Uniform Method of Operative Content Management in Web Systems

Aleksandr Gozhyj[1], Lyubomyr Chyrun[2],
Agnieszka Kowalska-Styczen[3], Olga Lozynska[4]

[1]Petro Mohyla Black Sea National University, Mykolayiv, Ukraine
[2,4]Lviv Polytechnic National University, Lviv, Ukraine
[3]Silesian University of Technologyty, Gliwice, Poland

alex.gozhyj@gmail.com

Abstract. In the given article content lifecycle support method in Web systems is proposed. The uniform method realizes information resources processing in the Web systems and simplifies the information technology of Web content automation management. In the given paper the main problems Web content function and management services in Web systems are analyzed. In this article, the authors proposed to apply the technique of content analysis of textual information in e-commerce systems for the automation of e-business and decision making by the appropriate person.

Keywords: content, content analysis, e-commerce, quantitative content analysis, qualitative content analysis, content management system, content lifecycle, information resources, Web system.

1 Introduction

The current level of complexity of tasks solved through e-commerce systems is constantly growing: from the automated information collection and interpretation to the management, design, modeling and forecasting of the various business processes, for example, forecasting by the e-commerce system the changes in demand for a certain type of products depending on the thematic discussions by forum users or the analysis of arrays of comments on these products. However, the most complex tasks are not solved automatically by e-commerce systems, but by persons who accompany the process of the corresponding system [1-3]. There was a need to search for the non-traditional approaches to the use of information technologies and mathematical methods in the processes of preparation and decision-making in e-commerce systems based on the information obtained from forums, comments of product users, electronic correspondence, search engine and agents results [4-6]. To date, most of the information obtained by the users from the e-commerce system is being studied by the moderators. For example, large text arrays of comments on the products presented on the site, or forum information, are mainly filtered by the moderators over a certain period of time by a certain circle of proxies [7-10]. To increase the demand for prod-

ucts, it is necessary to post only filtered user comments on the site, where not only obscene language is available (such comments are blocked), but where there is no advertising of competitors' products. In addition, a detailed analysis of the multitude of comments on products enables the manufacturer to analyze the market situation and adjust the demand for its products [11-16]. The study of information needs for the similar system operation shows that in the decision-making process, along with the use of factual materials, it is important to involve the textual information about the reflection of the relationships between various facts, events, persons, etc. and its automatic processing in a short period of time with minimal financial costs [17-21].

The modern market is characterized by an increase in demand for Web content and growth proposals in public enterprises and public institutions. [1] Using Web content helps optimize the development programs of management, trade and long-term and economic strategy for production. It is associated with an increase in the complexity of management and implementation of systems using the predictive Web content of character. Specialized information resources are most type's users of Web content, e.g., online publishing, online magazines, online newspapers, etc. For these users content is subject for making other Web content [2].

The article purpose is to design a typical architecture for content management systems. The work feasibility lies in the methods development of Web resources processing in such systems. It promotes the active development in Ukraine e-business and effective implementation of the Web systems [3]. The scientific problem is the lack of a standardized approach to architectural design and development of such systems and the subsystems of Web resources processing. There is no classification of the Web systems, which creates problems for the study of information resources processing methods in these systems. This creates problems for the appropriate software implementation. This also justifies the purpose, relevance, appropriateness and research directions. The special feature of Web systems is as follows [4]: openness, global character, unlimited in time, frankness, direct interaction with the user, automatic analysis of queries and users data tracking; reducing the cost of e-business functioning; additional information providing in interactively [5-15].

2   Recent research and publications analysis

The Web systems implementation contributes to the e-business success [1], including: business globalization increases the demand for Web content and fast access to it; uneven functioning business processes according to the regions. leads to increasing needs on speed, regularity and frequency of obtaining of the necessary Web content; saving time in the necessary Web content getting; personalization in services provide in the Web systems; the Web systems integration. Advantages and feasibility of the Web systems introduction are as follows [1-2]: increased efficiency of Web content receiving; the cycle reducing of Web content production and sale; costs reducing associated with the content exchange; the openness of the Web systems due to customers; automatically inform users of Web content; alternative sales channels creation, such as online newspaper. Problems of Web resources processing in the such systems associated with increasing content amounts on the Internet; the rapid prevalence of Internet access; active e-business development; the set expanding of Web services and

goods; the growth in demand for Web services and products; Internet technology and facilities creating and applications expansion of Web resources processing techniques [1-4]. In this area are actively working the world's leading manufacturers of Web resources processing, particularly developed such systems [1-3, 11-21]:

