

SET-THEORETIC MODEL FOR ELECTRONIC CONTENT COMMERCE SYSTEMS

© Berko A., Vysotska V., Chyrun L., 2014

General problems of electronic content commerce and functional services of content processing are analyzed in the article. Proposed model gives an opportunity to create a method and instrument of information resources processing in electronic commerce systems and to implement the subsystem of content formation, management and support.

Keywords - information resources, commercial content, content analysis, content monitoring, content search, electronic content commerce systems.

Introduction and the general problem formulation

Active development of the Internet promotes the needs growth in the operative reception of data for e-business conducting. It also facilitates the new forms implementation of information services through appropriate systems and information resources [1-2]. Commercial Web content is documented and according formatted unique data. They are prepared to meet the needs of information resources users [1-2, 5]. The characteristic feature of e-business systems is an opportunity to information resources processing. This increases the goods sales volume of regular user, the active involvement of potential users and expanding the boundaries of the target audience [4-5, 7-8]. The e-business principles and technology are used actively in the creation of electronic content commerce systems for on-line/off-line sales and Web content analysis/exchange/saving, online shop, cloud storage/computing, etc. [2]. The problems of design, development, implementation and maintenance of electronic content commerce systems are actual theme in view of factors such as: active development of research in the e-business area, the theoretical justification lack of standardized methods, and the need for software unification of information resources processing [2]. There are new approaches and solutions to this problem. But there is a mismatch between the known methods and means of commercial content processing and principles of e-commerce systems construction. Lacks are general approaches of e-commerce systems creation and standardized methods for commercial Web content processing in these systems.

Recent research and publications analysis

Practical factor of commercial Web content processing in e-business systems is closely linked with the problems addressing of the content amounts growth in the Internet, rapid prevalence of accessibility to the Internet, active development of e-business, set expanding of information products and services, growth of demand for information products and services, the technology and means creation, and areas expanding of methods application for information resources processing. In this direction are actively working leading global manufacturers of commercial content resources means. In particular, such systems are developed: for on-line content selling of online newspapers (New York Times, Washington Post), online journals (Wall Street Journal, Chip), distance learning (Ashford University), Internet publication as a dictionaries/directories (Oxford English Dictionary Online), Internet publishing (Online Publishers Association), portals of

entertaining (tochka.net), informative (Weather Channel, Karpaty.ua) and children (teremoc.ru, vshkolu.com) content; for off-line content selling of copywriting services (Textbroker, Sopywriting 911), Marketing Services Shop (The Copy Box, Local Internet Marketing), RSS Subscription Extension (Apple, Google, Intel, Microsoft); online shop for selling of eBooks, software, video, music, movies, picture, digital art, manuals, articles, certificates, forms, files, etc (Google, Amazon, Apple, Android, Opera, Yahoo!); for content saving on cloud storage and cloud computing (Google, Amazon, Apple, Micorsoft, Linux, iPhone, Android, Palm).

The theoretical factor of information resources processing in e-business systems is associated with the methods and means development of commercial Web content formation, management and support. In scientific studies D. Lande is mathematical models of electronic information flow studied and developed [3-4, 6-8]. G. Zipf is proposed an empirical regularity of words frequency distribution in natural language. And G. Salton and R. Papka is proposed new events detection in the content flow. B. Boiko, S. McKeever, A. Rockley, G. McGovern, J. Hackos, R. Nakano, B. Doyle, R. Woods, Halverson are a model of the commercial content lifecycle in the works described. A. Tenni, B. Matthews, D. Spiida, J. Kaiser, Glaser are the content analysis methodology launched. And H. Lasswell, O. Holsti, F. Joubish, K. Neuendorf and K. Krippendorff are its developed. EMC, IBM, Microsoft Alfresco, Open Text, Oracle and SAP have developed a Content Management Interoperability Services specification for Web-services. They are designed to provide interaction between subsystems of commercial Web content processing in e-business systems. Interoperable Content Application Tools interacts with content from different repositories via a service interface and special superstructure as Content Management Interoperability Services Implementation.

Problems extraction

The number of Web content streams in e-business systems is greater than the path of goods movement in the industry. Considerable part of content flow in electronic business systems consists of easily formalized and automated procedures. The main problem is the lack of a common approach to process of e-business systems modeling, design and development for the rapid processing of commercial content.

