

*commercial content, information resource, business-process,
content management system, content lifecycle, electronic content commerce system*

*Victoria VYSOTSKA **, *Lyubomyr CHYRUN ***

METHODS AND MEANS OF PROCESSING INFORMATION RESOURCES IN ELECTRONIC CONTENT COMMERCE SYSTEMS

Abstract

This paper presents the development of unified methods and software tools for processing information resources in the electronic content commerce systems. A model of electronic content commerce systems is proposed. The models of information resource processing in electronic content commerce systems are proposed.

1. INFORMATION

Rapid development of the Internet contributes to the increase of needs for the efficient data of the production / strategic nature and implementation of new forms of information services through modern information technologies (IT) of e-commerce [1–3]. Documented information prepared in accordance with users needs is a *commercial content*. Today e-commerce is a reality and a promising business process. Internet is the business environment, and commercial content is a commodity with the highest demand and selling rate. It is also the main object of the processes of electronic content commerce. Commercial content can be immediately ordered, paid and got on-line as a commodity. The entire spectrum of commercial content is sold via the Internet – scientific and publicist articles, music, books, movies, pictures, software etc. Well-known corporations that implement electronic content commerce are Google through Google Play Market, Apple – Apple Store, Amazon – Amazon.com. Most of the decisions and researches are conducted at the level of specific projects. The electronic content commerce systems (ECCS) are built on the closed principle as non-recurrent projects. Modern ECCS is focused on the realization of commercial

* Information Systems and Networks Department, Lviv Polytechnic National University, S. Bandery Str., 12, Lviv, 79013, UKRAINE, +38 (032) 258-26-38, victana@bk.ru

** Software Department, Lviv Polytechnic National University, S. Bandery Str., 12, Lviv, 79013, UKRAINE, +38 (032) 258-25-78, chyrunlv@mail.ru

content that is made outside of the system. Design, development, implementation and maintenance of ECCS are impossible without the use of modern methods and information technologies of formation, management and support of commercial content [1–13].

2. RECENT RESEARCH AND PUBLICATIONS ANALYSIS

Development of the technology of information resources processing is important in view of such factors as lack of theoretical grounding of methods of study of commercial content flows and the need for unification of software processing methods of information resources in ECCS. A practical factor of the processing of information resources in ECCS is related with the solution of problems of formation, management and support of growing volumes of commercial content in the Internet, rapid development of e-business, widely spreader availability of the Internet, the expansion of the set of information products and services, and increasing of a demand for commercial content. Principles and IT of electronic content commerce are used while creating on-line stores (selling of e-Books, Software, video, music, movies, picture), on-line systems (newspapers, magazines, distance education, publishing) and off-line selling of content (copywriting services, Marketing Services Shop, RSS Subscription Extension), cloud storage and cloud computing. The world's leading producers of means of processing of information resources as Apple, Google, Intel, Microsoft, and Amazon are working in this area.

A theoretical factor of information resources processing in ECCS is connected with the development of IT processing of commercial content. In scientific studies of D. Lande, V. Furasheva, S. Braychevskoho, A. Grigoriev mathematical models of electronic processing of information flows are investigated and developed [1–3]. G. Zipf proposed an empirical law of distribution of word frequencies in natural language text content for its analysis. In the works of B. Boiko, S. McKeever, A. Rockley models of the life cycle of content are developed [4–13]. The methodology of content analysis for processing textual data sets was initiated and developed by M. Weber, J. Kaiser, B. Glaser, A. Strauss, H. Lasswell, O. Holsti, Ivanov, M. Soroka, and A. Fedorchuk. In the works of V. Korneev, A. F. Gareev, S. V. Vasyutina, V. V. Reich were proposed methods of intellectual processing of text information. EMC, IBM, Microsoft Alfresco, Open Text, Oracle and SAP have developed specification of Content Management Interoperability Services based on Web-services interface to ensure interoperability of electronic content commerce system management. From the scientific point of view, this segment of IT is not investigated enough. Each individual project is implemented almost from the very beginning, in fact, based on the personal ideas and solutions. In literature, very few significant theoretical studies, research findings,

recommendations for the design of ECCS and processing of information results in such systems are highlighted. Appeared a need to analyze, to generalize and to justify existing approaches to implementation of e-commerce and ECCS building are proposed. The actual problem of the creation of technological products complex is based on the theoretical study of methods, models and principles of processing information resources in ECCS, based on the principle of open systems that allow managing the process of increase in sales of commercial content. Analysis of the factors enables us to infer the existence of a contradiction between the active development and extension of IT and ECCS on the one hand, and the relatively small amount of research on this subject and their locality on the other. This contradiction raises the problem of containment of innovation development in the segment of electronic content commerce through creation and introduction of appropriate new advanced IT that affects negatively the growth of this market. Within this problem there is an urgent task of developing scientifically based methods of processing information resources of electronic content commerce, and building process on the basis of software for the creation, dissemination and sustainability of ECCS. In this paper a study to identify patterns, characteristics and dependencies in the processing of information resources in ECCS was carried out. This paper detailed analysis of the system's control of the content and prospects for implementation of ECCS are proposed. Analysis of business processes, the torrent of content, tools, and models of the systems' control of the content had been made. The methods and the means of the control of commercial content, their preferences and drawbacks had been described. The basic terms and concepts had been defined and concretized. The following of them had been used in the work.