- Web content on-line sale: online edition as dictionaries/directories (The Mammals of Texas Online Edition, Dictionary Online, Oxford English etc.), distance learning (Ashford University, Capella, Argosy, Walden, etc.), Online Journals (Chip, Air-liner World, Wall Street Journal, etc.), Online newspapers (New York Post, Washington Post, New York Times, etc.), content entertainment portals (tochka.net, PokerNews, etc.), informative portals (Karpaty.ua, Weather Channel, etc.), online publishing (Online Publishers Association, iUniverse, etc.), and children content portals (vshkolu.com, teremoc.ru, etc.);
- off-line content sales: RSS Subscription Extension (Microsoft, Intel, Google, Apple, etc.); Marketing Services Shop (Local Internet Marketing,The Copy Box, etc.); copywriting services (Sopywriting 911, Textbroker, Apple Copywriting, etc.);\;
- Online shop for sale: video (Dailymotion, Yahoo! Directory, OnSiteVideos.com, Opera World, etc.); eBooks (Google eBookstore, PayLoadz, Amazon.com Inc. Etc.); Software (Apple, AppStore, Android, etc.); music (Amazon.com, MPI Home Video, MTI Home Video, Mass Music, etc.); picture (buy-images-online.com, PhotoBox, iStockPhoto, BigStockPhoto etc.); movies (Amazon.com, MovieSunlimited.com, DVDwarehouse.com.au etc.); digital art (Imagekind Inc.), files, forms, certificates, articles, manuals, etc.;
- content store of different types: cloud computing (Android, iPhone, Linux, Mac, Apple, Google, Palm); .cloud storage (Amazon, EMC, Google, Apple, Microsoft).

Web content is a business processes object for the Web systems (Table 1). Example is an information product or content for information resource as online newspapers, online publishing, marketing research and consulting services portal, etc.

Table 1. The features of Web systems

Name	Characterization
Virtuality	Personal contact lack of between the subjects in the buying/selling.
Interactivity	Adequate information support a user query in an interactive mode dumb dialogue.
Globality	Lack of time, space, assortment, administrative, social and demographic boundaries.
Dynamism	The on-line trade ability in momentary changes and adaptation to new conditions.
Efficiency	Demand, profits, economic benefits and social impact providing.

The content life cycle is a process that content takes place during management through the submission stages. A content lifecycle model does not take into account all stages of information resources processing.

3 Problems selection

Content streams number is greater than the ways number of goods moving to industrial enterprises. Much of the content flows consists easily formalized and automated procedures. The main problem is the common approach lack to process modeling, design and development of Web systems. There is a general and detailed classification lack of Web systems. This leads to the problem of the general methods definition for the architecture and algorithms development of these systems functioning. The existing Web systems do not support the entire content life cycle, do not solve information resources processing and content management problems (Table 2). This justifies the research purpose, relevance, appropriateness and directions.

Table 2. Characteristics comparison of Web systems

The system features	E-commerce	Content commerce
Automatic detection products aging within the meaning	–	+
Automatic detection and elimination of products duplication	–	+
Products storage lack	–	+
Product keeping in the database	–	+
The variety products growth	+/–	+
Constant products quantity	–	+
Product immateriality	–	+
Efficiency Product search by keyword	+/–	+
The promotion effectiveness by keyword	+/–	+
Automatic determination of the relevance product	+/–	+
Automatic analysis of audience	+/–	+
Digests automatic creation	–	+
Automatic creation product	–	+
Automatic formatting product	–	+
User experience Impact to drive sales	+/–	+

4 Goals formulation

Input information the functioning of Web systems is evidence of a system work appointment and conditions. They define the main purpose of modeling. They also make it possible to formulate the requirements for the system formal model S and content management models. Model of Web systems presented as $S = \langle X, C, V, H, Function, T, Y \rangle$, where $X = \{x_1, x_2, ..., x_{n_X}\}$ is the input data to the system, $C = \{c_1, c_2, ..., c_{n_C}\}$ is influences the content flow on the system, $V = \{v_1, v_2, ..., v_{n_V}\}$ is the environment influence, $H = \{h_1, h_2, ..., h_{n_H}\}$ is internal system parameters, $Y = \{y_1, y_2, ..., y_{n_Y}\}$ is the system output characteristics, $T = \{t_1, t_2, ..., t_{n_T}\}$ is the content management transaction time. The process S of Web systems functioning described by the function as $y_j(t_i + \Delta t) = Function(x_i, c_r, v_l, h_k, t_i)$, where x_i is the visitor/user

query to the system. Characteristics component y_j according to Google Analytics are the visits number for the time period Δt, the average Time Spent Online (min:s), bounce rate (%), achieved the goal, dynamics (%), total viewing pages, page views number per visit, new visits (%), total unique visitors, traffic sources in % (search engines, direct traffic or other sites).

