ANALYTICAL METHODS FOR COMMERCIAL WEB CONTENT PROCESSING OF INFORMATION RESOURCE IN ELECTRONIC BUSINESS SYSTEMS

Victoria Vysotska
Lviv Polytechnic National University, Lviv, Ukraine

Lyubomyr Chyrun
Lviv Polytechnic National University, Lviv, Ukraine

© MESTE NGO

JEL category: D84, L86

Abstract
The given article is devoted to the development of standardized analytical methods and software for Web content processing in e-business systems. The model of electronic content commerce systems is developed. The models of commercial Web content processing subsystems are constructed. A new approach of business processes application and implementation for the construction of electronic content commerce systems is formulated. Complex methods of formation, management and support commercial Web content are developed. Software for information resources processing in electronic content commerce systems is developed. The methods of e-business systems designing and implementation are presented as example on online newspaper and online magazine that reflect the theoretical studies results.

Keywords: content, web content, information resource, business process, online newspaper, online magazine, e-commerce systems, e-business.

1 INTRODUCTION
Active development of the Internet promotes the growth of needs in the operative reception of data for e-business conducting. It also facilitates the implementation of new forms of information services through appropriate systems and information resources (Berko, Vysotska, & Pasichnyk, 2009), (Jerk, 2001), (Uspensky, 2001), (Pasichnyk, Scherbyna, Vysotska, & Shestakevych, 2012), (Vysotska, & Chyrun, 2011). 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 (Jerk, 2001), (Evdokimov, 2007). The e-business principles and technology are used actively in the creation of electronic content commerce systems for online/off-line sales and Web content analysis/exchange/saving, online shop, cloud storage/computing, etc. (Ermakov, & Kiselev, 2005), (Ivanov, & Krukovskaya, 2004). (Ivanov, 2002), (Ivanov, 1994), (Clifton, 2009). 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 (Ivanov, & Krukovskaya, 2004), (Ivanov, 2002), (Clifton, 2009), (Korneev, Hareev, Vasyutyn, & Reich, 2000), (Krupnik, 2002). 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. There are the lacks of general approaches of e-commerce systems creation and standardized methods for commercial Web content processing in these systems.

2 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, expanding set of information products and services, growth of demand for information products and services, the technology and tools 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, vshkolou.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, 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 tools development of commercial Web content formation, management and support. In scientific studies D. Lande studied and developed mathematical models of electronic information flow (Bolshakova, Lande, Noskov, Klyshynskyy, Peskova, & Yahunova, 2011), (Braychevskyy, & Lande, 2005), (Grigoriev, & Lande, 2005), (Lande, Furashev, Braychevskyy, & Hryhorev, 2006), (Lande, 2005-2006), (Lande, & Braychevskyy, 2005-2006), (Lande, & Litvin, 2001), (Lande, & Morozov, 2004-2005), (Lande, & Furashev, 2006), (Lande, Furashev, Braychevskyy, & Hryhorev, 2006), (Lande, Furashev, & Grigor'yev, 2006), (Furashev, Lande, Grigor'yev, & Furashid, 2005), (Furashev, Lande, & Braychevskiy, 2005). G. Zipf proposed an empirical regularity of words frequency distribution in natural language (Zipf 1935, 1949). G. Salton and R. Papka have proposed new events detection in the content flow (Salton, 1979), (Papka, 1999). B. Boiko, S. McKeever, A. Rockley, G. McGovern, J. Hackos, R. Nakano, B. Doyle, Halverson have described a model of the commercial content lifecycle in the works (Boiko, 2005), (Doyle, 2005), (Hackos, 2002), (Halverson, 2009), (McGovern, & Norton, 2001), (McKeever, 2003), (Nakano, 2002), (Rockley, & Cooper, 2002), (Stone, 2003), (Vysotska, & Chyrun, 2011). In works (Pasichnyk, Scherbyna, Vysotska, & Shestakevych, 2012), (Soroka, & Tanatar, 1998), (Fedorchuk, 2005) have developed the content analysis methodology launched. EMC, IBM, Microsoft Alfresco, Open Text, Oracle and SAP a Content Management Interoperability Services specification for Web-services Berko, Vysotska, & Pasichnyk, 2009). They have 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.