The content is the totality of all data (commercial, service, extra, etc.) that implement a certain set of meta-models (a model that describes the structure and principles of a particular model) and the models of copies concentrated among information system [14–16]. The commercial content is a part of the general content, which is the subject of the purchase, the user's use and owner's profit; textual, visual or audio content as part of the user's experience according to the information resources (text, images, audio, video and software). The content control is control functions for receiving, analyzing, saving, searching and spreading of the content [14–16]. The information product is documented information prepared and designed to meet the needs of users. The information resource is an object of the means' action and information technology; set of documents in the information systems (libraries, archives, data banks, etc.) [1–3].

The e-commerce is a field of digital economy and of e-business, including all financial and commercial transactions over computer networks and business processes associated with conducting these transactions. E-commerce content is a field of e-commerce, where the commercial content is an object of financial and commercial transactions and business processes. The e-commerce content system is a system of processing of commercial content and related information,

human, technical, organizational and financial resources, to support and distribute commercial content. The content lifecycle is- a multi-complex process that takes place in the content control via the various stages/phases of the publication with a set of properties such as collaboration, records' control, digital asset and versions that are supported by various technologies [4–13].

Large torrents and volumes of different content are in ECCS. Most of these content's torrents are made up of easily formalized and automated procedures and commercial content. But there is no general approach to the process of modeling, design, development and implementation of ECCS.

The formal description of the e-commerce content is presented as

$$Y = \langle X, Q, C, V, H, Z, T, \delta \rangle, \quad (1)$$

where: $X = \{x_1, x_2, \dots, x_{n_x}\}$ – a set of content from various sources (information re-resources, authors, moderators, editors, visitors, journalists, users, administrators, analysts) (Fig. 1) [17–18],

$Q = \{q_1, q_2, \dots, q_{n_q}\}$ – set of users' information requests,

$C = \{c_1, c_2, \dots, c_{n_c}\}$ – a set of commercial content,

$V = \{v_1, v_2, \dots, v_{n_v}\}$ – a set of conditions content maintenance and external influences on the system environment,

$H = \{h_1, h_2, \dots, h_{n_h}\}$ – a set of processing content's conditions,

$Z = \{z_1, z_2, \dots, z_{n_z}\}$ – set of information resource's components,

$T = \{t_1, t_2, \dots, t_{n_t}\}$ – time of transaction processing content,

$Y = \{y_1, y_2, \dots, y_{n_y}\}$ – an outgoing characteristics' totality of system,

δ – an operator which form the statistics' analysis of ECCS's functioning [17–18].

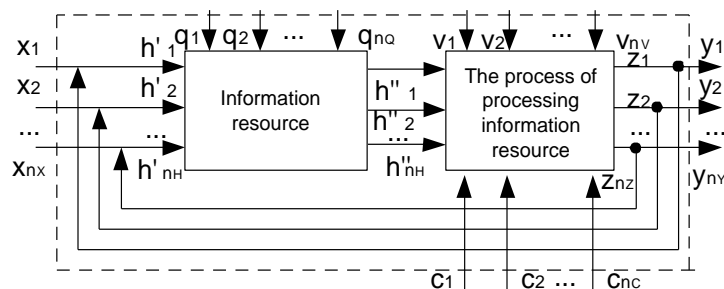


Fig. 1. Diagram of the process of e-commerce content [17–18]

The process which works up information resources (1) is described by

$$y_j(t_{p+1}) = \delta(x_i, q_d, c_r, v_l, h_k, t_p, z_w). \quad (2)$$

The value $y_j = \{y_{1j}, y_{2j}, \dots, y_{gj}\}$ is a totality of data over a specified period of time, where y_1 is number of visits, y_2 is average time of information resource's attendance (min: c), y_3 is a rate of refusals (%), y_4 is an achieved goal of a search, y_5 is content's dynamic (%), y_6 is the total number of viewed pages, y_7 is number of viewed pages per visit, y_8 are new visits, y_9 are absolute unique visitors, y_{10} is a traffic's source in % and so on.

The impact of the values x_i, q_d, c_r, v_l, h_k on the values z_w and y_j as a result of the e-commerce's content are unknown and unexplored. Connections between the input data, content, input data and the processing of information resources in the system are undisclosed. This justifies a goal, an actuality, expediency and a research's areas.

3. RESEARCH RESULTS ANALYSIS

The general design principles of ECCS's patterns make it possible to automate processing of resources to reduce production cycle, saving time and capacity of doing e-commerce. The main stages of the process of information resources' elaboration in ECCS are formation, control and maintenance of commercial content, with the following links: *content* → *content's formation* → *database* → *content's control* → *informational resource or user's request* → *content's control* → *informational resource* → *content's maintenance* → *database*. Then from (2) $\delta: X \rightarrow Y$ conveyed functions' superposition

$$\delta = \gamma \circ \beta \circ \alpha, \quad (3)$$

where: α – an operator of commercial content's formation,
 β – an operator of commercial content's control,
 γ – an operator of commercial content's maintenance.