Model of e-business system is given as

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

where $X = \{x_1, x_2, \dots, x_{n_x}\}$ – content set from various sources, $Q = \{q_1, q_2, \dots, q_{n_q}\}$ – set of information users requests, $C = \{c_1, c_2, \dots, c_{n_c}\}$ – commercial Web content set, $V = \{v_1, v_2, \dots, v_{n_v}\}$ – set of content support conditions, $H = \{h_1, h_2, \dots, h_{n_h}\}$ – the conditions set of content formation and management, $Z = \{z_1, z_2, \dots, z_{n_z}\}$ – the components set of an information resource, $T = \{t_1, t_2, \dots, t_{n_t}\}$ – transaction time of content processing, $Y = \{y_1, y_2, \dots, y_{n_y}\}$ – statistics data set of system operation, *Function* – the operator of the output data formation in the system [2].

The process of the e-business system operation is described by the *S* operator as

$$y_j(t_p + \Delta t) = Function(x_i, q_d, c_r, v_l, h_k, t_p, z_w), \quad (2)$$

where x_i – visitors/users requests to the e-business system. Value y_j is described as $y_j = \{a_1, a_2, \dots, a_g\}$, where a_1 – visits number over a time period Δt , a_2 – the average time spent on an information resource (min : c) for a time period Δt , a_3 – a refusals indicator (%) for a period time Δt , a_4 – achieved purpose of commercial content finding, a_5 – the dynamics of commercial content lifecycle (%), a_6 – the total number of page views for a time period Δt , a_7 – the number of page views per visit, a_8 – new visits, a_9 – absolute unique visitors, a_{10} – traffic sources (direct referrals, referrals from search engines, referrals from other sites, etc.) etc. [5].

The value influence x_i, q_d, c_r, v_l, h_k on z_w and y_j value as a result of the e-business system operation are unknown and unexplored. The model does not reveal relationships between the input data, commercial Web content, the original data and the information resources processing in e-business system S . This substantiates the purpose, relevance, appropriateness and research directions of the commercial content processing in e-business systems. It is necessary to solve the problems of the Web content formation, management and support as components of the information resources processing in e-business systems in the form of theory-based concept.

Goals formulation

This work aim is to expand the functionality and the efficiency mproving of e-business systems. This is achieved by system attendance increasing through the use of standardized methods and software for commercial Web content processing. The purpose is identified the need to the following tasks solving:

- models creation of e-business systems to the shortcomings determine of existing methods and tools for commercial Web content processing;
- standardized methods development of commercial Web content processing for the architecture designing of e-business systems;
- a generalized architecture development of e-business system for stages implementation of the commercial Web content lifecycle;
- software implementation for commercial Web content processing in e-business systems (for the time and costs reducing on commercial content formation, management and support) by the quality increasing of information resources processing.

Obtained scientific results analysis

The main stages of the commercial Web content processing in e-business systems is the content formation, management and support with the following links: *Web content* \rightarrow *Web content formation* \rightarrow *database* \rightarrow *Web content management* \rightarrow *information resource* or *user request* \rightarrow *Web content management* \rightarrow *information resource* \rightarrow *Web content support* \rightarrow *database*, ie

$$Function = Formation \cup Management \cup Support, \quad (3)$$

where *Function* – the operator of the output data formation in the system, *Formation* – the operator of the Web content formation, *Management* – operator of Web content management, *Support* – operator of Web content support.

Under the model of e-business system is presented as

$$S = \langle X, Q, Formation, H, C, V, Management, Support, Z, T, Y \rangle, \quad (4)$$

where X – Web content set from various sources, Q – set of user queries, *Formation* – the operator of the content formation, H – the conditions set of Web content formation and management, C – set of commercial Web content, V – the conditions set the content support and external influences on the system, *Management* – Web content management operator, *Support* – Web content support operator, Z – the components set of an information resource, T – transaction time of information resources processing, Y – statistical data set of the system operation.

Phase of commercial Web content formation is described operator *Formation* as follows

$$c_r = Formation(u_f, x_i, t_p), \quad (5)$$

where u_f – conditions set of commercial Web content formation, ie. $u_f = \{u_1(x_i), \dots, u_{n_U}(x_i)\}$.