5   Method of text information analysis

During the investigation of the mechanisms of textual information generation, it was revealed that the choice of the text information description model depends on how the probabilistic and linguistic test is constructed and how the extraction of the certain units is made from the text information [1-4]. The probabilistic modeling of textual information and its components is an introductory, preparatory stage to the description of the functions of linguistic units in the text. The study of the language and speech functioning through the probabilistic modeling of textual information relies upon the models of probability theory, mathematical statistics, and combinatorial analysis. In linguistic research, and especially in the implementation of the information retrieval algorithms [1-3], there are always problems associated with predicting the occurrence of a certain number of word forms or word combinations belonging to the certain classes in the segment of a given length. Probabilistic modeling of texts and syllables, word combinations, grammatical classes also makes it possible to determine the sample size necessary to provide, with a given probability, the occurrence of at least one time a corresponding linguistic unit [2].

Content analysis is one of the known analysis methods for the textual information. It is the standard method of research in the social sciences, the subject of which is the analysis of the textual arrays content and communicative correspondence products (for example, comments, forums, electronic correspondence, articles, etc.). Content analysis is the quantitative and qualitative analysis of textual information and textual arrays for the purpose of the subsequent meaningful interpretation of the obtained quantitative and qualitative regularities. *Content analysis* is used to study the sources that are invariant in structure or content, but which exist as a non-systematized, randomly organized textual material [1-5]. The method of content analysis is to form an abstract text content model from a variety of textual material. There are two types of content analysis: quantitative and qualitative.

Quantitative content analysis (content-related) is a research method of words, topics, and messages that focuses on a message content. So, before analyzing the selected linguistic units, it is necessary to predict their content and determine each possible observation result in accordance with the researcher expectations. As first step in conducting a content analysis of this type, the researcher has to create a dictionary in which each observation will be defined and assigned to appropriate category.

Qualitative content analysis (structural) is a research method in which the researcher is not so much interested in the message content as in the form and structure of this message. For example, you need to determine how much time or print space is allocated to a topic in a particular source, or how many words or columns have been given to each topic of the corresponding category. Then the relative density of each topic and category are calculated, and a comparative analysis of the corresponding

topics is carried out for the further prediction of events, processes, and the like. The relative density P of any category is calculated by the formula: $P=R/T$, where R is the number of the analysis units that sets this category, T is the total number of the analysis units. In [1-3] authors singled out the following components of the content-analytical study.

1. Content analysis deals with a mass of texts, using the typical sociological procedures for continuous reading or sampling observation while meeting the representativeness requirements.
2. Content analysis involves structuring, segmentation, decomposition of texts or defining the meaningful invariants that are repeated in all or in a number of texts related to the mass collection under investigation.
3. To ensure such a uniformity of segmentation and the isolation of invariants, content analysis involves a high degree of formalization, the use of harsh operational rules and formal algorithms to implement the content analysis procedures.
4. Content analysis is an analytical and synthetic procedure since it is supposed to formalize the division of the whole texts or the selection of their certain items for the subsequent collection.
5. Content analysis involves the use of methods of probability theory and mathematical statistics.

According to the Ukrainian researcher in [2], the content analysis is a qualitative and quantitative method of studying the documents, characterized by the conclusions objectivity and the procedure rigidity. It is the quantitative text processing with the subsequent results interpretation. The subject of the content analysis can be both problems of the social reality, expressed or concealed in documents, and the internal regularities of the study object itself. The key milestones of the formation of e-commerce information analysis system based on the textual information content analysis are the following:

1. Definition of the research topic, analysis purpose and object, its chronological and geographical frameworks, selection principles;
2. The classifier formation for the selection of key citations and the preparation of instructions for the encoder;
3. The coding of publication fragments;
4. Storing of publication fragments in the database;
5. Processing of publication fragments in an automated mode.

The implementation of these stages leads to the formation of thematically selected arrays of information, in which information on the coverage of all aspects of the problem under investigation is accumulated, taking into account the diversity of opinions and views.

The construction of the information analysis systems largely depends on the availability and compliance with the rules for the above steps:

6. The procedure of content-analytical selection should consist of precisely defined actions, which, without any changes, should be subject to all processing objects;
7. The sections of the classifier on which the publication fragments are coded should be clear and elementary to avoid ambiguity;

8. The interpretation of the research results should cover all the publication fragments obtained; conclusions should not be based on any part of the results, but rather take into account all of them without exception.

For the automated processing of textual information is of great importance not only what the occurrence frequency of a particular category of the linguistic unit, but generally its presence in the test text or absence. Quantitative estimation makes it possible to draw the objective conclusions about the material focus concerning the number of uses of the analysis units (key quotes) in the studied texts, such as the number of positive/negative responses to a certain type of product. Qualitative analysis allows us to draw the objective conclusions, where there is a certain linguistic unit in the text and in what context.

6   The main stages of content analysis in e-commerce

The main stages of content analysis of textual information in e-commerce systems are the following items.

1. *Identification of an array of sources under investigation or messages* through a set of specified criteria to which each message corresponds:

- a given source type (forum, email, online newspaper, chat, online magazine);
- one message type (article, e-mail, banner, comment);
- specified parties involved in the communication process (addresser, addressee, recipient);
- confronted size of messages (minimum scope or length);
- message occurrence frequency;
- message distribution method;
- message distribution place;
- time of message emergence and so on.