The e-commerce content's system is presented as

$$Y = \langle X, Q, H, C, V, Z, T, \alpha, \beta, \gamma \rangle. \quad (4)$$

The operator of commercial content's formation α is a content's reflection c_r into new state c_{r+1} , that differs from the previous due to emergence of a new piece of content Δc which complements the previous state $c_{r+1} = c_r + \Delta c$, then

$$\alpha : (c_r, t_p, X, u_f) \rightarrow (c_{r+1}, t_{p+1}), \quad (5)$$

where $u_f = \{u_{1f}, u_{2f}, \dots, u_{n_{Uf}}\}$ is set of conditions of content's formation c_r .

Commercial content c_r is presented as

$$c_r = \left\{ \bigcup_i^{n_X} x_i \left| \begin{array}{l} \forall x_i \in X_{u_f}, x_i \notin X_{u_f}^-, \exists u_f \in U_{x_i}, u_f \notin U_{x_i}^-, \\ X = X_{u_f} \cup X_{u_f}^-, U = U_{x_i} \cup U_{x_i}^-, f = \overline{1, n_U} \end{array} \right. \right\}, \quad (6)$$

where the set of conditions u_f commercial content's formation c_r is defined as

$$u_f = \left\{ \bigcup_j^k u_{jf} \left| \begin{array}{l} \forall u_{jf} \in U_{x_i}, \exists x_i \in X_{u_f}, u_{jf} \notin U_{x_i}^-, \\ U = U_{x_i} \cup U_{x_i}^-, X_{u_f} \subseteq X, f = \overline{1, n_U}, i = \overline{1, m} \end{array} \right. \right\}. \quad (7)$$

The operator of commercial content's control β is a reflection of commercial content c_r into new state c'_r , which is different from the previous state due to values of the defining parameters $h_k \rightarrow h'_k$ (actuality, completeness, relevance, authenticity, trustworthiness) that satisfy predefined requirements

$$\beta : (q_d, z_w, c_r, h_k, u_M, t_p) \rightarrow (c'_r, h'_k, z_{w+1}, t_{p+1}), \quad (8)$$

where $q_d \in Q$, $h_k \in H$, $h_k = \{h_{1k}(c_r, q_d), \dots, h_{n_H k}(c_r, q_d)\}$ is set of conditions of commercial content's control, with it is presented as

$$z_w = \left\{ \bigcup_{r=1}^{n_C} c_r \left| \begin{array}{l} \forall c_r \in C_{q_d}, \exists q_d \in Q, \exists h_k \in H_{c_r}, c_r \notin C_{q_d}^-, h_k \notin H_{c_r}^-, \\ C = C_{q_d} \cup C_{q_d}^-, H = H_{c_r} \cup H_{c_r}^-, d = \overline{1, n_Q}, k = \overline{1, n_H} \end{array} \right. \right\}, \quad (9)$$

where the set of defining parameters' values form as $h'_k = h_k + \Delta h$.

The operator of commercial content's maintenance γ is a commercial content reflection c_r in the collection of values y_i , which is formed as result of the analysis, monitoring, evaluation of user's interaction, searching engines and other information resources that are the basis for making decisions about development and commercial content's control

$$\gamma: (c_r, q_d, v_l, h_k, z_w, u_s, t_p) \rightarrow y_i, \quad (10)$$

where $v_l = \{v_{1l}(q_i, h_k, c_r, z_w, t_p), \dots, v_{n_v l}(q_i, h_k, c_r, z_w, t_p)\}$ is set of conditions of content's maintenance and impact of the environment on the system. Outgoing data is implemented in

$$y_j = \left\{ \bigcup_l^{n_v} v_l \left| \begin{array}{l} \forall v_l \in V_{q_d} \cup V_{z_w}, \exists q_d \in Q, \exists z_w \in Z, \exists h_k \in H_{c_r}, \\ v_l \notin V_{q_d}, v_l \notin V_{z_w}, V_{q_d} \subset V, V_{z_w} \subset V, d = \overline{1, n_Q}, \\ w = \overline{1, n_Z}, r = \overline{1, n_C}, k = \overline{1, n_H} \end{array} \right. \right\}. \quad (11)$$

4. METHOD OF COMMERCIAL CONTENT FORMATION

The process of content's formation for an information resource provides a mapping of the input data from different sources to the set of formed commercial content and saved in an appropriate database in the electronic commerce content

$$S(x_i) \rightarrow x_i \rightarrow X \rightarrow \alpha(u_f, x_i, t_p) \rightarrow c_r \rightarrow C \rightarrow D(C), \quad (12)$$

where $S(x_i)$ is a data source, $D(C)$ is database of commercial content.