Commercial Web content is presented as follows:

$$c_r = \left\{ \bigcup_f u_f \mid (x_i \in X) \wedge (\exists u_f \in U), U = U_{x_i} \vee U_{x_i}^-, i = \overline{1, m}, f = \overline{1, n} \right\}. \quad (6)$$

Stage of commercial Web content managing is described operator *Management* as follows

$$z_w = \text{Management}(q_d, c_r, h_k, t_p), \quad (7)$$

where Q – set of information users requests, H – set of commercial Web content management conditions, ie. $H = \{h_1(c_{i+1}, q_d), \dots, h_{n_H}(c_{i+n_H}, q_d)\}$. Commercial Web content management is presented as

$$z_w = \left\{ \bigcup_{k=1}^{n_H} h_k(c_{i+1}, q_d) \mid \begin{array}{l} (c_{i+k} \in C) \wedge (q_d \in Q) \wedge (h_k \in H_q), \\ H = H_{q_d} \vee H_{q_d}^-, i = \overline{1, n_C}, d = \overline{1, n_Q}, k = \overline{1, n_H} \end{array} \right\}. \quad (8)$$

Stage of commercial Web content support is described operator *Support* as follows

$$y(t_p + \Delta t) = \text{Support}(v_l, h_k, c_r, z_w, t_p, \Delta t), \quad (9)$$

where v_l – conditions set of commercial Web content support and the environment impact on the system, ie. $v_l = (v_1(q_i, h_k, c_r, z_w, t_p), \dots, v_{n_V}(q_i, h_k, c_r, z_w, t_p))$. Output data are implemented

$$y_j = \left\{ \bigcup_l v_l \mid \begin{array}{l} (\exists q_d \in Q) \wedge (\exists z_w \in Z) \wedge (\forall v_l \in V) \wedge (\forall (c_r \wedge q_d) \in h_k), \\ V = V_{q_d} \vee V_{q_d}^-, d = \overline{1, n_Q}, l = \overline{1, n_V}, w = \overline{1, n_Z}, r = \overline{1, n_C}, k = \overline{1, n_H} \end{array} \right\}. \quad (10)$$

A complete functional electronic business system is characterized a complex system of interrelated operations, methods and techniques of commercial Web content processing.

Method of commercial Web content forming

The commercial Web content formation for an information resource provides communication between inputs data set from different sources and generated content set that is stored in the appropriate database of e-business systems, ie $Source(x_i) \rightarrow x_i \rightarrow X \rightarrow Formation(u_f, x_i, t_p) \rightarrow c_r \rightarrow C \rightarrow DataBase(C)$, where $Source(x_i)$ – the x_i content source, x_i – the i -th content from data source, X – content set from corresponding sources of data, $Formation(u_f, x_i, t_p)$ – the operator of the a commercial content formation at x_i in a fixed time t_p at the conditions u_f , c_r – formed under conditions u_f of r -th commercial content, C – a set of commercial content, $DataBase(C)$ – database of commercial Web content.

Model of commercial Web content formation is presented as

$$Formation = \left\langle \begin{array}{l} X, Gathering, Formatting, KeyWords, Backup, \\ Caterizati on, BuDigest, Dissemination, T, C \end{array} \right\rangle, \quad (11)$$

where X – set of input data $x_i \in X$ from different information resources or moderators, *Gathering* – operator of content collecting/creating from various sources, *Formatting* – content formatting operator, *KeyWords* – the

operator of content keywords and concepts identification, *Caterization* – the operator of automatic categorization of content, *Backup* – the operator of content duplication identification, *BuDigest* – the operator of the content digest formation, *Dissemination* – operator of selective distribution of Web content, T – the transaction time of $t_p \in T$ the content formation, C – a Web content set $c_r \in C$.

The optimal solution can help navigate in the dynamic input data from different sources. This optimal solution provide the process of information gathering from the sources and its fragments distribution according to the users needs $C = \text{Gathering}(X, U_G, T)$, where X – content set from different sources of data, U_G – a conditions set of data gathering from a different sources, *Gathering* – the operator of content gathering/creation, T – the content gathering/creation time.

Content duplication identification is described by the operator $C = \text{Backup}(\text{Gathering}(X, U_G, T), U_B)$, where X – content set from different sources of data, U_B – conditions set content duplication identification, *Backup* – the operator of content duplication identification, C – a content set. Content duplication identification in the e-business systems implements linguistic statistical methods of the general terms serching, which form chain of a verbal signatures in content. Content syndicate is in data gathering programs training of structural features at individual sources (with information resources, by moderators, users, visitors, journalists, editors), direct scanning of commercial content and his adduction to a common XML format $C = \text{Formatting}(\text{Backup}(\text{Gathering}(X, U_G, T), U_B), U_{FR})$, where *Formatting* – the operator of content formatting, U_{FR} – a conditions set of commercial content formatting.