2. *The formation of the sample multitude of messages.* In some cases, the entire set of sources identified at the first stage is studied, since the message to be analyzed is often quantitatively limited and available. However, sometimes the content analysis should be based on a limited sample, extracted from a larger array of information.

3. *Identification of the linguistic units under analysis,* for example, words or topics. The correct choice of the linguistic units to be analyzed is an important part of the entire work. A word is a simple message element. A topic is the other unit, it is a separate statement about any object. There are clear requirements for the choice of the possible linguistic unit to be analyzed:

- large enough to interpret the meaning;
- small enough to not interpret many meanings;
- easily identified;
- the number of units is large enough for sampling.

If accepted a topic for the unit analysis, the following rules are taken into account:

- the size of the topic does not go beyond a paragraph;
- a new topic arises if there is a substitution of goal, category or someone who perceives or acts.

4. *Allocation of calculation units,* which can coincide with the meaningful units or be specific in nature. In the first case, the analysis procedure is reduced to the frequency occurrence count of the allocated meaningful unit; in the other, the researcher, on the basis of the analyzed material and the research objectives, sets forth the calculation units that can be:

- the physical length of the texts;
- the text area filled with the meaningful units;
- number of lines (paragraphs, characters, text columns);
- file size and type;
- a number of pictures with a certain content, plot, etc.

In some cases, researchers use other calculation elements. At this stage of the content analysis, the fundamental value has a rigorous definition of its operators.

5. *The calculation procedure itself.* In general, it is similar to the standard methods of classification by the selected groups from the formulas of mathematical statistics and the probability theory. There are also special counting procedures for the content analysis, for example, the formula for calculating the Janis coefficient c, designed to calculate the ratio of positive and negative (towards the selected position) evaluations, opinions, arguments. The Janis coefficient can be used, for example, to calculate the ratio of positive and negative thoughts, highlighted in user comments about products that are sold through the e-commerce system. If the number of positive evaluations exceeds the negative ones, the Janis coefficient is calculated by the formula: $c = \dfrac{f^2 - f \cdot n}{r \cdot t}$, where f is the number of positive evaluations; n is the number of negative evaluations; r is the volume of the text content that has a direct relationship to the problem that is being investigated; t is the total volume of the analyzed text. If the number of positive evaluations is less than the negative, the Janis coefficient is found by a formula: $c = \dfrac{f \cdot n - n^2}{r \cdot t}$.

6. *Interpretation of the obtained results* according to the goals and objectives of the case study. Usually, at this stage, the characteristics of the text material are identified and evaluated, which allows one to draw conclusions about what the author wanted to emphasize or hide. Or, based on a statistical set of calculated Janis coefficients for a certain period of time for a specific category of products, you can predict changes in demand for the same products.

7 Advantages of content analysis in e-commerce

The use of content analysis of textual information in e-commerce systems, according to the authors, provides a number of advantages for simplifying the business dealing and solves many problems facing business process participants. Here are a few basic

advantages of content analysis of the text information in e-commerce systems:

- the automation of the textual information filtering, which is placed on the website of the e-commerce system by a user;
- the possibility to automatically create a *portrait* of a permanent user based on his/her comments;
- the possibility to automatically create a *portrait* of the target audience based on the analysis of *portraits* of permanent users;
- the reduction of the number of moderators who serve the e-commerce system;
- the reduction of the time for filling the site with the text information of a permanent user due to the automatic processing of this information and the absence of an intermediate link as a moderator;
- the language barrier elimination due to the automatic generation of dictionaries of the permanent user and the use of automatic translation.

The content analysis of the user text information distribution makes it possible to qualitatively evaluate the content flow in e-commerce systems for further decision-making by the relevant person. The comparison results shown in Fig. 1 approve that there is a need to pay attention to the content analysis in e-commerce systems.
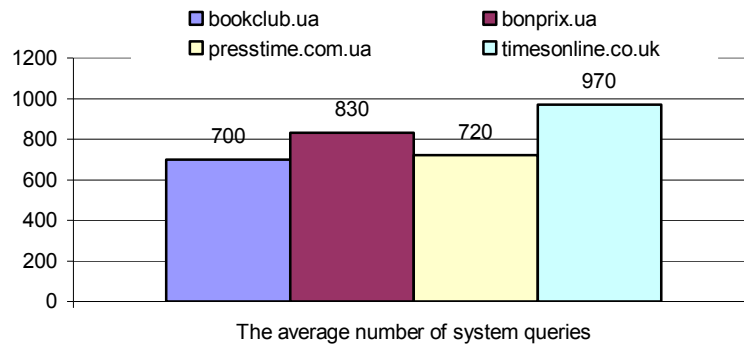

Fig. 1. The average number of the information queries to the system

8 Model and Method of Operative Content Management in Web Systems

The Web content life cycle presented in the form next major processes communication as *Source* $\rightarrow$ *content formation* $\rightarrow$ *content management* $\rightarrow$ *content implementation* $\rightarrow$ *database*.