The process of commercial content's formation is presented as

$$\alpha = \langle X, T, U, C, \alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7 \rangle, \quad (13)$$

where $\alpha: X \rightarrow C$ is presented by superposition of functions

$$\alpha = \alpha_7 \circ \alpha_6 \circ \alpha_5 \circ \alpha_4 \circ \alpha_3 \circ \alpha_2 \circ \alpha_0, \text{ or } \alpha = \alpha_7 \circ \alpha_6 \circ \alpha_5 \circ \alpha_4 \circ \alpha_3 \circ \alpha_2 \circ \alpha_1. \quad (14)$$

The set $\{\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7\}$ is adequate in the content formation.

1. The operator of the commercial content's establishment is mapping of input data from various sources of information into content that is different from the previous state of the content due to its actuality.

$$\alpha_0 : (X, U_C, T) \rightarrow C_0. \quad (15)$$

2. The operator of the content's collecting is reflection of input data from the authors or systems' moderators in content that is different from the previous state of the content due to its trustworthiness and actuality.

$$\alpha_1 : (X, U_G, T) \rightarrow C_0. \quad (16)$$

3. The operator which identifies commercial content's duplication is a reflection of a commercial content into a new state that is different from the previous state according to its uniqueness.

$$\alpha_2 : (C_0, T, U_B) \rightarrow C_1. \quad (17)$$

4. The operator of content's formatting is display it's in a new state that is different from the previous state according to its format of presentation.

$$\alpha_3 : (C_1, U_{FR}, T) \rightarrow C_2. \quad (18)$$

5. The operator which identifies content's keywords is a content's reflection into a new state that is different from the previous state due to the presence of the set of keywords that describe the general content.

$$\alpha_4 : (C_2, U_K, T) \rightarrow C_3. \quad (19)$$

6. The operator which categorizes commercial content is content's reflection into a new state due to its validation, which is different from previous state due to it's belonging to the set of thematic content.

$$\alpha_5 : (C_3, U_{CT}, T) \rightarrow C_4. \quad (20)$$

7. The operator which forms content's digests is a content's reflection into a new state that is different from the previous state due to the emergence of a new piece of content as a summary of its complement previous state.

$$\alpha_6 : (C_4, U_D, T) \rightarrow C_5. \quad (21)$$

8. The operator of commercial content's selective distribution is a commercial content's reflection into a new state that is different from the previous state due to its purpose and spread among the target audience.

$$\alpha_7 : (C_5, U_{Ds}, T) \rightarrow C_6. \quad (22)$$

It is a complex of measures of providing data control from different sources to create commercial content with a set of additional values (relevance, credibility, uniqueness, completeness, accuracy etc.). The creation of commercial content is described by the operator $C_0 = \alpha_0(X, U_C, T)$, where U_C is a set of the conditions for creation of commercial content. The task of gathering information from the sources is described by the operator $C_0 = \alpha_1(X, U_G, T)$, where U_G is the set of conditions of data collection from various sources. The task of identifying of doubling commercial content matter is described by the operator α_2 in form $C_1 = \alpha_2(\alpha_0(X, U_C, T), U_B)$ and $C_1 = \alpha_2(\alpha_1(X, U_G, T), U_B)$, or $C_1 = \alpha_2(C_0, U_B)$, where U_B is the set of conditions of identifying and doubling matter of commercial content.

Identifying of duplicate commercial content by content in ECCS is made with help of linguistic-statistical methods of finding common terms, lines of which form a verbal signature of commercial content (the text is unique in the factor of uniqueness $\geq 80\%$).

The task of scanning the commercial content and the conversion to a common format in XML is described by the operator α_3 in form

$$C_2 = \alpha_3(\alpha_2(C_0, U_B), U_{FR}), \quad (23)$$

where U_{FR} is the set of conditions of content formatting. Processing of the content set C_2 to identify relevant keywords in meaning (terms) is based on the Zipf law and is reduced to the choice of words with an average frequency of occurrence (the most used words are ignored by using stop-dictionary, and rare words are not taken to account from text messages. Identifying keywords and concepts with the use of dictionaries is determined by the operator $\alpha_4(C_2, U_K)$ in form $C_3 = \alpha_4(\alpha_3(\alpha_2(C_0, U_B), U_{FR}), U_K)$ when $U_K = \{U_{K1}, U_{K2}, U_{K3}, U_{K4}\}$, where U_K is the collection of conditions to identify keywords and concepts in the text, U_{K1} is the set of all terms (a term is basis of the noun, noun, word combinations of the nouns or an adjective with a noun), U_{K2} is the set of frequencies of term use in the text of commercial content, U_{K3} is the set of coefficients of terms use that is based on the number of symbols without space

(2000–3000 symbols the frequency of keywords is in the range of 4–6%, up to 2000 symbols – 6–8%, more than 3000 symbols – 2–4%), U_{K4} is the set of terms which comply with the conditions of belonging to the keywords and concepts.