The content set processing C for meaningful keywords identification are based on the principles of the keywords finding in content (terms). This are based on Zipf's law and reduced to the words choice wich an average frequency of appearance (the most used words are ignored by stop-dictionary using and rare words from messages text not included). The keywords and concepts identification are defined by the operator $C = \text{KeyWords}(\text{Formatting}(\text{Backup}(\text{Gathering}(X, U_G, T), U_B), U_{FR}), U_K)$, where U_K – conditions set of key words and concepts identification. Commercial Web content classification and distribution implement through a information retrieval system of selective content distribution (the containing router). Content is analyzed content categorization operator $C_C = \text{Categorization}(\text{KeyWords}(C, U_K), U_C)$ for compliance with thematic requests, where *Categorization* – operator of content categorization according to the detected keywords, U_C – a conditions set of an automatic categorization, C_C – set of classified relevant content.

The digest set C_D form depending $C_D = \text{BuDigest}(C_C, U_D)$, where *BuDigest* – the digest formation operator, U_D – conditions set of digest formation, ie $C_D = \text{BuDigest}(\text{Categorization}(\text{KeyWords}(C, U_K), U_C), U_D)$. Relevant content send users and are loaded into the database. Selective distribution of commercial content is described $C_{Ds} = \text{Dissemination}(C_D, U_{Ds})$.

Method of commercial Web content management

The process of commercial Web content management is presented this general scheme links: $User(q_d) \rightarrow q_d \rightarrow Q \rightarrow H(c_{i+1}, q_d) \rightarrow \text{Management}(q_d, c_r, h_k, t_p) \rightarrow z_w \rightarrow User(z_w)$. Methods for commercial content management is divided into different classes according to the model of pages generating of information resources in e-business systems.

1. Model of commercial Web content management such as pages generating by the user requests is presented as

$$\text{Management}_Q = \langle H, C, Q, Presentation, Edit, Weight, T, Z \rangle, \quad (12)$$

where C – Web content set, Z – set of generated pages on information resource, Q – a set of user queries, $Presentation$ – the operator page formulation and presentation on information resource, $Edit$ – operator of content editing and modification, $Weight$ – the weight of the content block, T – transaction time of commercial Web content management, H – a conditions set of commercial content management.

Step of content editing and modification is presented by the operator $c_j(t_{r+1} + \Delta t) = Edit(c_j(t_r), t_r, \Delta t)$. Stage a pages set forming is described operator $Z(t_r + \Delta t) = Presentation(q_i, C, Weight(C), t_r, \Delta t)$, where $Weight(C)$ – the total weight of content block on information resource, ie

$$z_i = \left\{ \begin{array}{l} \bigcup c_j(q_i, t_r) | (\forall c_j \in C_q) \wedge (\exists q_i \in C_q) \wedge Weight(C_q), \\ C = C_q \vee C_{\bar{q}}, j = \overline{1, m}, i = \overline{1, n} \end{array} \right\}. \quad (13)$$

2. Model of pages generation while the information resource editing in the system is presented as

$$Management_E = \langle C, H, Edit, Weight, T, Z \rangle, \quad (14)$$

where C – Web content set, H – conditions set of content management, Z – static pages set of information resource, $Weight$ – the weight of the Web content block, $Edit$ – operator of Web content editing/updating. Phase of pages formation is described operator $Z(t) = Edit(C, Weight, H, t)$.

3. Model of mixed management of commercial content is presented as

$$Management_M = \langle C, Q, H, Presentation, Edit, Weight, Caching, T, Z \rangle, \quad (15)$$

where C – content set, Z – generated pages set of information resource, Q – the set of user queries, H – the set of content management conditions, $Presentation$ – the operator of page formulation and presentation on information resource, $Edit$ – operator of content editing and modification, $Caching$ – the operator of the cache or information content blocks formation where the cache is presented as $Cache = Caching(C, Weight, t, \Delta t)$ or $Cache = Caching(Z, Weight, t, \Delta t)$ ie

$$Cache = \left\{ \bigcup c_i | (\forall c_i \in C) \wedge Weight(c_i), t = t + \Delta t, i = \overline{1, n} \right\}, \quad (16)$$

$$Cache = \left\{ \bigcup z_i | (\exists z_i \in Z) \wedge (\forall c_j \in C) \wedge Weight(c_j) \wedge (\forall c_j \in z_i), t = t + \Delta t, i = \overline{1, n} \right\}. \quad (17)$$

Method of commercial Web content supporting

Process of commercial Web content support is presented as follows scheme links $User(q_d, z_w) \rightarrow q_d \rightarrow z_w \rightarrow V(q_d, z_w, t_p) \rightarrow Support(v_l, h_k, c_r, z_w, t_p, \Delta t) \rightarrow y_j \rightarrow Moderator(y_j)$.

