elements of the online strategy that are most effective for attracting visitors, generating conversions, conversions, and e-commerce profits. It will also help determine the optimal structure of the website to improve the efficiency of its use and increase the volume of regular visitors and customers. Thus, you can identify many ineffective web pages. When analyzing data about visitors, it is necessary to optimize the pages of the site for the effectiveness of users on it. In many cases, you can improve your site, for example, by fixing broken links, changing login URLs to effectively visit the pages you want, or adjusting the content of the page to deliver the required advertising message.

The algorithm for identifying problem areas of the site structure for further optimization is as follows:

1. Identify the set of ineffective web pages through the analysis of their usefulness.
2. Determine the set of popular login pages through the analysis of failure rates.

3. Analyse the entry sources (search engines, paid advertising, links in e-mails, links to other sites, direct access to the address, for example, from the history of previous visits of user or the first direct visit).
4. Analyse the login keywords.
5. Visualize the site conversions by the user to achieve the goal.
6. Evaluate the success of the search on the site

To identify many ineffective pages with web analytics tools by analyzing the list of indicators:

- The value of the measure of usefulness of the page  $\$I_{dx}$ ;
- Many of the Top Landing and Exit Pages;
- Funnel Visualization tree.

The usefulness of the page is calculated as

$$\$I_{dx} = \frac{R_{cv} + R_{ec}}{N_{upv}}, \quad (23)$$

where  $R_{cv}$  is usefulness of the purpose of the visit (based on the usefulness of the goals) and usefulness of the visit (based on the data of e-commerce transactions);  $R_{ec}$  is income from e-commerce;  $N_{upv}$  is unique page views.

That is, if page  $x_i$  is viewed by visitors who reach the goal, then the usefulness of this goal increases the usefulness of page  $x_i$ . The more often the  $x_i$  page is viewed by visitors who reach the goal, and the higher the usefulness of the goal, the greater the value of  $\$I_{dx}$ . This method of evaluating the usefulness of pages has nothing to do with goals and conversions. Ranking pages by the value of  $\$I_{dx}$  sets the order of their optimization. Unexpected pages in this set (which are not related to the goals) indicate a problem with the structure and content of the website.

In parallel, you need to analyse the most popular pages. The main value in the analysis of many popular pages is the failure rate; if visitors get to the login page  $x_i$  and immediately leave the site, it is a sign of low user involvement in the site. If the  $x_i$  login page has a high bounce rate, then the content of the  $x_i$  page does not meet user expectations. It is necessary to investigate the source of conversions both in the middle of the site, and conversions from other sources. The statistics of low indicators of the last causes to intensify work depending on transitions in the following directions:

- Search engine optimization (SEO);
- Campaigns with paid search results;
- Offline / Online marketing activities;
- Advertising and maintaining pages on social networks.

Keyword analysis is a real market research, i.e., visitors report what content they expected to receive when visiting the site. Visualization of conversions on the site by the user to achieve the goal will assess the problem areas of the site structure, where the potential visitor / buyer faces problems, such as incorrect or unclear or difficult stages of payment / ordering. Site search is an internal search engine that visitors often replace with a site or menu navigation system. For large websites with hundreds or thousands of pages of content, the internal search engine is an important component for visitors to quickly find the content they need. Internal search engines typically use the same architecture and mechanisms as external search engines like Google. Evaluation of the success of the search on the site is to analyse the failure rate, as well as a few other indicators, in particular:

- Conversion rate achieved  $P_{cv}$ :

$$P_{cv} = \frac{N_{cv}}{N_{vt}} \cdot 100\%, \quad (24)$$

where  $N_{cv}$  is number of conversions,  $N_{vt}$  is number of visits.

- Income indicator  $P_{ip}$ :

$$P_{ip} = R_{cv} + R_{ec}, \quad (25)$$

where  $R_{cv}$  is usefulness of the purpose of the visit,  $R_{ec}$  is usefulness of e-commerce.

- Indicator of average utility  $S_{cv}$ :

$$S_{cv} = \frac{R_{cv} + R_{ec}}{N_{cv} + N_{tr}} \quad (26)$$

where  $R_{cv}$  is the usefulness of the goal,  $R_{ec}$  is usefulness of e-commerce,  $N_{cv}$  is number of conversions,  $N_{tr}$  is number of transactions.

- Conversion rate in e-commerce  $R_{ecc}$ :

$$R_{ecc} = \frac{N_{tr}}{N_{vt}} \cdot 100\% \quad (27)$$

where  $N_{tr}$  is number of transactions,  $N_{vt}$  is number of visits.

- An indicator of the usefulness of the visit  $P_{uv}$ :

$$P_{uv} = \frac{R_{cv} + R_{ec}}{N_{vt}}, \quad (28)$$

where  $R_{cv}$  is the usefulness of the goal,  $R_{ec}$  is usefulness of e-commerce,  $N_{vt}$  is number of visits.

A visitor who uses site search is several times more valuable than a visitor who does not use the site search. Therefore, the development and development of site search service effectively affects the performance of site visits by increasing the volume of a regular audience.

To do this, use the calculation of the impact on income of the search function on the site  $I_{ssp}$ :

$$I_{ssp} = (R_{ssv} - R_{snv}) \cdot N_{ssv}, \quad (29)$$

where  $R_{ssv}$  is usefulness of visiting with site search,  $R_{snv}$  is usefulness of visiting without searching the site,  $N_{ssv}$  is number of visits with site search.

This indicator regulates plans and strategies for further investment in the development of site search service. This figure should be 80% of the monthly revenue for the website.

Search Engine Marketing (SEM) Optimization Algorithm is as follows:

1. Keyword research (for paid / unpaid search results).
  - a. Visitors who came according to natural search results.
  - b. Visitors who use internal site search.
2. Campaign optimization (paid search results).
3. Login page optimization and SEO (search engine optimization) (for paid / unpaid search results).
4. Optimize your ad positions for your AdWords campaign (paid search results):
  - a. Optimization of positions per visit (pages / visits, average length of stay on the site);
  - b. Position optimization by percentage of new visits (bounce rate, conversion rate achieved for goal 1 [for goals 2-4], conversion rate achieved, [profit, transactions, average utility, e-commerce conversion rate, visit utility]);
  - c. Optimization of positions for the usefulness of the visit;
  - d. Daytime optimization in AdWords;
5. Optimize ad versions for your AdWords campaign (paid search results).

For keywords that drive conversions, you need to optimize your investment by setting a maximum cost-per-click (CPC) in AdWords. The amount of return on investment (ROI) must be positive, ie the income received must exceed the costs, i.e.:

$$ROI = \frac{Income - Expenses}{Expenses} \cdot 100\% > 0, \quad (30)$$

where  $Income$  – profit,  $Expenses$  – costs.

The ROI for gross profit is

$$ROI_{vp} = \frac{(Income \cdot A_p)/100 - Expenses}{Expenses} \cdot 100\%, \quad (31)$$

where  $A_p$  is the amount of profit.

That is, you can calculate how many percent ( $q\%$ ) more money should be allowed to spend on a particular keyword in AdWords, without the risk of getting a negative ROI. At the start of a campaign, ROI may be negative as long as the brand and website are unknown. Visitors typically need several visits to a new website before they can convert. But such a situation (with a negative ROI) can only be acceptable for a short period of time - about a few weeks, depending on the situation.

To calculate the maximum amount that can be spent on attracting visitors - the maximum cost of attracting ( $C_{amax}$ ), you must use the formula:

$$C_{amax} = \frac{Income \cdot A_p}{\frac{100}{ROI_{vp}} + 1} \quad (32)$$

Knowing the conversion rate for each keyword, you can now calculate the maximum cost-per-click ( $C_{max}$ ) for that keyword.

$$C_{cmax} = C_{amax} \cdot \frac{R_{ecc}}{100} \quad (31)$$

The result of this system is that you do not have to overpay for AdWords keywords.

Keyword topics are a term used in search engine marketing to describe a set of keywords that accurately describe the content of a page. Properly defined keyword topics for search engines significantly improve the effectiveness of user visits because of search.