The set of digests C_5 is formed depending on $C_5 = \alpha_6(C_4, U_D)$, where U_D is the set of conditions of formation commercial content digests, that is $C_5 = \alpha_6(\alpha_5(\alpha_4(C_2, U_K), U_{CT}), U_D)$. Relevant content is sent to a user and loaded into the database. The sampling distribution of the content is described like $C_6 = \alpha_7(C_5, U_{DS})$, where U_{DS} is the set of conditions of the sampling distribution of content. Selective distribution of commercial content list $C_7 = \max(C_6)$ depends on the level of demand for this content. Associative rule of formation content list consists of original content list C_6 and the content list which is selected from original and derived list C_7 , in other words $C_6 \rightarrow C_7$. The formation of associative rule is a formation of content list that was formed through the merger of the original and derived lists. Associative rule of the appearance of the content from the original list along with the content from the list in the database is the following operator

$$P = \varphi(C_6, C_7) = \frac{\max(C_7 \cup C_6)}{\max(C_6)}, \quad (24)$$

where $d_s = \max(C_6)$ is the maximum of the set of content C_6 when $\forall c_{6i} \in C_6 \Rightarrow c_{6i} \leq d_s$; $d = \max(C_7 \cup C_6)$ is the maximum $C = C_6 \cup C_7$ with $\forall c_i \in C = (C_6 \cup C_7) \Rightarrow c_i \leq d$, then

$$d_s = \max(C_6) \Leftrightarrow d_s \in D_{C_6} \wedge \forall c_{faset} \in D_{C_6} : d_s \leq c_{faset}, \quad (25)$$

$$d = \max(C_6 \cup C_7) \Leftrightarrow d \in D_{C_6 \cup C_7} \wedge \forall c_{faset} \in D_{C_6 \cup C_7} : d \leq c_{faset}, \quad (26)$$

where $D_{C_6} = \{c_{faset} \in C_{faset} \mid \forall c_{6i} \in C_6 : c_{6i} \leq c_{faset}\}$ is the set of maxima values for C_6 at values of cardinalities set $\rho_{faset} = |C_{faset}|$, $\rho_6 = |C_6|$, $\rho_{faset} \geq \rho_6$; $D_{C_6 \cup C_7} = \{c_{faset} \in C_{faset} \mid \forall c_i \in (C_6 \cup C_7) : c_i \leq c_{faset}\}$ is the set of maxima for $C_6 \cup C_7$ with $C_{faset} \geq (C_6 \cup C_7)$.

Indicators of profitability and growth of demand for commercial content are used to determine its relevance and calculate respectively as

$$I_p = \kappa(C_7, P) = P \frac{\max(C_6 \rightarrow C_7)}{\max(C_7)}, \quad (27)$$

$$I_g = \chi(C_7, P) = \frac{\max(C_6 \rightarrow C_7) - \max(C_7)}{\max(C_6 \rightarrow C_7)[1 - P]}. \quad (28)$$

Derived list of rules determines the moderator. The list, the formation of which exceeds the minimum level is the most used content list. Even if this condition is formed by a large number of lists of frequently used content in accordance with requests. The result is limited when the sample by operators

$$I_r = \psi(C_6, C_7, P) = \min(\forall C'_6 \subset C_6, P - P'), \quad (29)$$

$$I_r = \min(\forall C'_6 \subset C_6, \varphi(C_6, C_7) - \varphi(C'_6, C_7)). \quad (30)$$

5. METHOD OF COMMERCIAL CONTENT MANAGEMENT

The process of content control is presented by the following scheme of links:

$$User(q_d) \rightarrow q_d \rightarrow Q \rightarrow H(c_r, q_d) \rightarrow \beta(q_d, c_r, h_k, t_p) \rightarrow z_w \rightarrow User(z_w), \quad (31)$$

where $User(q_d)$ is a formation of user's request; $User(z_w)$ is browsing by user's answers to a request q_d . The operator of commercial content's control $\beta: C \rightarrow Z$ is presented as superposition of functions

$$\beta = \beta_4 \circ \beta_3 \circ \beta_2 \circ \beta_1. \quad (32)$$

The management process of commercial content is presented as

$$\beta = \langle C, Q, H, U, T, Z, \beta_1, \beta_2, \beta_3, \beta_4 \rangle. \quad (33)$$