The method of commercial Web content support looks like

$$Support = \left\langle \begin{array}{l} Q, C, H, V, T, BuInfPort, IdThemTop, \\ ConCorrTablConc, CalRankConc, Z, Y \end{array} \right\rangle, \quad (18)$$

where Q – set of user queries, C – the Web content set, $BuInfPort$ – the operator of the information portraits formation, $IdThemTop$ – the operator of thematic storyline identification, $ConCorrTablConc$ – operator content relationship tables building, $CalRankConc$ – the operator of content ratings calculation, H – set of internal parameters in the system, V – the parameters set of the external environment, Z – page set of information

resource, T – set of transaction time for information resources processing, Y – set of statistical data of the system operation.

A analysis result of the S e-business systems 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 $Y_P = Y_{P_c} \vee Y_{P_q}$ – a subset of the information portraits of Y_{P_c} content and Y_{P_q} users, Y_T – a subset of thematic storyline of content, Y_C – subset of content relationship tables, Y_R – a subset of the ratings content, $V_P = V_{P_c} \vee V_{P_q}$ – the conditions set of information portraits formation, V_T – a conditions set for thematic storyline identification, V_C – the conditions set of the content relationship construct tables, V_R – the parameters set of the content ratings calculation.

The information portraits set Y_{P_c} of content is presented as $Y_{P_c} = BuInfPort(V_{P_c}, C, H, Q, T)$, and a Y_{P_q} set of the users portraits are given as $Y_{P_q} = BuInfPort(V_{P_q}, Q, H, Z, T)$, where $V_P = V_{P_c} \vee V_{P_q}$ – set of the portraits formation conditions, $BuInfPort$ – the operator of the portraits formation as $Y_P = Y_{P_c} \vee Y_{P_q}$. The thematic storyline set for the Y_T content is presented as $Y_T = IdThemTop(C, H, Q, V_T, T)$, where V_T – conditions set for content storyline identification, $IdThemTop$ – the operator of Y_T thematic storyline definition.

The Y_C set of relationship content tables is presented as $Y_C = ConCorrTablConc(C, V_C, T)$, the where V_C – the conditions set of the relationship tables building, $ConCorrTablConc$ – tables constructing operator.

The Y_{R_c} set of the content ratings is presented as $Y_{R_c} = CalRankConc(C, Q, H, Y_C, V_{R_c}, Spam, Tonality, T)$, and a Y_{P_q} set of a moderators ratings is presented as $Y_{R_m} = CalRankConc(C, Q, H, Y_C, V_{R_m}, T)$, where $V_R = V_{R_c} \vee V_{R_m}$ – the parameters set for the content ratings calculation, $Tonality(Q^+, Q^0, Q^-, T, H)$ – the tonality criteria for commercial content, $Spam(Q, T)$ – operator of comments filtering definition, $CalRankConc$ – operator of content and moderators ratings identification $Y_R = Y_{R_c} \vee Y_{R_m}$.

The Y set of output statistical data are is presented as

$$Y = \{Y_P, Y_T, Y_C, Y_R\} = Support(V, C, Q, H, Z, T, \Delta T), \quad (19)$$

$$Y = \{Y_P, Y_T, Y_C, Y_R\} = Support(V_P, V_T, V_C, V_R, C, Q, H, Z, T, \Delta T), \quad (20)$$

where $Y_P = Y_{P_c} \vee Y_{P_q}$ – the information portraits subset of content and users, Y_T – a subset of the thematic storyline in content, Y_C – a subset of content relationship tables, $Y_R = Y_{R_c} \vee Y_{R_m}$ – a subset of the content and moderators ratings, $Support$ – operator of content support.

The obtained data taken into account when an information resource creation or updating and the architecture improvement of e-business systems. Actuality of subsystem development for commercial content processing consists in the necessity obtain operational/objective assessment of the competition level in the financial market segment of commercial content, assess the competition level and their competitiveness degree in the financial market with the content distribution. In Table. 1 lists the developed systems with the lifecycle support for commercial content.