Typically, topics usually contain 5-10 phrases per page, in which keywords intersect. More than ten common phrases weaken the impact and effectiveness of the page - in terms of user experience and search engine rankings. If you already have a page with more than 10 keyword phrases, it is best to create a separate page dedicated to additional keywords. Basic tips:

1. It is always necessary to put the interests of visitors and customers first.
2. For campaigns, you need to use special login pages - for visitors who came for both paid and unpaid search results.
3. Login pages should be next to the call to action.
4. Website content should be built around the topic of keywords with 5-10 keywords and intersecting phrases.
5. You should place keyword-rich content closer to the top of the page.
6. You should use keywords in HTML tags <title>.
7. You need to use keywords in anchors, i.e., in HTML-tags <a>.
8. Avoid placing text in images, Flash, or other embedded content.
9. You must use a robots.txt file to control which pages should be indexed by search engines.
10. Do not abuse keywords and do not spam search engines.

An algorithm for site promotion and calculation of its efficiency is:

1. Assign utility to goals.
2. Activate e-commerce reports.
  - a. Define an unlimited number of goals (standard number - 4 goals for each profile).
  - b. Determine the amount of time and visits a user needs to convert.
  - c. Investigate the contribution of each goal (product) to the overall revenue of the website.
  - d. Group goals by category.
  - e. Generate lists of individual transactions as individual targets.
3. Track non-commercial content of the site as elements of e-commerce (downloading pdf-files, images, etc.).
4. Track offline marketing activities or offline visitors.
  - a. Prestigious URLs - in the case of a well-known brand, all web content should be hosted on one central domain.
  - b. Coded URLs - in the case of a well-known brand or if the products already have separate websites.
  - c. Combination with search - brand awareness is less than product or service awareness, or the target audience is more price-oriented than brand-oriented.

## 4. Experiments, results, and discussion

For the detailed textual content analysis and monitoring in social network user profile or the Web site, such as the online blog or forum and online forum, eight different systems were developed and implemented, respectively, each supporting different number of stages of the content life cycle. That is, not all components have been developed for different implemented systems or subsystems of Web resources processing as content creation; management and maintenance have not been developed at all. Table 3 presents a list of implemented Web sites with indication of the availability of implemented subsystems of Web resources processing with the textual content life cycle support. The table 4 presents the results of the developed systems according to Google Analytics. The analysis of the results of text content support allows to determine the reasons for the formation of the target audience using a set of characteristics of the operation of Web resource. By regulating the thematic set of text content, its uniqueness, efficiency of its formation and adequate management according to the individual needs of the regular user, you can model the boundaries of the target social audience and the number of unique

visitors from search engines. Fig. 1-3 presents the results of the developed systems in the form of graphs, which show that in the presence of all stages of the content life cycle significantly increases the number of visits and unique users.

**Table 3**

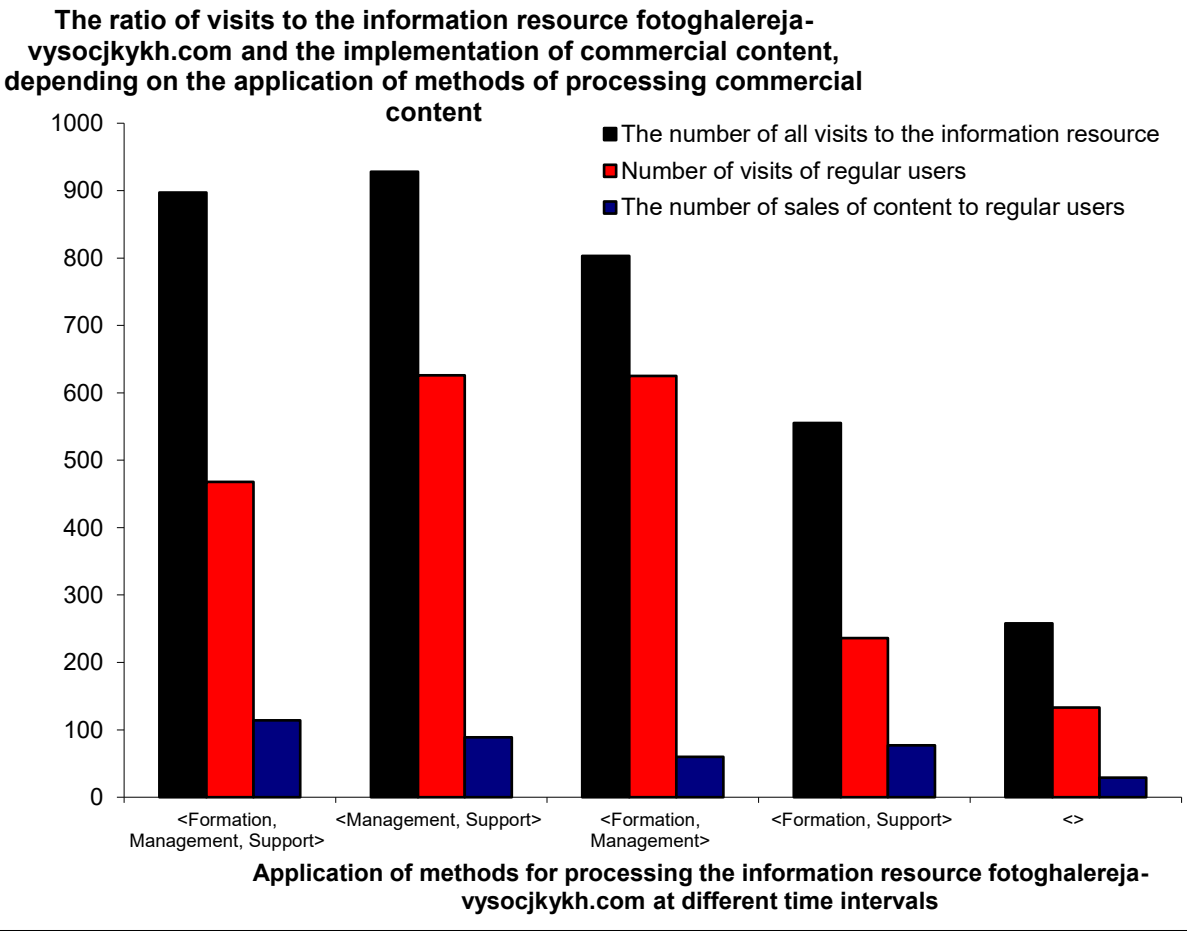
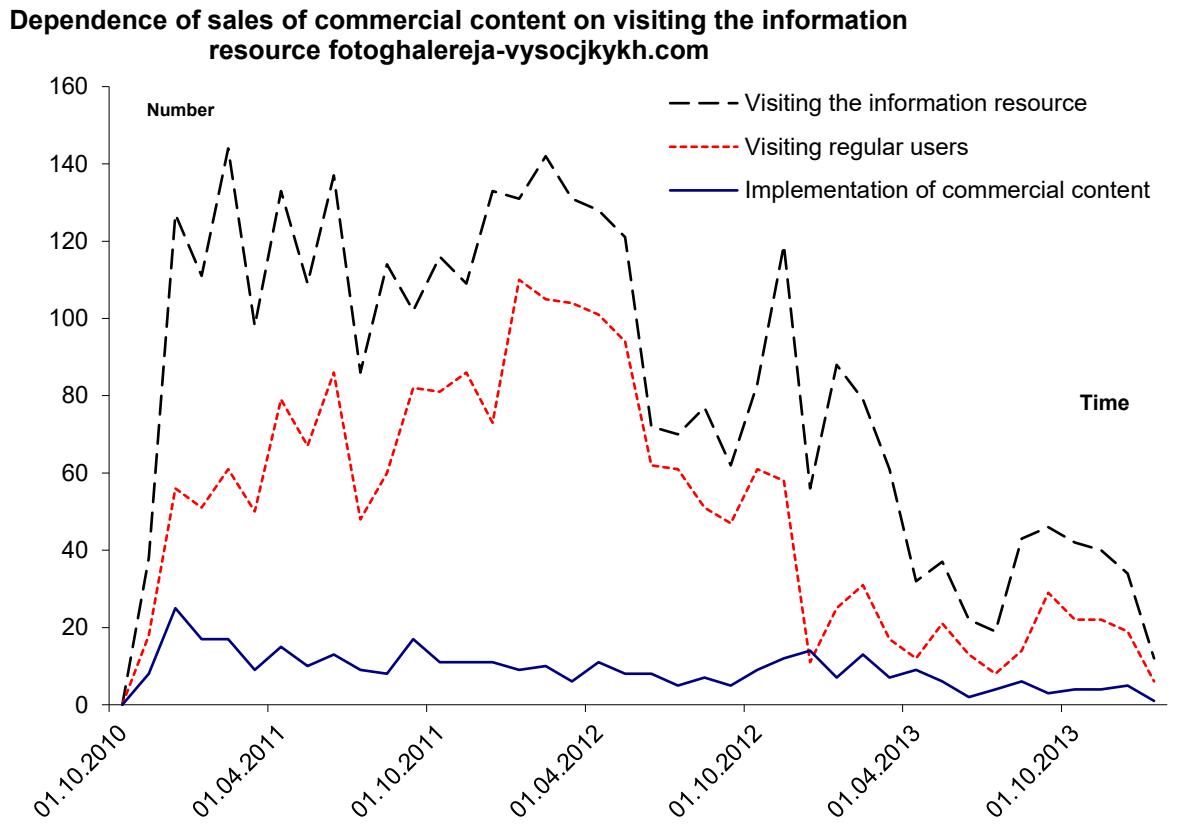
Implementation of stages of Web resources processing in the developed systems