1. The operator of the editing and modification of commercial content

$$\beta_1 : (c_r, h_k, u_l, t_p) \rightarrow c'_r. \quad (34)$$

2. The determination operator of the block weight and formation the base search images of commercial content

$$\beta_2 (c'_r, y_j, u_l, t_p) \rightarrow c''_r. \quad (35)$$

3. The operator of the determinant meanings control parameters formation

$$\beta_3 : (c''_r, h_k, u_l, t_p) \rightarrow h'_k. \quad (36)$$

4. The operator of information resource page formation and presentation

$$\beta_4 (c''_r, h'_k, z_w, q_d, t_p) \rightarrow z_{w+1}, \quad (37)$$

where $h_k \in H$, $h_k = \{h_{1k}, h_{2k}, \dots, h_{mk}\}$ is the set of process parameters that control commercial traditional content (h_{1k} – actuality, h_{2k} – relevance, h_{3k} – completeness, h_{4k} – authenticity, h_{5k} – authenticity of commercial content); $u_l \in U$, $u_l = \{u_{1l}, u_{2l}, \dots, u_{nl}\}$ is multiple criteria process to control commercial content (u_{1l} – the coefficient of the block location in the commercial content, u_{2l} – the coefficient of keywords in the block, u_{3l} – coefficient of the key words statistical importance, u_{4l} – the coefficient of keywords from the user request, the coefficient of the keywords volume from the request). It is a support measures for determining parameters of commercial content (actuality, completeness, relevance, authenticity, reliability) in accordance with certain requirements by the set of criteria for the management of commercial content.

Processes classification of commercial content management.

1. The management of commercial content to generate pages on demand of the user of the electronic content is commerce is field as

$$\beta_Q = \langle C, Q, H, U, T, Z, \beta_1, \beta_2, \beta_3, \beta_4 \rangle. \quad (38)$$

The stage of editing and modification of commercial content is served by the operator $c_j(t_{r+1}) = \beta_1(c_j, t_r, h_k, u_l)$ where $c_j(t_{r+1}) \in C$. The stage of development of pages set is described by the operator, where

$$z_i = \left\{ \bigcup_{j=1}^m c_j(q_i, t_r) \left| \begin{array}{l} \forall c_j \in C_q, c_j \notin C_q^-, C_q = \beta_3(\beta_2(C_q)), \exists q_i \in Q_c, \\ \exists h_k \in H_c, h_k \notin H_c^-, C = C_q \cup C_q^-, Q_c \subset Q, \\ H = H_c \cup H_c^-, k = \overline{1, n_H}, i = \overline{1, n}, r = \overline{1, w} \end{array} \right. \right\}. \quad (39)$$

The block weight is defined as the sum of the coefficients of the scales of commercial content:

$$\omega = \|C\| = \beta_2(C, \omega_1, \omega_2, \omega_3, \omega_4, \omega_5), \quad (40)$$

where: $\omega_1(c_j)$ – the coefficient of the block location in the content,
 $\omega_2(c_j)$ – the coefficient of keywords in the block,
 $\omega_3(c_j)$ – the coefficient of statistical significance of terms,
 $\omega_4(c_j)$ – the coefficient of the availability of additional terms,
 $\omega_5(c_j)$ – the coefficient of the volume of terms from the user request.

2. Management of commercial content with pages generation during editing information resource by moderator is presented as

$$\beta_E = \langle C, H, T, Z, \beta_1, \beta_2, \beta_3 \rangle. \quad (41)$$

Stage of pages set forming is described by operator: $Z(t_r) = \beta_3(C, H, t_r, \beta_1, \beta_2)$.

3. Mixed type of management of commercial content is presented as

$$\beta_M = \langle C, Q, H, T, Z, W, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5 \rangle, \quad (42)$$

where: W – the set of commercial cached content,
 β_5 – the operator of formation of the set of cached commercial content or information blocks when $W = \beta_5(C, \beta_3(\beta_2(\beta_1(C, t_r, H, U))), t_{r+1})$ or $W = \beta_5(Z, \beta_3(\beta_2(\beta_1(C, t_r, H, U))), t_{r+1})$,

$$w_l = \left\{ \bigcup_{i=1}^n c_i \mid \forall c_i \in C_Q, C_Q \subset C, C_Q = \beta_3(\beta_2(C)) \right\}, \quad (43)$$

$$w_l = \left\{ \bigcup_{j=1}^m z_j \mid z_j \in Z_c, \forall c_j \in C_z, \exists c_j \in Z_c, \forall c_j \in z_j, C_z = \beta_3(\beta_2(C)), C_z \subset C, Z_c \subseteq Z, i = \overline{1, n} \right\}. \quad (44)$$

6. METHOD OF COMMERCIAL CONTENT SUPPORT

The process of commercial content support is presented as scheme of links $User(q_d, z_w) \rightarrow q_d \rightarrow z_w \rightarrow V(q_d, z_w, t_p) \rightarrow \gamma(v_l, h_k, c_r, z_w, t_p) \rightarrow y_j \rightarrow User(y_j)$.