Model of Web systems $S = \langle X, Formation, C, Management, Realization, Y \rangle$, where $X = \{x_1, x_2, \ldots, x_{n_X}\}$ is input data set, *Formation* is content formation operator, $C = \{c_1, c_2, \ldots, c_{n_C}\}$ is content set, *Management* is content management operator, *Realization* is content implementation operator, $Y = \{y_1, y_2, \ldots, y_{n_Y}\}$ is output data

set. Below is the content management models classification.

1. Pages generate per request is submitted in the form of the following main stages connection as $Content \rightarrow content\ editing \rightarrow Database \rightarrow content\ presentation \rightarrow informational\ resource$. Pages generate model on demand as $Management_Q = \langle X, C, Q, R, Edit, Y \rangle$, where X is input data set, C is content set, Y is pages generated set, Q is query set, R is pages formulation and submission function, $Edit$ – content editing and updating function.

2. Pages generate model while editing is presented as the next major stage of communication as $Content \rightarrow content\ editing \rightarrow database \rightarrow informational\ resource$.

When making changes to the site content creates a static *pages* set. Not is taken into account interactivity site between visitors and content. Pages generation system model while editing as $Management_E = \langle C, Edit, Y \rangle$, where C is content set, Y is static pages set, $Edit$ is content editing function. The pages formation described function as $\vec{y}(t) = Edit(\vec{c}, Weight, t)$.

3. Pages generate mixed model combines the advantages of the first two types and is presented as a communications major stages as $Content \rightarrow content\ editing \rightarrow Database \rightarrow content\ analysis \rightarrow blocks\ collection \rightarrow content\ presentation \rightarrow informational\ resource$. This model is as $Management_M = \langle X, C, Q, R, Edit, Caching, Y \rangle$, where X is input data set, C is content set, Y is pages generated set, Q is query set, R is pages formulation and submission function, $Edit$ is content editing and updating function, $Caching$ is cache formulation function. Cache is update automatically (after a certain period or when amending certain site sections) or manually (team administrator). Another approach is to maintain information blocks on the editing site stage. Then the pages collected from these units when requesting user. The process is implementing caching. The module generates submission page once. Then it is downloading from the cache several times faster.

User queries content analysis allows to qualitatively assess the content flow in the system. This facilitates the subsequent decisions by the moderator as follows: the problem situation description and study purpose search; precise definition of the study object and subject; the object preliminary analysis; concepts substantial clarification and empirical interpretation; the procedures description for the properties and phenomena registration; the overall study plan determining; the definition of the sample type, sources collection and so on. Qualitative content analysis is intended to provide the necessary means moderator for results analysis (Table 3). With their help identify the content properties and test them on the general content stream. Then apply the content stream general properties on its specific thematic part.

Table 3. Qualitative content analysis stages

Stage name	Stage characteristics
Text wrapping on blocks	Integrated content units are formation for encoding and processing.
Content stream reconstruction	The values, thoughts, views systems and arguments are reconstruction of each source text.

Stage name	Stage characteristics
Conclusions forming	Generalizations are withdrawal by comparing individual system values.

Quantitative content analysis consists of the stages presented in Table. 4.

Table 4. Quantitative content analysis stages

Stage name	Stage characteristics
The analysis unit selecting	Linguistic unit convert in the form for processing.
Units frequency counting	Relationships are identifying between linguistic units.
Categorization	Categories finite and excess aggregate are determining to obtain quantitative data of their appearance. Categories irregular sequence is clustering (into groups and classes division). And on the basis of new generalized categories is received.
Data Mining	New knowledge is identifying in the content flow through multiple quantitative evaluations. Next qualify them as categories.
Results interpretation	Content and semantically-filled results are getting. For this purpose use various statistical mathematical methods and semantic formalisms.

Main task for content management process are the following items: the operational and retrospective databases forming; databases rotation; users work personalization; a database create and access to it; personal needs and sources protection; work statistics keeping; search ensuring in database; output forms generation; interaction with databases of other subsystems. In Table 5 presents the content management key stages in the Web systems.

Table 5. Web content management stages

Stage name	Process name	Process features
Content editing	content themes definition	creating goal, the content and structure formation;
	form definition of content presentation	graphic information; the text (article, news release, job descriptions); HTML templates; back-end code, etc.;
	management tools selection	HTML editors, processors word; visual tools for creating objects;
Content analysis	rights access assignment	full or limited access to content;
	process identifying	standard processes of new information content creation/publishing;
	content saving	in a database or repository;
	processes logging	creation, transmission and storage processes;
	information interactive	information about the next performer content;
	events audit	content versions save;

Stage name	Process name	Process features
	text content analysis	quantitative or qualitative;
	versions access	support the possibility of users appealing to previous content versions;
	business process analysis	objectives, roles and tasks definition; roles default to user groups; business processes development for all content;
Content presentation	static	without any logic behavior;
	dynamic	personalization (rules/filters), globalization, localization.