No	Resource address	Type	formation	management	support
1	fotoghalereja-vysocjkykh.com	forum	+/-	+	+/-
2	vgolos.com.ua	forum	+	+	+
3	victana.lviv.ua	blog	+/-	+	+
4	tatjana.in.ua	blog	-	+/-	+/-
5	www.autochip.vn.ua	blog	-	+	+/-
6	kursyvalyut.com	blog	+	+	-
7	dobryjranok.com	blog	+/-		+/-
8	goodmorningua.com	blog	+/-	+/-	-
9	зсшЗльвiв.in.ua	blog	-	-	-
10	presstime.com.ua	forum	+/-	+/-	+/-

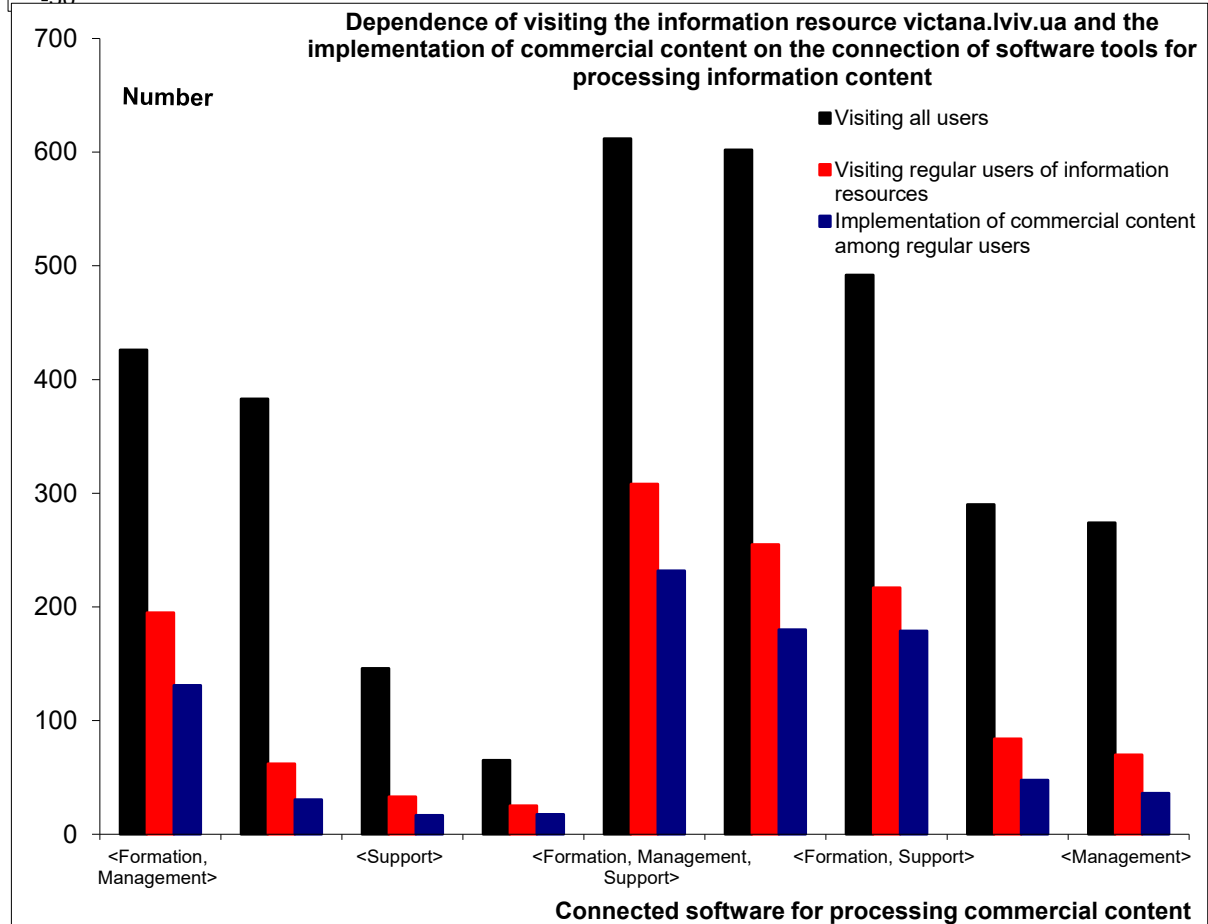
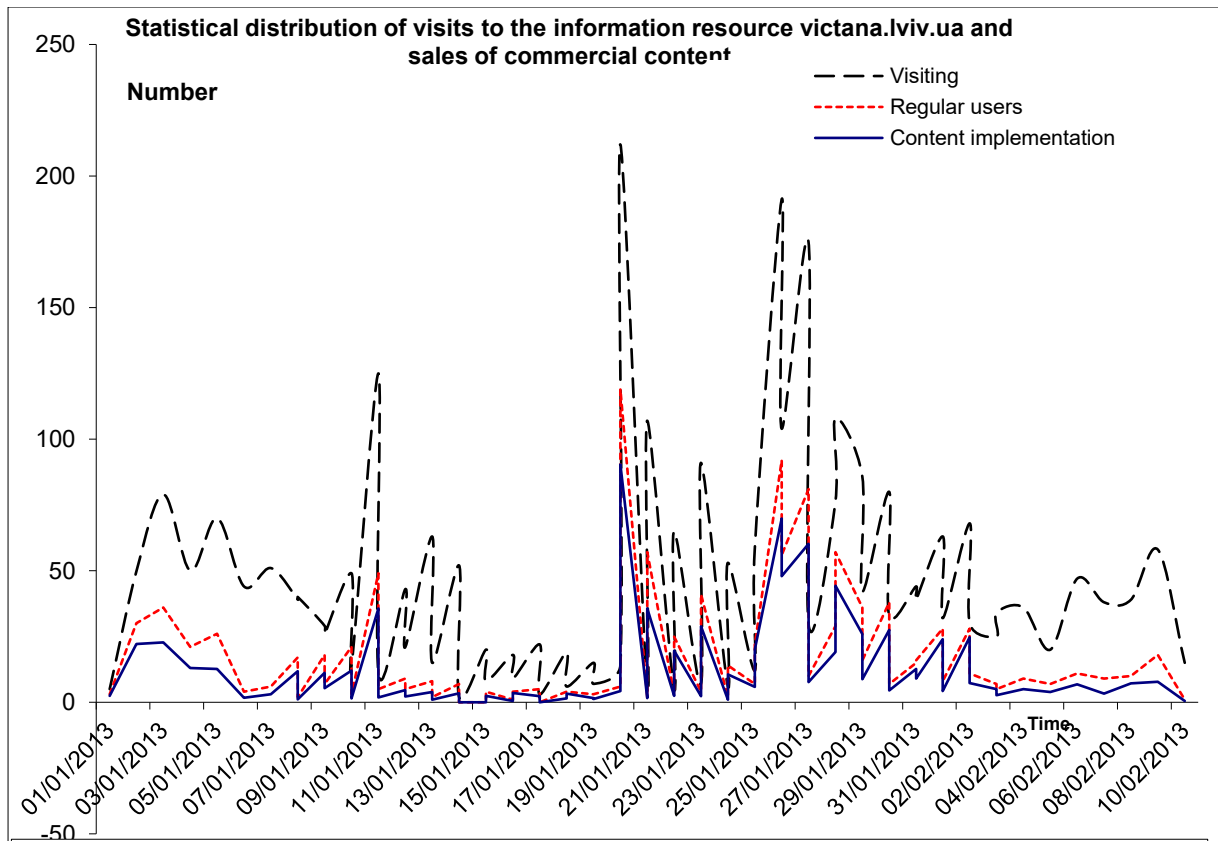
**Table 4**