The process of commercial content engineering support is presented as

$$\gamma = \langle Q, C, H, V, T, Z, Y, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7, \gamma_8 \rangle. \quad (45)$$

Content support $\gamma : Z \rightarrow Y$ is represented by a superposition of functions

$$\gamma = \gamma_8 \circ \gamma_6 \circ \gamma_5 \circ \gamma_3 \circ \gamma_1, \text{ or } \gamma = \gamma_8 \circ \gamma_7 \circ \gamma_5 \circ \gamma_4 \circ \gamma_2. \quad (46)$$

1. The operator of formation digital flows portraits of commercial content is a mapping of the set of relevant commercial content in a variety of settings, which describe the thematic needs of the target audience according to certain criteria that is defined by the moderators.

$$\gamma_1 : (V_{Pc}, C, H, Q, T) \rightarrow Y_{Pc}. \quad (47)$$

2. The operator of formation portraits of regular users is the mapping of the relevant commercial content set in a variety of settings which describe the thematic needs of the target audience according to certain criteria that is defined by the moderators.

$$\gamma_2 : (V_{Pq}, Q, H, Z, T) \rightarrow Y_{Pq}. \quad (48)$$

3. The operator of identification of thematic subjects in the plural of new content which is mapping of new commercial content set from a variety of sources of information in the set of keywords for new rubric of commercial content which describe a topic sentence of these reliable sources according to certain criteria that is defined by the moderators.

$$\gamma_3 : (C, H, X, V_T, T) \rightarrow Y_T. \quad (49)$$

4. The operator of content thematic subjects identification with a set of user requests is the mapping of multiple user requests to the set of keywords for the new content rubric which describe thematic needs of registered users according to certain criteria that is defined by moderators.

$$\gamma_4 : (C, H, Q, V_T, T) \rightarrow Y_T. \quad (50)$$

5. The operator of tabulation of the commercial content relations by keywords and frequency of visits is the mapping of commercial content in a new state, which is different from the previous large number of links content based on criteria such as thematic, the relevance factor rating, sequence and frequency of viewing, popularity, actuality, authorship.

$$\gamma_5 : (C, V_c, T) \rightarrow Y_C. \quad (51)$$

6. The operator of calculating the ratings of commercial content is the mapping of content to a new state which is different from the previous state of commercial content by the emergence of new content in the form of ratings on certain criteria that complements the previous state.

$$\gamma_6 : (C, Q, H, Y_C, V_{Rc}, T, \theta, \vartheta) \rightarrow Y_{Rc}. \quad (52)$$

7. The operator of calculating the ratings of regular users is the mapping of the set permanent portraits of classified users in a new state, which is different from the previous condition of commercial content by the emergence of a new part of the characteristics of these users in the form of ratings on certain criteria that complements the previous state.

$$\gamma_7 : (C, Q, H, Y_C, V_{Rm}, T) \rightarrow Y_{Rm}. \quad (53)$$

8. The operator of the statistical analysis of system functioning is the mapping of statistic system functioning in a collection of values, which create as result of analysis, monitoring, evaluation of user interaction, search engines and other information resources, which is the basis of making decisions regarding to the creation and management the content.

$$\gamma_8 : (Y_P, Y_T, Y_C, Y_R, Z, H, V, T) \rightarrow Y. \quad (54)$$

A analysis result of the S e-business system functioning and C commercial Web content support is formed set $Y = \{Y_P, Y_T, Y_C, Y_R\}$ under the conditions $V = \{V_P, V_T, V_C, V_R\}$, where $Y_P = Y_{Pc} \cup Y_{Pq}$ is a subset of the information portraits of Y_{Pc} content and Y_{Pq} users, Y_T is a subset of thematic storyline of content, Y_C is subset of content relationship tables, $Y_R = Y_{Rc} \cup Y_{Rm}$ is a subset of the rating content Y_{Rc} and moderators Y_{Rm} , $V_P = V_{Pc} \cup V_{Pq}$ is the conditions set of information portraits formation, V_T is a conditions set for thematic storyline identification, V_C is the conditions set of the content relationship construct tables, V_R is the parameters set of the content ratings calculation. The information portraits set of content is presented as $Y_{Pc} = \gamma_1(V_{Pc}, C, H, Q, T)$, and set of the users` portraits are given as $Y_{Pq} = \gamma_2(V_{Pq}, Q, H, Z, T)$.

The thematic storyline set for the content is presented as $Y_T = Y_{TC} \cup Y_{TQ}$, where $Y_{TC} = \gamma_3(C, H, X, V_T, T)$ is condition set for content storyline identification in the of new commercial content and $Y_{TQ} = \gamma_4(C, H, Q, V_T, T)$ is thematic storyline definition. The set of relationship content tables is presented as $Y_C = \gamma_5(C, V_c, T)$. The set of the content rating is presented as $Y_{Rc} = \gamma_6(C, Q, H, Y_C, V_{Rc}, T, \theta, \xi)$, and a set of a moderators ratings Y_{Pq} is presented as $Y_{Rm} = \gamma_7(C, Q, H, Y_C, V_{Rm}, T)$, where $V_R = V_{Rc} \cup V_{Rm}$ is the parameters set for the content ratings calculation, $\theta(Q^+, Q^0, Q^-, T, H)$ is the tonality criteria for commercial content, $\xi(Q, T)$ is operator of comments filtering definition. The Y set of output statistical data is presented as

$$Y = \{Y_P, Y_T, Y_C, Y_R\} = \gamma(V_P, V_T, V_C, V_R, C, Q, H, Z, T). \quad (55)$$

Commercial Web content management subsystem is implemented through caching (representation module generates a page once; then it is several times faster loaded from the cache, which is updated automatically after a certain period of time or when making changes to specific sections of an information resource, or manually by administrator command) or information blocks formation (blocks conservation in the information resources editing stage and page collection from these blocks at the user request of the relevant page).