Table 1

Subsystems of commercial content processing in the developed e-business systems

| № | Information Resource Name | System Appointment | Address | Subsystem availability | | |
|---|---------------------------|--------------------|-----------------------------|------------------------|------------|---------|
| | | | | Formation | Management | Support |
| 1 | Fotogalereja-vysocjkykh | Online shop | fotogalereja-vysocjkykh.com | +/- | + | +/- |
| 2 | Vgolos | Online Newspaper | vgolos.com.ua | + | + | + |
| 3 | Tatjana | Online Newspaper | tatjana.in.ua | - | +/- | +/- |
| 4 | Presstime | Online Newspaper | presstime.com.ua | +/- | + | + |
| 5 | AutoChip | Online shop | autochip.vn.ua | - | + | +/- |

| | | | | | | |
|---|----------------------------|------------------|-------------------|-----|-----|-----|
| 6 | Kursyvalyut | Online Journal | kursyvalyut.com | + | + | - |
| 7 | Good morning, accountant! | Online Newspaper | dobryjranok.com | + | + | +/- |
| 8 | Victana | Online Journal | victana.lviv.ua | +/- | +/- | +/- |
| 9 | Information for Businesses | Online Journal | goodmorningua.com | - | - | - |

There are developed software for commercial Web content formation, management and support. There are posted software of developed systems implementation with subsystems of information resources processing in e-business organization over the Online Newspaper and Online Journal.

Table 2

The results of systems operation in the time period from 10.2010 to 03.2013

| Characteristics | The developed e-business systems | | | | | | | | |
|------------------------|----------------------------------|----------|-------|----------|-------|-------|-------|-------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Visiting | 4144 | 5997052 | 646 | 3654456 | 7838 | 9124 | 8724 | 517 | 25 |
| Visits time | 4:50 | 2:14 | 6:22 | 2:04 | 1:58 | 1:30 | 2:27 | 8:38 | 8:12 |
| Failures/refusals | 55,62 | 71,90 | 47,5 | 83,08 | 54,45 | 82,76 | 68,15 | 35,8 | 48,0 |
| Goal Conversion | 4,35 | 0,00 | 0,00 | 0,00 | 10,97 | 0,00 | 0,00 | 0,00 | 0,00 |
| Unique visitors | 2674 | 2501402 | 256 | 1501202 | 5704 | 7632 | 4996 | 250 | 7 |
| Repeat visits | 1470 | 3495641 | 393 | 2153254 | 2134 | 1500 | 3728 | 267 | 18 |
| Pageviews | 19082 | 11588861 | 2924 | 11588861 | 20828 | 16952 | 18892 | 3760 | 81 |
| Pages/visit | 4,60 | 1,93 | 4,51 | 1,67 | 2,66 | 1,86 | 2,17 | 7,27 | 3,24 |
| % New Visits | 64,48 | 41,68 | 39,3 | 39,88 | 72,75 | 83,56 | 57,23 | 47,9 | 28,0 |
| New Visitor | 64,53 | 41,71 | 39,5 | 34,76 | 72,77 | 83,58 | 57,27 | 48,4 | 28,0 |
| Returning Visit | 35,47 | 58,29 | 60,5 | 65,24 | 27,23 | 16,42 | 42,73 | 51,6 | 72,0 |
| Search traffic | 56,47 | 36,10 | 32,5 | 26,12 | 69,67 | 79,68 | 24,34 | 23,0 | 20,0 |
| Direct Traffic | 19,50 | 11,20 | 38,5 | 31,22 | 17,15 | 14,17 | 48,73 | 32,9 | 80,0 |
| Transitions (%) | 24,03 | 52,48 | 28,97 | 42,46 | 13,18 | 6,15 | 26,94 | 44,10 | 0,00 |
| Language – ru | 45,15 | 55,57 | 33,90 | 43,54 | 66,87 | 54,39 | 69,01 | 49,7 | 52,0 |
| Language – uk | 36,20 | 23,70 | 57,3 | 33,34 | 10,23 | 25,51 | 10,49 | 39,3 | 48,0 |
| Ukraine | 81,25 | 91,22 | 83,9 | 87,56 | 67,17 | 66,01 | 89,47 | 87,8 | 60,0 |
| Russia | 9,10 | 1,84 | 6,78 | 4,67 | 24,04 | 5,57 | 0,62 | 0,39 | 0,00 |
| Poland | 0,19 | 0,29 | 0,01 | 0,12 | 0,11 | 0,61 | 0,11 | 0,19 | 0,00 |
| United States | 1,21 | 0,66 | 0,01 | 0,32 | 0,01 | 0,57 | 0,07 | 0,01 | 0,00 |
| Germany | 0,34 | 0,32 | 0,15 | 0,11 | 0,20 | 0,26 | 0,08 | 0,01 | 0,00 |
| City – Kyiv | 17,25 | 23,62 | 29,1 | 25,45 | 20,09 | 22,86 | 36,16 | 28,4 | 48,0 |
| City – L'viv | 38,92 | 43,41 | 48,4 | 34,54 | 10,17 | 14,43 | 9,70 | 45,1 | 12,0 |
| Dnipropetrovsk | 2,82 | 2,72 | 0,00 | 0,45 | 3,22 | 2,51 | 6,63 | 1,35 | 0,01 |
| City – Kharkiv | 1,50 | 2,82 | 0,62 | 0,65 | 3,56 | 1,65 | 6,08 | 1,16 | 0,01 |
| City – Ternopil | 2,24 | 0,01 | 0,01 | 0,01 | 0,01 | 2,17 | 0,01 | 0,01 | 0,01 |