The results of the systems

Name	1	2	3	4	5	6	7	8	9	10
Visiting	20132	5997052	3654456	1381	8724	4865	9606	25	7	3138
Formation	90%	100%	40%	10%	70%	30%	20%	60%	0%	50%
Management	40%	100%	80%	50%	60%	90%	30%	20%	0%	70%
Support	50%	100%	40%	10%	80%	30%	20%	60%	0%	70%
Content	20%	80%	70%	100%	50%	100%	30%	40%	100%	60%
Uniqueness										
Average duration of visit	1:02	2:14	2:04	3:56	2:27	4:41	1:51	8:12	0:46	4:15
Unique visitors	16586	2501402	1501202	728	4996	3215	7105	7	5	1345
Pages / visits	1,59	1,93	1,67	3,96	2,17	4,54	2,59	3,24	1,67	5,78
% new visits	82,39	41,68	39,88	52,57	57,23	65,45	73,88	28,0	97,32	42,86
% repeated visits	17,61	58,32	60,12	47,43	42,77	34,55	26,12	72,0	2,68	57,14
Page views	31982	11588861	769923	5464	18892	22071	24908	81	12	18132
Failure rate (%)	82,92	71,90	83,08	53,15	68,15	56,14	55,67	48,0	97,02	32,92
Social network (%)	0,06	0,02	0	4,56	0,01	0,06	0,02	0	0,01	7,81
Organic search (%)	34,83	36,10	31,22	22,23	24,34	4,67	8,91	0	0,13	26,04
Not set/non-organic search results (%)	58,22	52,48	42,46	62,49	26,984	91,14	88,13	23	6,72	31,90
Direct traffic (%)	5,34	11,20	26,12	7,53	48,73	2,14	2,35	77	93,12	27,88
Other sites (%)	1,55	0,20	0	3,19	0,04	1,99	0,58	0	0,02	6,37
Conversion rate achieved	0	0	0	0	0	7,83	12,51	0	0	0



**Figure 1:** Statistical analysis of the functioning of the "Photo Gallery Vysotsky"



**Figure 2:** Statistical analysis of the functioning of "Victana" site

The service of keeping visits statistics to the Web resource allows to estimate the increase in sales of textual content in direct proportion to the increase in the number of visits to the information resource, the number of regular users, the prospects of marketing activities (Fig. 3). The presence of subsystems for the formation, management, and maintenance of textual content in Web resource processing systems increases the sales of textual content to the regular user by 9%, actively attracting unique visitors, potential users and expanding the target and regional audience by 11%, viewed pages by 12%, time of visiting information resources by 7%.

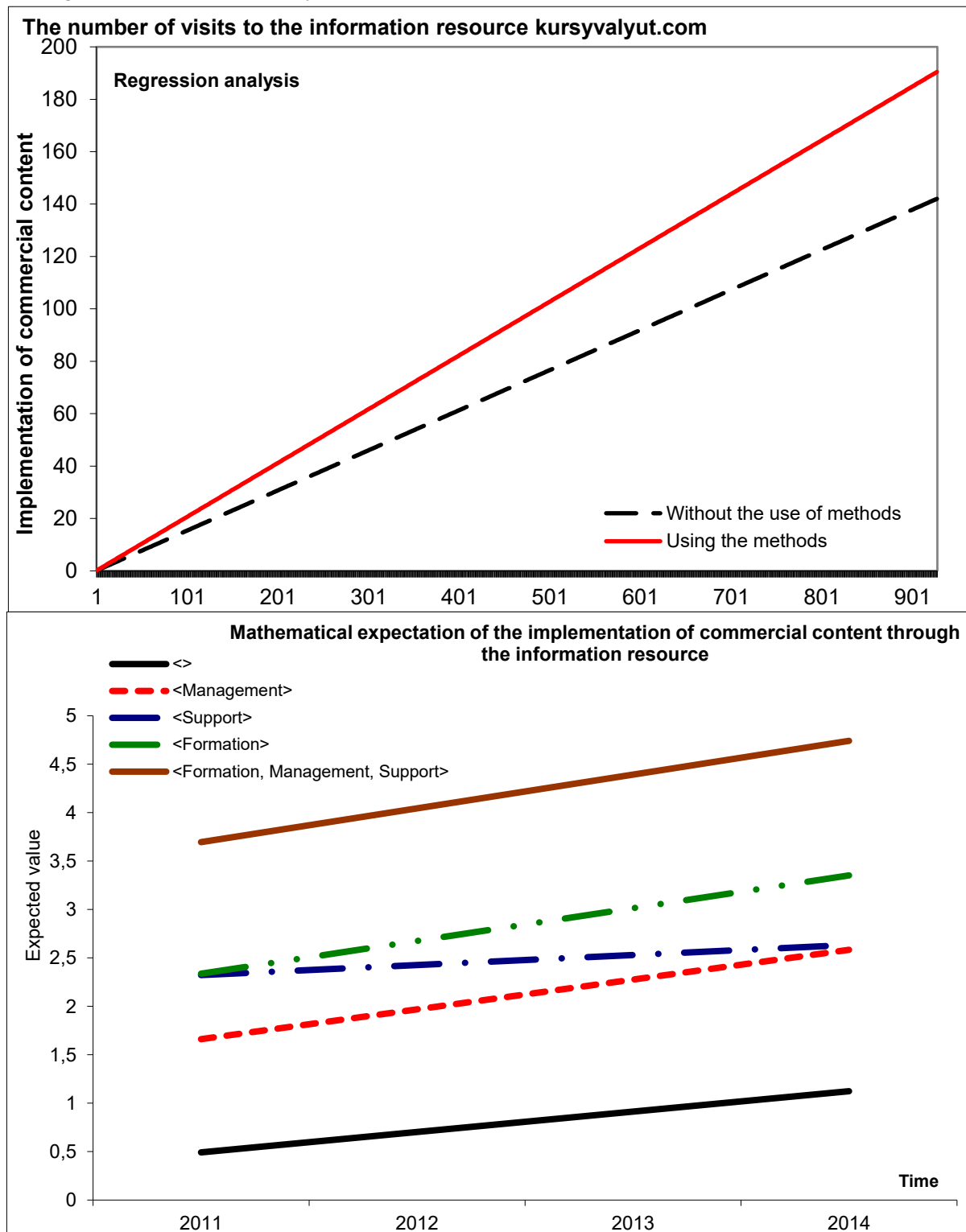


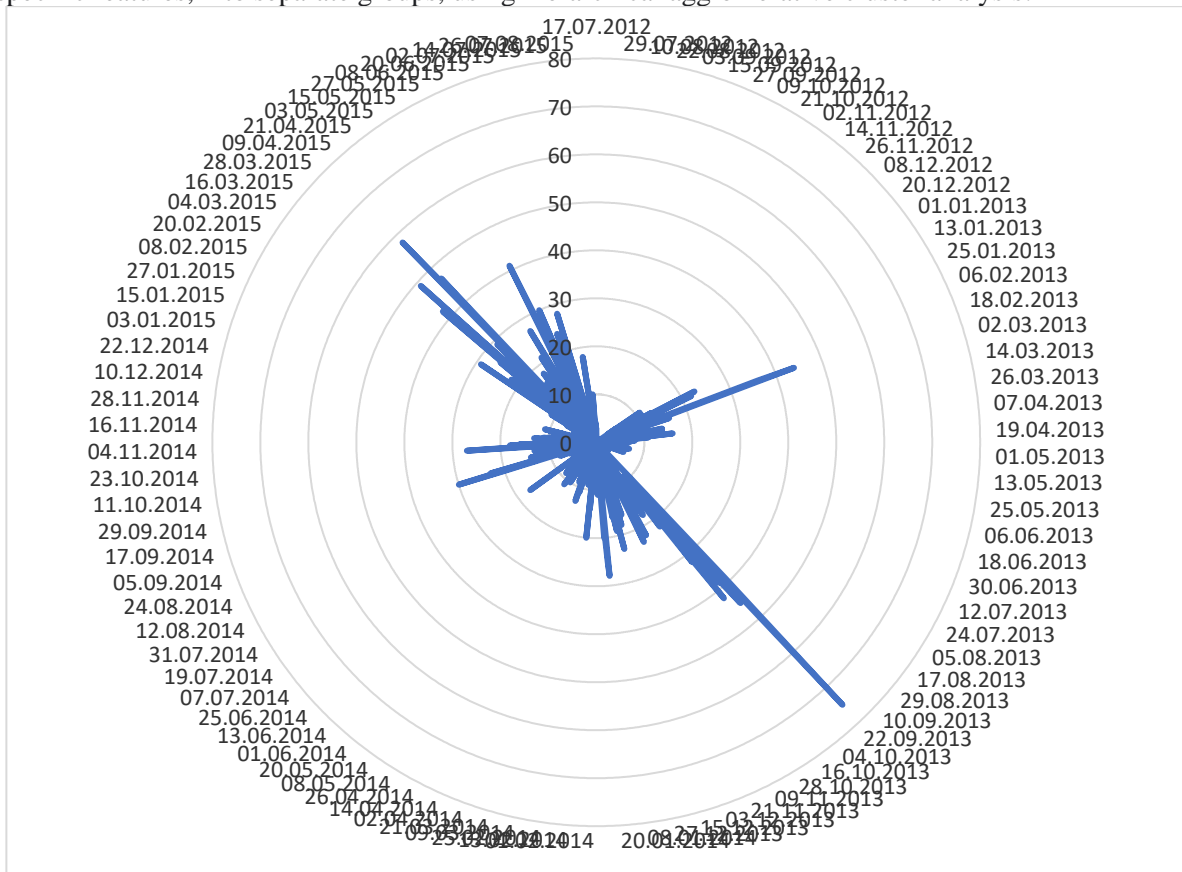
Figure 3: Regression analysis of increased sales of textual content