As information technology basis considered annotated database in search engines. It contains an index, inverse, dictionary tables, etc. In content management systems creates a database search primary content pattern (PCP). They used clustering technology (automatic forming groups with similar content on the criteria PCP). In content management systems formed database annotations for used in the search process. Clusters database each record is corresponding cluster definitions and containing its description. Database record is performing automatic abstracting methods (digest is formation of text statistical analysis methods). These methods are used to create the PCP and descriptions of available users (Fig. 2).
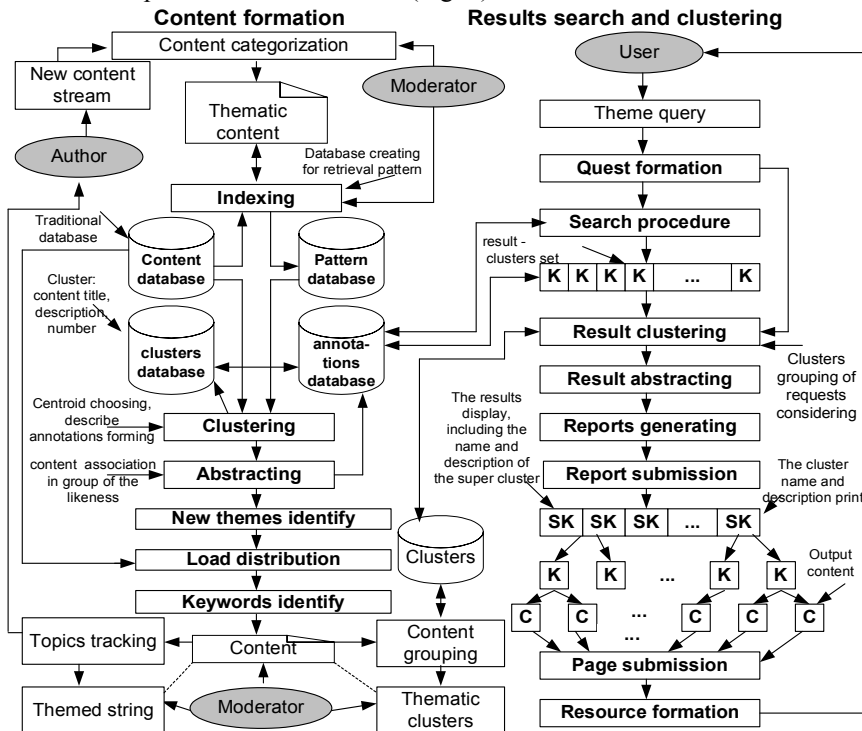

Fig. 2. Scheme of the Web systems functioning from annotated database

Personalization based rules is the content provision to specific users or users groups of conventional business logic using. For example, using a rule where all those interested in children's books fall into the group that focused advertising children's clothing. Rules developed on the content basis that type users in a registration card. In content management systems use algorithms categorization with personalization using filters (intelligent agents). Also, algorithms use based on the content analysis of user behavior. In particular, he analyzes the content to which the user accesses, the sites visited and more. That is constantly conducted analyzed registered user and user's group history with priorities an overwhelming number for interest.