Table 2 shows the comparative results characteristics of the developed e-business systems operation over the time period 10-11.2012. Statistical data is obtained from Google Analytics, which provides advanced features for data analysis. Google Analytics is a free service for visits maintenance to Web-sites. It allows us to estimate the information resources traffic in e-business systems and the effectiveness of marketing activities, such as for Online Newspaper and Online Journal. Fig. 1-2 are present the work results of the developed systems in the form of graphs. So the all stages presence of the commercial content lifecycle significantly increases the visits and unique users amount of information resources in e-business systems.

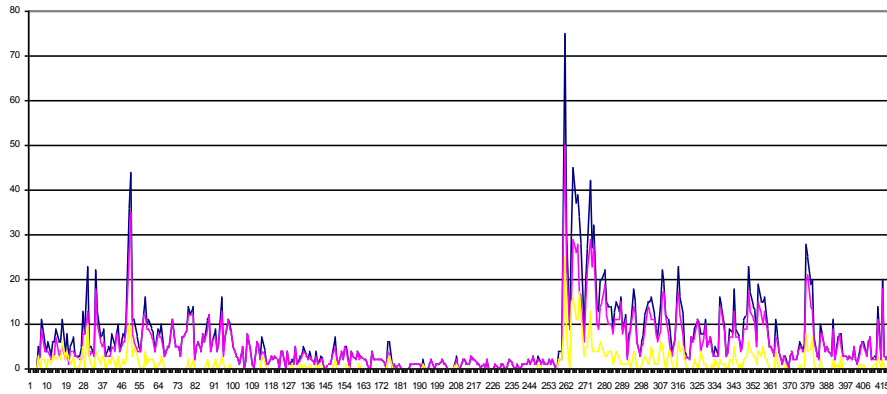


Fig. 1. Distribution of a) time of arrival at the site and b) traffic

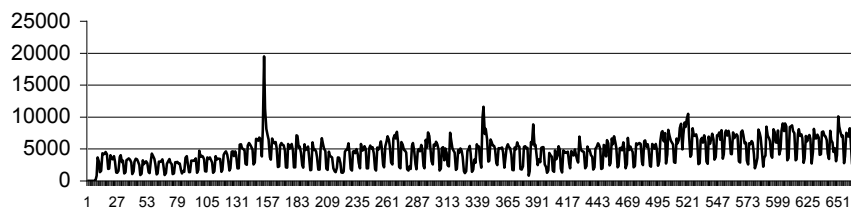


Fig. 2. Distribution of new visits and failures in information resources visiting

Service for accesses counting of information resource allows us to estimate the information resource traffic and marketing activities efficiency, such as for the Online Newspaper Vgolos (Fig. 3).