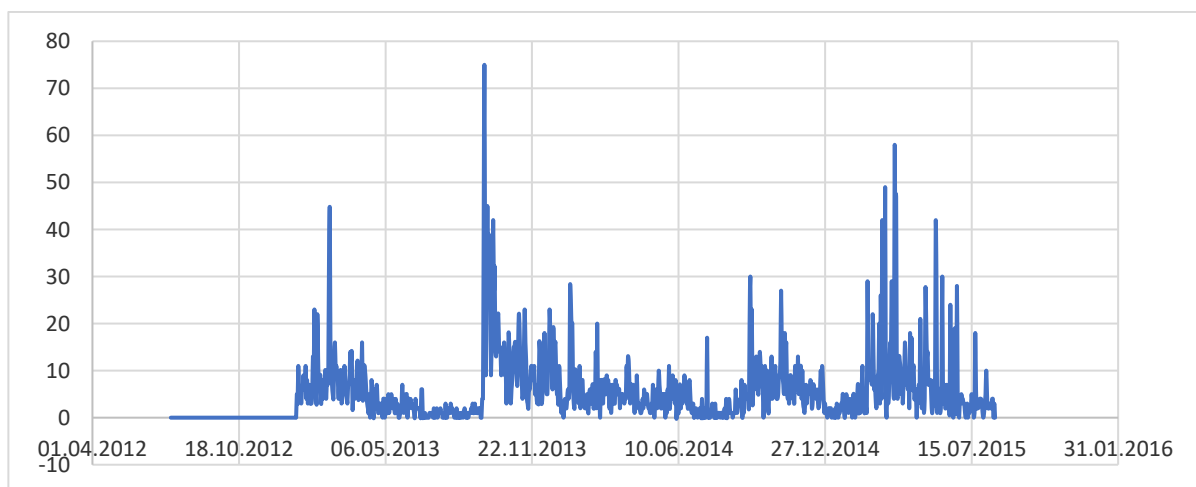
Results from further analysis of Victana blog (Fig 4-5):

- Detection of time series trends by smoothing methods (Fig 6-12).
- Constructing the correlation field and determining the value of the correlation coefficient.
- Calculating the correlation ratio and building graphs of autocorrelation functions.
- Dividing one of the sequences into three equal parts.
- Constructing a correlation matrix for them.
- Finding the coefficients of multiple correlation (Fig. 13).
- Obtaining the result of cluster data analysis (Fig 14-15).

It is necessary to divide a given set of objects, each of which is characterized by the same set of specific features, into separate groups, using hierarchical agglomerative cluster analysis.



**Figure 4:** Graph in the Polar coordinate system of the functioning of "Victana" site



**Figure 5:** Graph in the Cartesian coordinate system of the functioning of "Victana" site

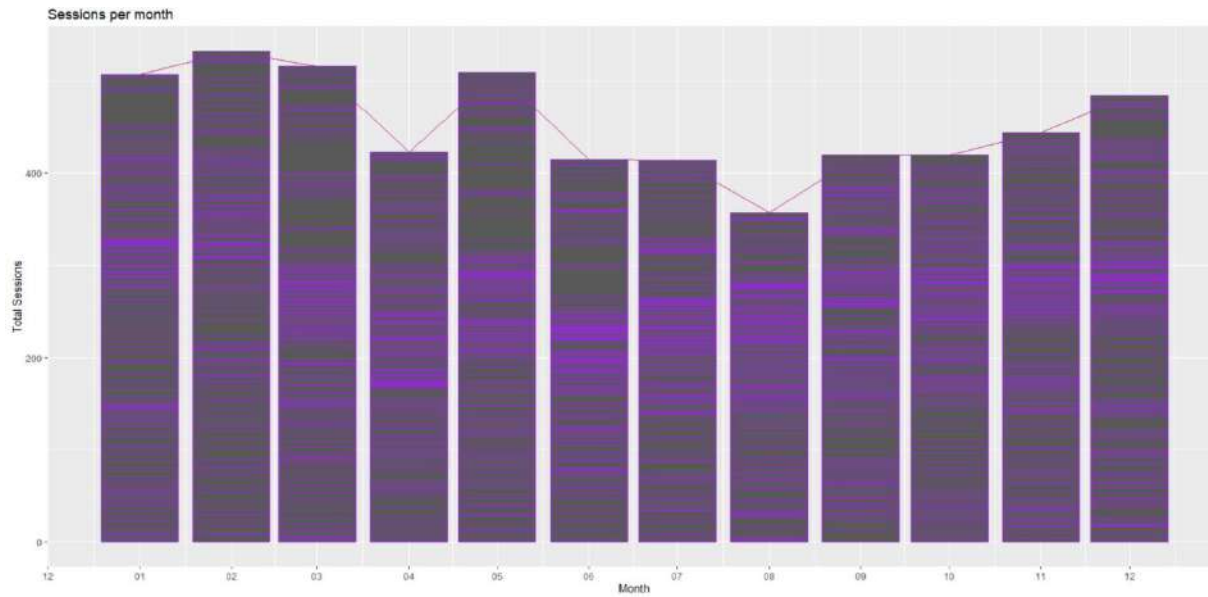


Figure 6: Cumulative results according to histogram data

Arithmetic mean	5,269094
Mode	0
General	5933
Median	3
Maximum	75
Minimum	0
Scope	75
Standard deviation	7,199607
Coefficient of variation	137%
Asymmetry	3,366589
Excess	17,35012
Dispersion	51,83434
Sturges's formula	14,53519
Scott's formula	1,391934
Pocket interval	0,012641