The full-text search problem in large content arrays is ineffective. The annotated content search solves the problem exactly: instead of the full content searching to search on annotations (pattern). Digest remotely similar content and often not perceived by person. But as the search content pattern with weighted keywords and phrases it leads to adequate results with full-text search. Digest constructed from content fragments with the largest weight values. Content analysis is used for digests automatic generation, the concepts (categories) relationship automatic detection, relationships automatic clustering to the most important identify, the relationships automatic detection (e.g., positive and negative). One of the most important tasks in content analysis is the categorization process. It sets the conceptual grid. In its terms is the content flow analyzing and new categories generate. In Fig. 2-6 presented the results of the developed Web systems "Good morning, accountant" (dobry-jranok.com).
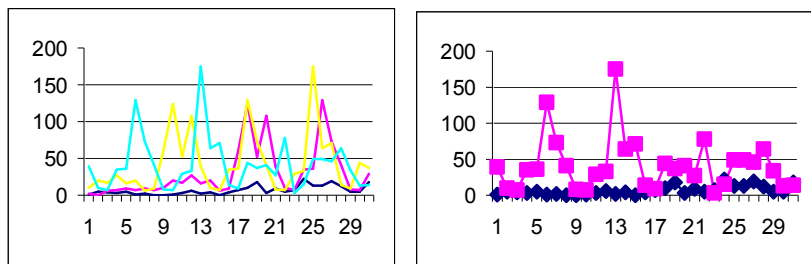
Fig. 3. Resource "Good morning, accountant" visit distribution for the 07-12/2012 period
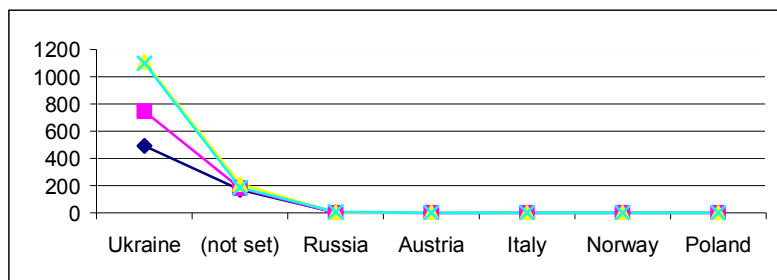
Fig. 4. Resource visit distribution from different countries for the 07-12/2012 period
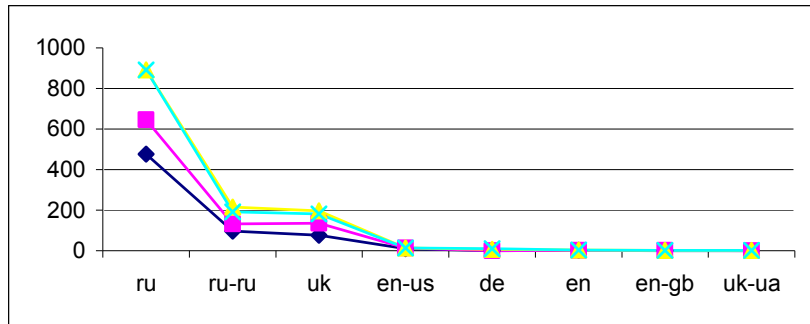
Fig. 5. Resource visit distribution from different languages for the 07-12/2012 period
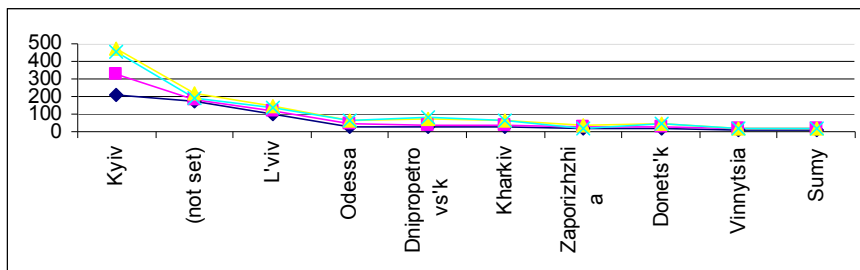

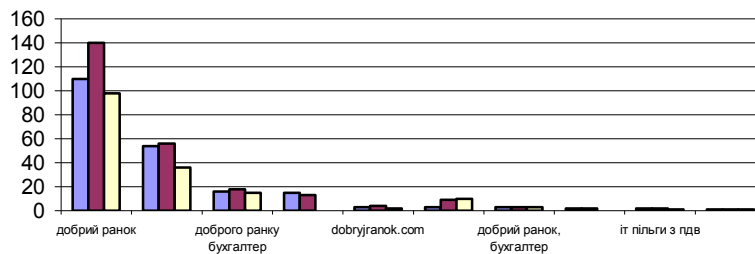**Fig. 6.** Resource visit distribution from different Ukraine cities for the 07-12/2012 period d

Fig. 7. Resource visit distribution of the search engines under the keyword content-search for the 09-12/2012 period

Globalization and localization is one of the modern market requirements. For successful marketing studies is use statistical basis. Globalization is more than a simple Web-pages conversion from one language to another. It still is to promote local brands (localization). Within globalization and localization is using content strategy, information hierarchy and navigation structure. For appropriate systems implement the following tasks: create content in languages that its list is predefined; global/local content preservation in different databases; support services for text automatic translation; content changes track and its approved.

9 Conclusion

In the article the justification is needed in the information resources processing methods and means development in electronic commerce content. This is implemented by the architecture improving of Web systems. Improving such systems is to automate processes of the Web content formation, management and implementation. This paper is analyzes the terminology and classification of Web systems. This is done to determine the characteristic design and simulation patterns, trends, process of Web systems. Also, the article outlines the shortcomings of existing content management methods and means. For Web systems is developed and described a content management model. This made it possible to develop a generic architecture for typical Web systems and standardized methods of information resources processing in these systems. The paper improved the overall architecture of Web systems. It is different from the existing modules presence of information resources processing. This is enabled the stages implementation of the content lifecycle. The content analysis of textual information in e-commerce systems allows us to determine the incidence of any feature of the studied set of texts. In such a case, it is important not so much absolute as the relative importance of the feature, that is, the characteristic of the place (share) among other features. For example, the percentage point of discussions by the forum users on the economic issues in relation to the political issues, or the percentage of positive comments on products in relation to the negative ones and with respect to all comments on this product category in the online store. Measuring the ratio between the features in the text gives the empirical material for understanding the functional relationships between the elements of reality reflected in the texts, for example, determining the mood of the audience of Forum on the economic or political situation in the country and the world. If there are chronologically ordered texts, you can have a number of portraits of the studied reality fixed over time (the change in demand for the product category depending on the season, for example, fantastic fiction is read more in the winter, and detective novels − in the summer) or portraits of the target audience (the change in demand for the product category according to the gender, for example, the demand for women's clothing is greater in September than in March). It allows you to hypothesize about the predictive nature of system elements functioning.