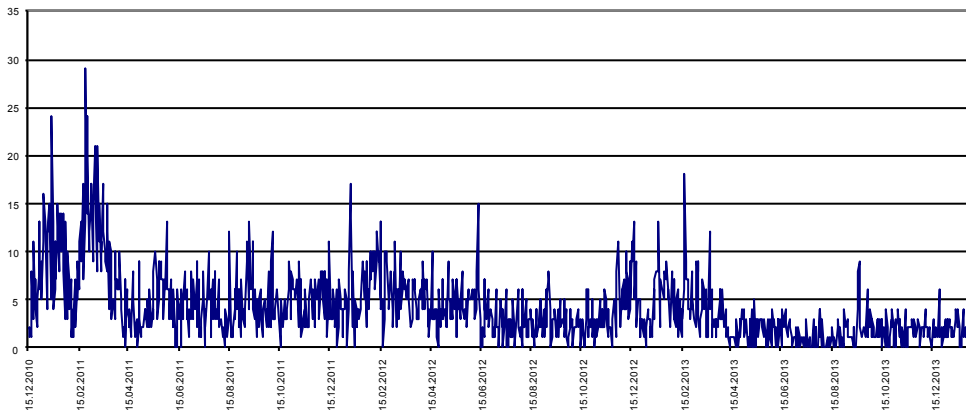


Fig. 3. Overall average number of visitors for years 2010-2012

The subsystems presence of commercial Web content creation, management and support in e-business systems increases the number of unique visitors, viewed pages, visiting time and information resources, the target and regional audiences increasing.

The subsystem of Web content formation is implemented as a content-monitoring complex for content gathering from different sources of data that provide a content database creation of according to the information needs of users. As a gathering and primary processing result of content is reduced to a single format, classified according to the specified rubricator. And he is credited descriptors with keywords. This facilitates the process implementation of content management.

Task of Web content management subsystem are: database formation, rotation and providing access to it; the operational and retrospective databases formation; the user experience personalization; personal user

queries and sources storing; operation statistics analysis; search providing in database; initial forms generation on information resources; information interaction with other databases; the an information resource formation. 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 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.

Conclusions and recommendations for further scientific studies

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. It is substantiated the necessity of methods and means development for the content processing in e-business systems by the system architecture improving to the processes automate of the content formation, management and support. In the work is analysed e-business terminology for determination of their characteristic patterns, trends, process of e-business systems design and simulation and the shortcomings of existing methods and tools for content processing. It is developed e-business system model, which is helped to develop the commercial content lifecycle, the generalized system architecture and standardized methods for content processing in these systems. There is improved overall system architecture for e-business systems, which is different from the existing architectures of availability of commercial content processing modules. This is enabled the stage implementation of the content lifecycle. In this paper is developed a complex methods of commercial Web content formation, management and support to achieve the work effect at the developer level (the time and costs reducing for development, quality improvement through the use of proven solutions). It is developed the general guidelines for the architecture designing of e-business systems, which are different from the existing systems more detailed steps and the presence of commercial Web content processing modules They allow effectively implement a information resources processing at the developer level of e-business systems (the time and resources reducing to the development, the quality improvement of e-business systems operation). In this work is developed the modules architecture of e-business systems for the stages implementation of the commercial Web content lifecycle. It is developed and implemented a software application of commercial Web content formation, management and support to achieve the effect of work at the level of the owner (profitability increasing, user interest growing) and user (clarity, the interface simplification standardization, unification, choice expanding) of e-business systems.

1. Береза А. *Електронна комерція* / А. Береза, І. Козак, Ф. Левченко. – К: КНЕУ, 2002. – 326 с.
2. Берко А. *Системи електронної контент-комерції* / А. Берко, В. Висоцька, В. Пасічник. – Л: НУЛП, 2009. – 612 с.
3. Большакова Е. *Автоматическая обработка текстов на естественном языке и компьютерная лингвистика* / Е. Большакова, Д. Ландэ, А. Носков, Э. Клышинский, О. Пескова, Е. Ягунова. – М: МИЭМ, 2011. – 272 с.
4. Брайчевский С. *Современные информационные потоки* / С. Брайчевский, Д. Ландэ // *Научно-техническая информация*. – 2005. – № 11. – С. 21-33.
5. Клифтон Б. *Google Analytics: профессиональный анализ посещаемости веб-сайтов* / Б. Клифтон. – М: Вильямс, 2009. – 400 с.
6. Корнеев В. *Базы данных. Интеллектуальная обработка информации* / В. Корнеев, А. Гареев, С. Васютин, В. Райх. – М: Нолидж, 2000. – 352 с.
7. Ландэ Д. *Основы моделирования и оценки электронных информационных потоков* / Д. Ландэ, В. Фурашев, С. Брайчевский, О. Григорьев. – К: Інжиніринг, 2006. – 348 с.
8. Ландэ Д. *Основы интеграции информационных потоков: монографія* / Д. Ландэ. – К: Інжиніринг, 2006. – 240 с.