Figure 7: Statistical value for visits of Victana resource

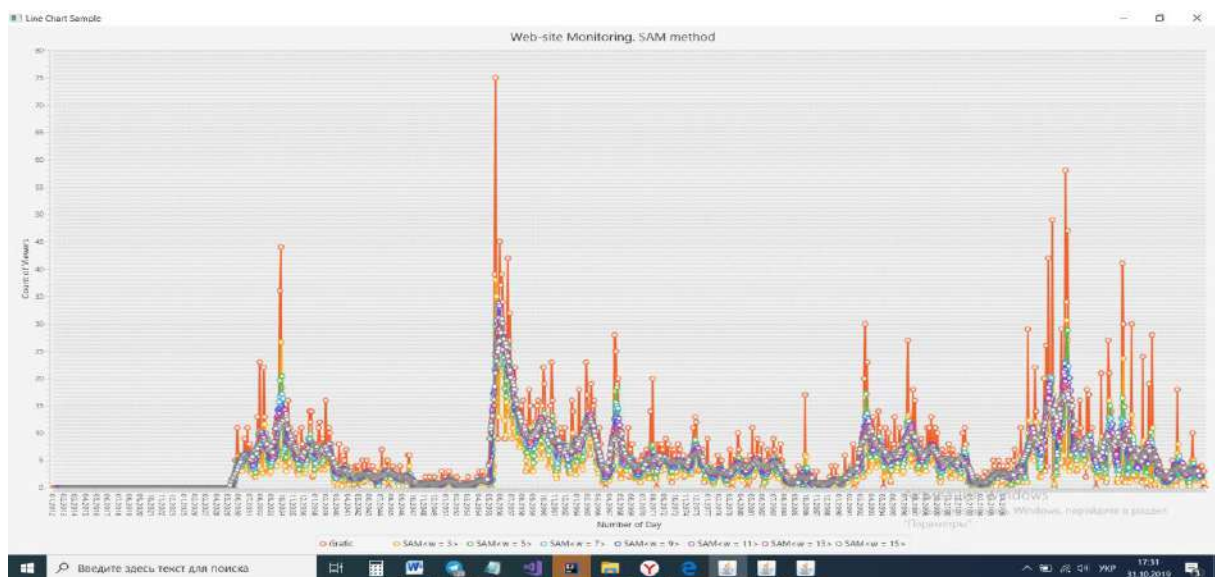


Figure 8: Analysis of Web site monitoring by days based on SAM method

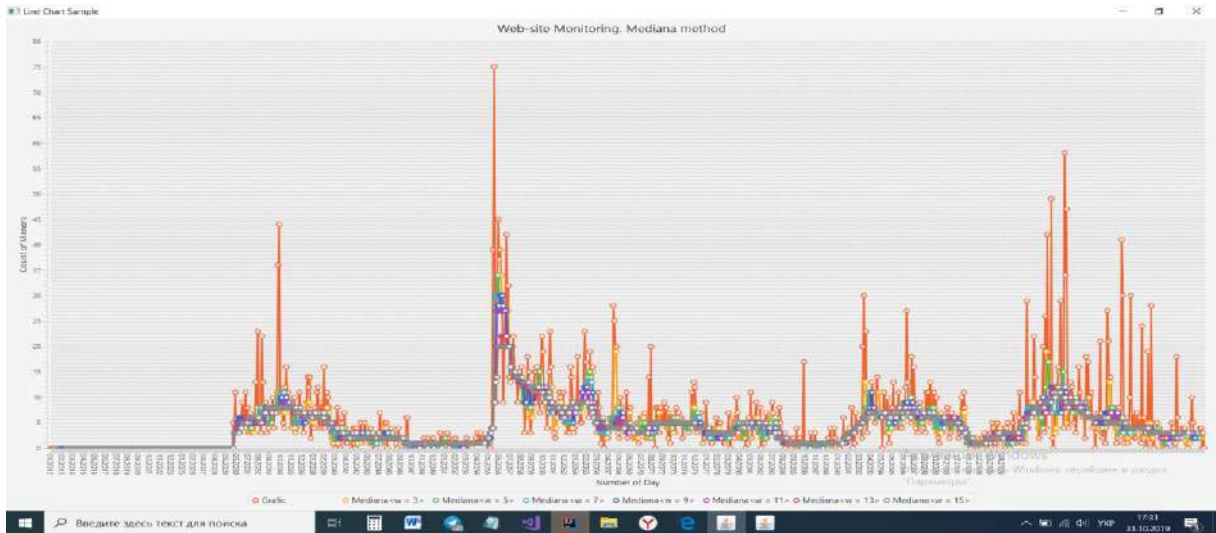


Figure 9: Analysis of Web site monitoring by days based on Medians method

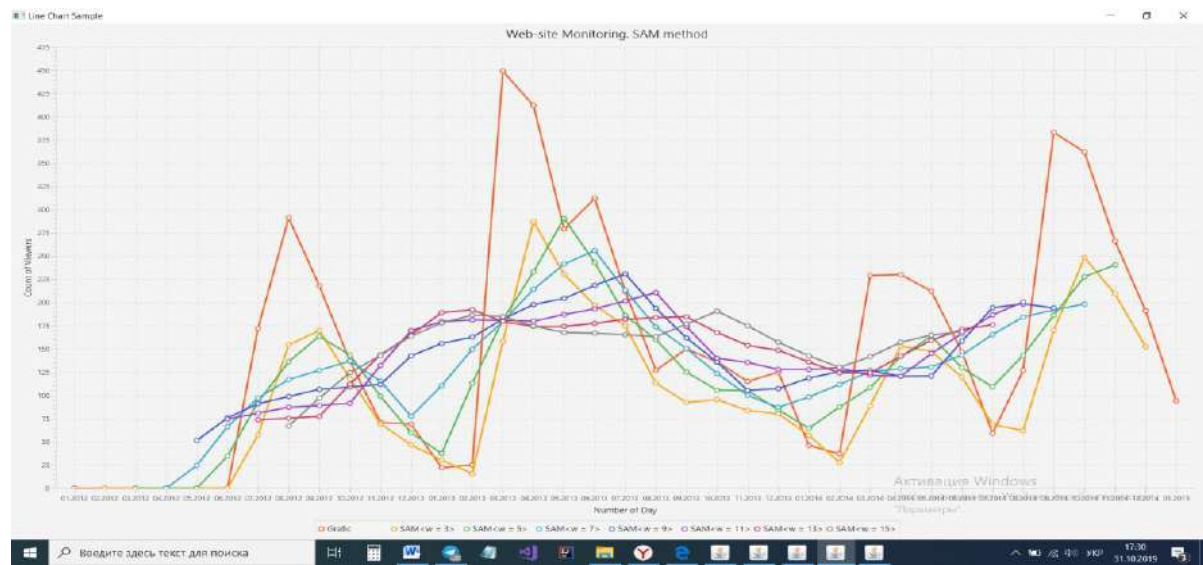


Figure 10: Analysis of Web site monitoring by months based on SAM method



Figure 11: Analysis of Web site monitoring by months based on Medians method

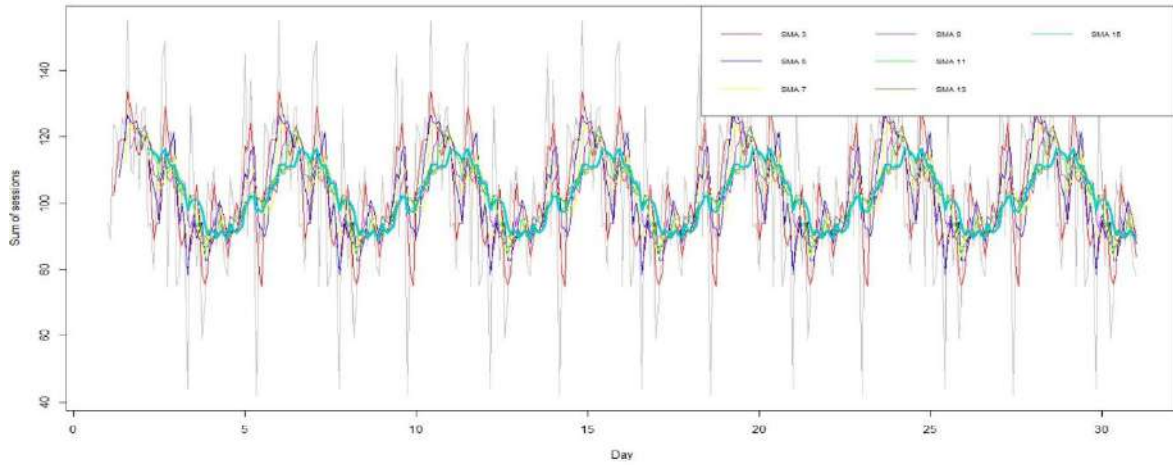


Figure 12: Detection of the trend of the time series by smoothing methods

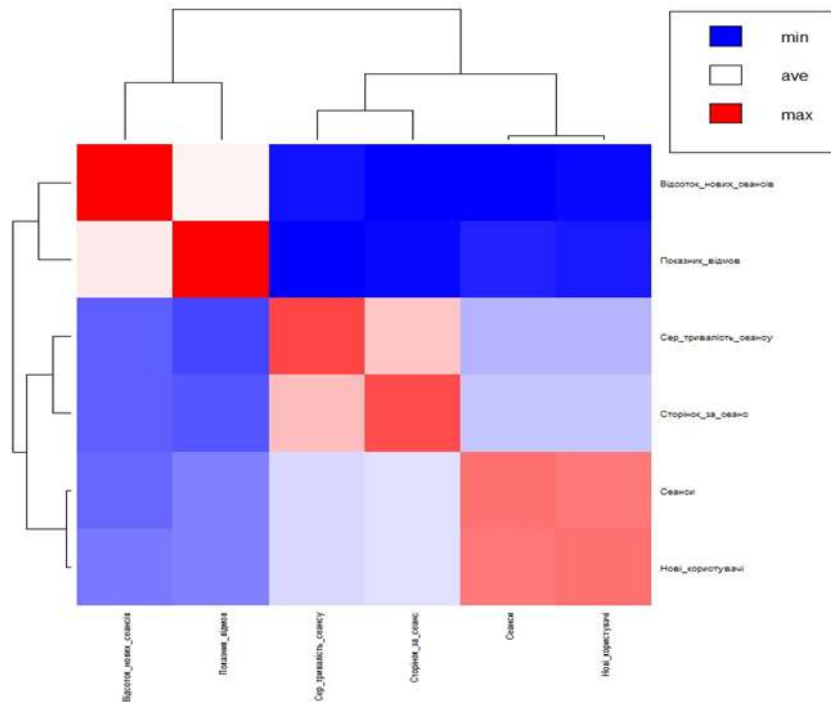


Figure 13: Correlation matrix

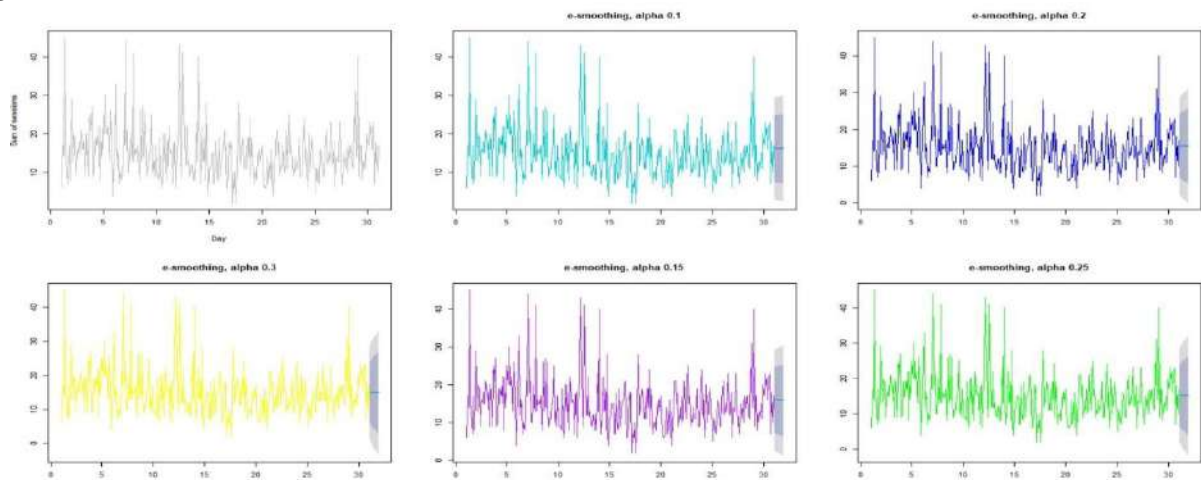
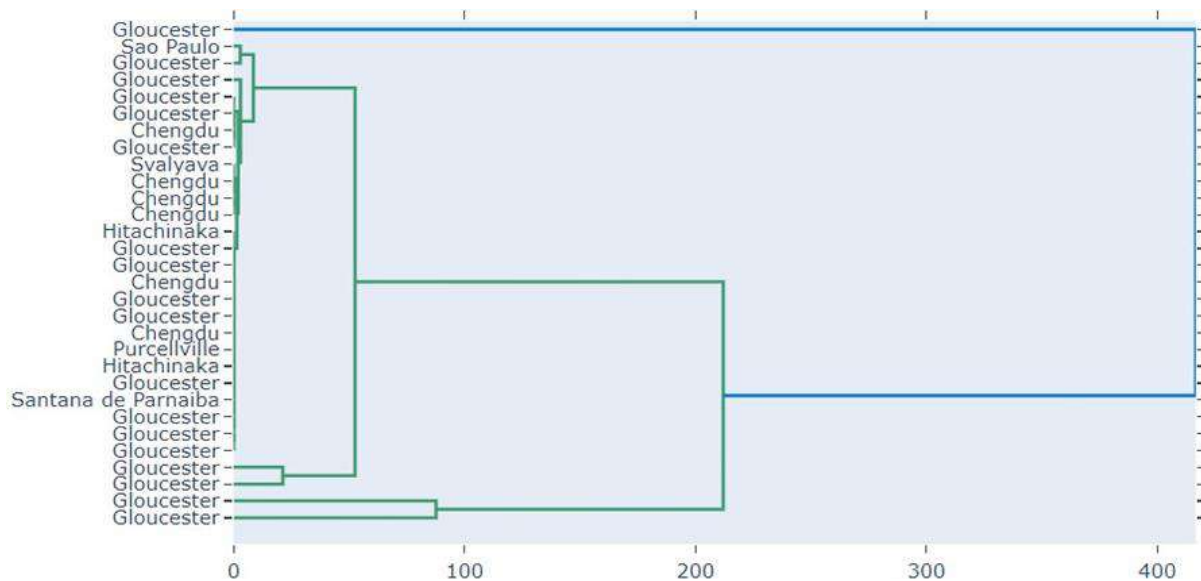


Figure 14: Visiting forecast (highlighted in gray)



**Figure 15:** Dendrogram of attendance as a result of cluster analysis

Accuracy of model is 22.033898305084744%.

## 5. Conclusion

General recommendations for the design of information resource processing systems have been developed, different from the existing by providing more detailed stages and the availability of information resource processing subsystems that allow for effectively implement information resource processing at the system developer level (reducing resources and development time and improving information processing systems resources). Structures of modules of the system of information resources processing for the realisation of stages of the life cycle of text content were proposed. The application software for the formation, management, and maintenance of textual content to achieve the effect at the level of the owner (increasing the profitability, user interest) and user (clarity, simplification of the interface, unification, expansion of choice) of information processing systems are developed and implemented. A method of text content support was developed based on the analysis of statistics about the functioning of the information resources processing system to change the values of management parameters and requirements for the formation of text content. It allowed increasing sales of text content to the regular user by 9%. The structure of the information resources processing system was improved based on the analysis of information resources processing processes, different from the existing ones by the subsystems of text content formation, management, and maintenance, which made it possible to implement the stages of text content life cycle and develop recommendations for designing standard systems. Recommendations for designing the structure of the information resources processing system, different from the existing stages by the detail and the availability of subsystems of information resources processing, which allow maintaining the life cycle of textual content at the system developer level (reducing resources and development time and improving system quality). Software for creating, managing, and maintaining textual content to increase the active involvement of potential users and expand the target audience by 11% to improve the functioning of information resources at the level of the owner (increasing profitability, user interest) and user (intelligibility) are developed and implemented, provided simplification of the interface, automation of information resources processing and expanding the choice of functionality).

## 6. References

- [1] A. Wongkoblak, M. A. Vadillo, V. Curcin, Researching mental health disorders in the era of social media: systematic review, *Journal of medical Internet research* 19(6) (2017) e228. doi: 10.2196/jmir.7215.