References

1. Lytvyn, V., Vysotska, V.: Designing architecture of electronic content commerce system. In: Computer Science and Information Technologies, Proc. of the X-th Int. Conf. CSIT'2015, 115-119 (2015)
2. Vysotska, V., Hasko, R., Kuchkovskiy, V.: Process analysis in electronic content commerce system. In: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT 2015, 120-123 (2015)
3. Vysotska, V.: Linguistic Analysis of Textual Commercial Content for Information Resources Processing. In: Modern Problems of Radio Engineering, Telecommunications and Computer Science, TCSET'2016, 709–713 (2016)
4. Chen, J., Dosyn, D., Lytvyn, V., Sachenko, A.: Smart Data Integration by Goal Driven Ontology Learning. In: Advances in Big Data. Advances in Intelligent Systems and Computing. – Springer International Publishing AG 2017. P. 283-292 (2017).

5. Mykich, K., Burov, Y.: Algebraic Framework for Knowledge Processing in Systems with Situational Awareness, Advances in Intelligent Systems and Computing, 217-228 (2016)
6. Mykich, K., Burov, Y.: Algebraic model for knowledge representation in situational awareness systems. In: Computer Sciences and Information Technologies, Proc. of the Int. Conf. CSIT, 165-167 (2016)
7. Chernukha, O., Bilushchak, Y.: Mathematical modeling of random concentration field and its second moments in a semispace with erlangian disrtibution of layered inclusions. In: Task Quarterly, 20(3), 295-334 (2016)
8. Kravets, P.: Game Model of Dragonfly Animat Self-Learning. In: Perspective Technologies and Methods in MEMS Design, 195-201 (2016)
9. Kravets, P., Kyrkalo, R.: Fuzzy logic controller for embedded systems. In: Proc. of the 5th Int. Conf. on Perspective Technologies and Methods in MEMS Design, (2009)
10. Kravets, P.: The control agent with fuzzy logic, Perspective Technologies and Methods in MEMS Design, MEMSTECH'2010, 40-41 (2010)
11. Vysotska, V., Chyrun, L., Chyrun, L.: Information Technology of Processing Information Resources in Electronic Content Commerce Systems, Computer Science and Information Technologies, 212–222 (2016)
12. Shakhovska, N., Vysotska, V., Chyrun, L.: Features of E-Learning Realization Using Virtual Research Laboratory. In: Computer Science and Information Technologies, Proc. of the XI–th Int. Conf. CSIT'2016, 143–148 (2016)
13. Khomytska, I., Teslyuk, V.: The Method of Statistical Analysis of the Scientific, Colloquial, Belles-Lettres and Newspaper Styles on the Phonological Level. In: Advances in Intelligent Systems and Computing, 512, 149-163 (2017)
14. Vysotska, V., Chyrun, L.: Analysis features of information resources processing. In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 124-128 (2015)
15. Vasyl, Lytvyn, Victoria, Vysotska, Dmytro, Dosyn, Roman, Holoschuk, Zoriana, Rybchak: Application of Sentence Parsing for Determining Keywords in Ukrainian Texts. In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 326-331 (2017)
16. Lytvyn, V., Vysotska, V., Uhryn, D., Hrendus, M., Naum, O.: Analysis of statistical methods for stable combinations determination of keywords identification. In: Eastern-European Journal of Enterprise Technologies, 2/2(92), 23-37 (2018)
17. Davydov, M., Lozynska, O.: Information System for Translation into Ukrainian Sign Language on Mobile Devices. In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 48-51 (2017).
18. Davydov, M., Lozynska, O.: Linguistic Models of Assistive Computer Technologies for Cognition and Communication. In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 171-175 (2017)
19. Khomytska, I., Teslyuk, V.: Specifics of Phonostatistical Structure of the Scientific Style in English Style System. In: Computer Science and Information Technologies, Proc. of the XI–th Int. Conf. CSIT'2016, 129-131. (2016)
20. Lytvyn, V., Vysotska, V., Pukach, P., Bobyk, I., Pakholok, B.: A method for constructing recruitment rules based on the analysis of a specialist's competences. In: Eastern-European Journal of Enterprise Technologies, 6/2(84), 4-14 (2016)
21. Zhezhnych, P., Markiv, O.: Linguistic Comparison Quality Evaluation of Web-Site Content with Tourism Documentation Objects. In: Advances in Intelligent Systems and Computing, 689, 656-667 (2018)