Web content support subsystem provides information portraits formation, thematic storyline identification in content flows, the content relationship tables building, content rankings calculation, new events identification in their content flows, their tracking and clustering. Analysis of commercial support content helps identify causes of the formation of the target audience for a set of characteristics of functioning of e-commerce content.

The subsystems presence of commercial Web content creation, management and support in an e-commerce content increases sales volume of commercial content to the permanent user at 9%, active involvement of unique visitors, prospective users and expand the limits of the target and regional audience by 11%, viewed pages by 12% , visiting time and information resources by 7%.

7. CONCLUSIONS

The paper is solved the actual scientific problem of methods and means research and development for commercial Web content processing in e-business systems by using the developed mathematical software for the appropriate systems creation, which made it possible maintain the life cycle of commercial content on the level of developer (the time and costs reducing for development, quality improvement through the use of proven solutions).

Software tools for content creation, management and support are developed. Design and implementation methods of electronic content commerce systems are based on online newspapers, which reflect the results of theoretical research, are developed. From the perspective of a systemic approach, the principles of applying information resources processing in electronic content commerce systems for content lifecycle implementation made the development of methods for the commercial content formation, management and support possible. An integrated method of commercial content formation for the time and resources reduction of content production is developed. This makes it possible to create a means of information resources processing and implement subsystem of automatically generated content. A method of commercial content management for the time and resources reduction of content sales was created, which makes it possible to implement commercial content management subsystem. A method of commercial content support for the time and resource reduction of the target audience analysis in electronic content commerce systems is implemented, which makes it possible to develop a commercial content support subsystem.

REFERENCES

- [1] BOLSHAKOVA E., LANDE D., NOSKOV A., KLYSHINSKY E., PESKOVA O., YAGUNOVA E.: *Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i kompyuternaya lingvistika*. (in Russian). Moskva: MIEM 2011, pp. 272.
- [2] LANDE D., FURASHEV V., BRAYCHEVSKY S., GRIGOREV O.: *Osnovy modelirovaniya i otsenki elektronnyh informatsionnyh potokov*. (in Ukraine). Kyiv: Inzhiniring, 2006, pp. 348.
- [3] LANDE D.: *Fundamentals information streams integration*. (in Ukraine). Kyiv: Engineering Publ., 2006, p. 240.
- [4] BOIKO B.: *Content Management Bible*. (in USA). Hoboken, 2004, pp. 1176.

- [5] CM LIFECYCLE POSTER: *Content Management Professionals*. (in USA). Retrieved 20 July 2010, <http://www.cmprosold.org/resources/poster/>.
- [6] DOYLE B.: *Seven Stages of the CM Lifecycle*. (in USA). EcontentMag.com. 2010, www.econtentmag.com/Articles/ArticleReader.aspx?ArticleID=13554&AuthorID=155.
- [7] HACKOS J.: *Content Management for Dynamic Web Delivery*. (in USA). Hoboken, NJ: Wiley, 2002, pp. 432.
- [8] HALVERSON K.: *Content Strategy for the Web*. (in USA). Reading, Mass: New Riders Press, 2009, pp. 192.
- [9] MCGOVERN G., NORTON R.: *Content Critical*. (in USA). Upper Saddle River, NJ: FT Press, 2001, pp. 256.
- [10] MCKEEVER S.: *Understanding Web content management systems: evolution, lifecycle and market* (in USA). Industrial Management & Data Systems (MCB UP), 2003, 103 (9), pp. 686–692.
- [11] NAKANO R.: *Web content management: a collaborative approach* (in USA). Boston: Addison Wesley Professional, 2002, pp. 222.
- [12] WOODS R.: *Defining a Model for Content Governance*, 2010, www.contentmanager.net/magazine/article_785_defining_a_model_for_content_governance.html.
- [13] ROCKLEY A.: *Managing Enterprise Content: A Unified Content Strategy*. (in USA). Reading, Mass: New Riders Press, 2002, pp. 592.
- [14] ISO/IEC/IEEE 24765: 2010 (E): *Systems and software engineering – Vocabulary*. (in USA). Michigan State University, 2010, pp. 418.
- [15] ISO/IEC 15474-1: 2002: *Information technology – CDIF framework – Part 1: Overview*. (in Switzerland). ISO copyright office, 2002, pp. 8.
- [16] ISO/IEC 2382-1: 1993: *Information technology -- Vocabulary -- Part 1: Fundamental terms*. (in USA). 1993, http://www.iso.org/iso/catalogue_detail.htm?csnumber=7229.
- [17] SOVETOV B., YAKOVLEV S.: *Systems modeling* (in Russian). Moskva: VS Publ., 2001, pp. 343.
- [18] BERKO A., VYSOTSKA V., PASICHNYK V.: *Systemy elektronnoyi content-komertsiyi* (in Ukraine). Lviv: NULP, 2009, pp. 612.