- [2] K. Krippendorff, *Content analysis: An introduction to its methodology*. Sage publications, 2018.
- [3] Z. Xiang, Q. Du, Y. Ma, W. Fan, A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism, *Tourism Management* 58 (2017) 51–65. doi: 10.1016/j.tourman.2016.10.001.
- [4] J. C. Fagan, The suitability of web analytics key performance indicators in the academic library environment, *The Journal of Academic Librarianship* 40(1) (2014) 25–34. doi: 10.1016/j.acalib.2013.06.005.
- [5] T. Toivonen, et al., Social media data for conservation science: A methodological overview, *Biological Conservation* 233 (2019) 298–315. doi: 10.1016/j.biocon.2019.01.023.
- [6] S. J. Miah, H. Q. Vu, J. Gammack, M. McGrath, A big data analytics method for tourist behaviour analysis, *Information & Management* 54(6) (2017) 771–785. doi: 10.1016/j.im.2016.11.011.
- [7] J. Li, L. Xu, L. Tang, S. Wang, L. Li, Big data in tourism research: A literature review, *Tourism Management* 68 (2018) 301–323. doi: 10.1016/j.tourman.2018.03.009.
- [8] A. R. Alaei, S. Becken, B. Stantic, Sentiment analysis in tourism: capitalizing on big data, *Journal of Travel Research* 58(2) (2019) 175–191. doi: 10.1177/0047287517747753.
- [9] V. Lytvyn, V. Danylyk, M. Bublyk, L. Chyrun, V. Panasyuk, O. Korolenko, The lexical innovations identification in English-language eurointegration discourse for the goods analysis by comments in e-commerce resources, in: *Proceedings of 16th Int. Conf. on Computer science and information technologies*, Lviv, 2021, pp. 85–97. doi: 10.1109/CSIT52700.2021.9648594.
- [10] W. Duan, Q. Cao, Y. Yu, S. Levy, Mining online user-generated content: using sentiment analysis technique to study hotel service quality, in: *IEEE 46th Hawaii International Conference on System Sciences*, January 2013, pp. 3119–3128. doi: 10.1109/HICSS.2013.400.
- [11] B. Clifton, *Advanced web metrics with Google Analytics*. John Wiley & Sons, 2012.
- [12] J. R. Saura, P. Palos-Sánchez, L. M. Cerdá Suárez, Understanding the digital marketing environment with KPIs and web analytics, *Future Internet* 9(4) (2017) 76. doi: 10.3390/fi9040076.
- [13] V. Podobnik, An analysis of facebook social media marketing key performance indicators: The case of premier league brands, in: *Proceedings of the IEEE 12th International Conference on Telecommunications*, June 2013, pp. 131–138.
- [14] S. A. Rios, F. Aguilera, J. D. Nuñez-Gonzalez, M. Graña, Semantically enhanced network analysis for influencer identification in online social networks, *Neurocomputing* 326 (2019) 71–81. doi: 10.1016/j.neucom.2017.01.123.
- [15] M. Saqr, U. Fors, M. Tedre, J. Nouri, How social network analysis can be used to monitor online collaborative learning and guide an informed intervention, *PloS* 13(3) (2018) e0194777. doi: 10.1371/journal.pone.0194777.
- [16] V. D. Păvăloaia, E. M. Teodor, D. Fotache, M. Danileț, Opinion mining on social media data: sentiment analysis of user preferences, *Sustainability* 11(16) (2019) 4459. doi: 10.3390/su11164459.
- [17] A. Arora, A. Srivastava, S. Bansal, Business competitive analysis using promoted post detection on social media, *Journal of Retailing and Consumer Services* 54 (2020) 101941. doi: 10.1016/j.jretconser.2019.101941.
- [18] M. Bublyk, A. Kowalska-Styczeń, V. Lytvyn, V. Vysotska, The Ukrainian economy transformation into the circular based on fuzzy-logic cluster analysis, *Energies* 14(18) (2021) art. 5951. doi: 10.3390/en14185951.
- [19] R. Levus, A. Berko, L. Chyrun, V. Panasyuk, M. Hrubel, Intelligent System for Arbitrage Situations Searching in the Cryptocurrency Market, *CEUR Workshop Proceedings Vol-2917* (2021) 407–440.
- [20] O. Prokipchuk, L. Chyrun, M. Bublyk, V. Panasyuk, V. Yakimtsov, R. Kovalchuk, Intelligent System for Checking the Authenticity of Goods Based on Blockchain Technology, *CEUR Workshop Proceedings Vol-2917* (2021) 618–665.
- [21] V. Husak, L. Chyrun, Y. Matseliukh, A. Gozhyj, R. Nanivskyi, M. Luchko, Intelligent Real-Time Vehicle Tracking Information System, *CEUR Workshop Proceedings, Vol-2917* (2021) 666–698.
- [22] A. Berko, M. Bublyk, L. Chyrun, Y. Matseliukh, R. Levus, V. Panasyuk, O. Brodyak, L. Dzyubyk, O. Garbich-Moshora, Models and Methods for E-Commerce Systems Designing in the Global Economy Development Conditions Based on Mealy and Moore Machines, *CEUR Workshop Proceedings Vol-2870* (2021) 1574–1593.

- [23] L. Chyrun, The E-Commerce Systems Modelling Based on Petri Networks, CEUR Workshop Proceedings Vol-2870 (2021) 1604–1631.
- [24] A. Berko, V. Andrunyk, L. Chyrun, M. Sorokovskyy, O. Oborska, O. Oryshchyn, M. Luchkevych, O. Brodovska, The Content Analysis Method for the Information Resources Formation in Electronic Content Commerce Systems, CEUR Workshop Proceedings 2870 (2021) 1632–1651.
- [25] V. Kuchkovskiy, V. Andrunyk, M. Krylyshyn, L. Chyrun, A. Vysotskyi, S. Chyrun, N. Sokulska, I. Brodovska, Application of Online Marketing Methods and SEO Technologies for Web Resources Analysis within the Region, CEUR Workshop Proceedings 2870 (2021) 1652–1693.
- [26] I. Dyyak, L. Chyrun, N. Antonyuk, A. Berko, V. Andrunyk, A. Khudyi, Development of Internet Auction Systems, in: Proceedings of the IEEE 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT, 2020, 1, 55–61. doi: 10.1109/CSIT49958.2020.9321844.
- [27] L. Chyrun, N. Antonyuk, V. Andrunyk, I. Dyyak, A. Berko, K. Mykich, Features of Internet Auction Systems Design, in: Proceedings of the IEEE 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT, 2020, 1, pp. 134–140..
- [28] N. Antonyuk, M. Medykovskyy, L. Chyrun, M. Dverii, O. Oborska, M. Krylyshyn, A. Vysotsky, N. Tsiura, O. Naum, Online Tourism System Development for Searching and Planning Trips with User's Requirements, Advances in Intelligent Systems and Computing 1080 (2020) 831–863.
- [29] Y. Kis, L. Chyrun, T. Tsymbaliak, L. Chyrun, Development of System for Managers Relationship Management with Customers. In: Advances in Intelligent Systems and Computing 1020 (2020) 405–421. doi: 10.1007/978-3-030-26474-1\_29.
- [30] Y. Tverdokhlib, V. Andrunyk, L. Chyrun, L. Chyrun, N. Antonyuk, I. Dyyak, O. Naum, D. Uhryn, V. Basto-Fernandes, Analysis and estimation of popular places in online tourism based on machine learning technology, CEUR Workshop Proceedings Vol-2631 (2020) 457–470.
- [31] A. Berko, L. Chyrun, I. Dyyak, V. Andrunyk, L. Chyrun, N. Antonyuk, E-Commercial Systems Designing Methods for Virtual Enterprise, in: Proceedings of the IEEE 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT, 2020, 1, pp. 119–124. doi: 10.1109/CSIT49958.2020.9321939.
- [32] B. Rusyn, L. Pohreliuk, O. Kapshii, J. Varetskyy, A. Demchuk, I. Karpov, A. Gozhyj, V. Gozhyj, I. Kalinina, An Intelligent System for Commercial of Information Products Distribution Based SEO and Sitecore CMS, CEUR workshop proceedings Vol-2604 (2020) 760–777.
- [33] A. Demchuk, B. Rusyn, L. Pohreliuk, A. Gozhyj, I. Kalinina, L. Chyrun, N. Antonyuk, Commercial content distribution system based on neural network and machine learning, CEUR Workshop Proceedings Vol-2516 (2019) 40–57.
- [34] R. Yurynets, Z. Yurynets, I. Myshchysyn, N. Zhyhaylo, A. Pekhnyk, Optimal Strategy for the Development of Insurance Business Structures in a Competitive Environment, CEUR Workshop Proceedings Vol-2631 (2020) 79–94.
- [35] R. Yurynets, Z. Yurynets, M. Kokhan, Econometric Analysis of the Impact of Expert Assessments on the Business Activity in the Context of Investment and Innovation Development, CEUR workshop proceedings Vol-2604 (2020) 680–694.
- [36] O. Cherednichenko, O. Yanholenko, O. Iakovleva, O. Kustov, Models of research activity measurement: Web-based monitoring implementation, Lecture Notes in Business Information Processing 193 (2014) 75–87. doi: 10.1007/978-3-319-11373-9\_7.
- [37] A.Y. Berko, Methods and models of data integration in E-business systems, Actual Problems of Economics 10, (2008) 17–24.
- [38] A. Berko, Consolidated data models for electronic business systems. In: The Experience of Designing and Application of CAD Systems in Microelectronics, CADSM 2007, pp. 341–342.
- [39] I. Oksanych, I. Shevchenko, I. Shcherbak, S. Shcherbak, Development of specialized services for predicting the business activity indicators based on micro-service architecture, Eastern-European Journal of Enterprise Technologies, 2(2-86) (2017) 50–55. doi: 10.15587/1729-4061.2017.98991.