3 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

\[ S = \{X, Q, C, V, H, F, \text{Function}, T, Z, Y\}, \] (1)

where:
- \( X = \{x_1, x_2, ..., x_n\} \) – content set from various sources;
- \( Q = \{q_1, q_2, ..., q_m\} \) – set of information users requests;
- \( C = \{c_1, c_2, ..., c_c\} \) – commercial Web content set;
- \( V = \{v_1, v_2, ..., v_v\} \) – set of content support conditions;
- \( H = \{h_1, h_2, ..., h_h\} \) – the conditions set of content formation and management,
- \( Z = \{z_1, z_2, ..., z_r\} \) – the components set of an information resource,
- \( T = \{t_1, t_2, ..., t_t\} \) – transaction time of content processing,
- \( Y = \{y_1, y_2, ..., y_y\} \) – statistics data set of system operation,
- Function – the operator of the output data formation in the system (Berko, Vysotska, & Pasichnyk, 2009), (Sovetov, & Yakovlev, 1998).

The process of the e-business system \( S \) operation (1) is described by the operator as:

\[ y_j(t_p + \Delta t) = \text{Function}(x_i, q_d, c_r, v_l, h_k, t_p, z_w) \] (2)

where \( x_i \) – visitors/users requests to the e-business system. Value \( y_j \) is described as:

\[ y_j = \{a_1, a_2, ..., a_g\} \]

The value influences \( x_i, q_d, c_r, v_l, h_k, t_p, z_w \), 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 \( S \) from (1). 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.

4 GOALS FORMULATION

The aim of this work is to expand the functionality and the efficiency improving 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 by 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 and formatted 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.

5 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, i.e. due to (2) as

$$\text{Function} = \text{Formation} \cup \text{Management} \cup \text{Support} \quad (3)$$

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

Due to (1) and (3) the model of e-business system is presented as

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

where $X$ – Web content set from various sources, $Q$ – set of user queries, \text{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, \text{Management} – Web content management operator, \text{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 by the operator \text{Formation} as follows

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

where $u_f$ – conditions set of commercial Web content formation, i.e. $u_f = \{u_1(x_1),...,u_n(x_1)\}$.

Commercial Web content is presented as follows:

$$c_r = \left\{ \bigcup_{f} u_f \left(x_i \in X \land \exists u_f \in U \right), \quad U = U_x \lor U_{-x} \quad i = 1, m, f = 1, n \right\} \quad (6)$$

Stage of commercial Web content managing is described by the operator \text{Management} as follows

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

where $Q$ – set of information users requests, $H$ – set of commercial Web content management conditions, i.e. $H = \{h_k(c_{i+1}, q_d),...,h_{n_k}(c_{j+1}, q_d)\}$.

Commercial Web content management is presented as

$$z_w = \bigcup_{l=1}^{n_k} h_l(c_{i+1}, q_d) \land (q_d \in Q) \land (h_k \in H_d), \quad H = H_{q_d} \lor H_{-q_d}, \quad i = 1, n_c, \quad k = 1, n_H \quad (8)$$

Stage of commercial Web content support is described by the operator \text{Support} as follows

$$y(t_p + \Delta t) = \text{Support}(v_1, h_k, c_r, z_w, t_p, \Delta t) \quad (9)$$

where $v_1$ – conditions set of commercial Web content support and the environment impact on the system, i.e.

$$v_1 = (v_1(q_1, h_k, c_r, z_w, t_p),...v_{n_p}(q_1, h_k, c_r, z_w, t_p))$$

Output data are implemented

$$y_j = \bigcup_{l=1}^{V_1} \left\{ \exists q_d \in Q \land \exists z_w \in Z \land \big(V \land V_{-q_d} \lor V_{-q_d} \land d = 1, n_q, l = 1, n_v, \right\} \quad (10)$$

A complete functional electronic business system is characterized a complex system of interrelated operations, methods and techniques of commercial Web content processing. This is shown in Fig. 1.
6 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, i.e.: 

\[ \text{Source}(x_i) \rightarrow x_i \rightarrow X \rightarrow \text{Formation}(u_f, x_i, t_p) \rightarrow c_r \rightarrow C \rightarrow \text{DataBase}(C), \]

where:
- \( \text{Source}(x_i) \) – the \( x_i \) content source,
- \( x_i \) – the \( i \)-th content from data source,
- \( X \) – content set from corresponding sources of data,
- \( \text{Formation}(u_f, x_i, t_p) \) – the operator of the commercial content formation at \( x_i \) at 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,
- \( \text{DataBase}(C) \) – database of commercial Web content.

Model of commercial Web content formation is presented as

\[ \text{Formation} = \{X, \text{Gathering}, \text{Formatting}, \text{KeyWords}, \text{Backup}, \text{Caterization}, \text{BuDigest}, \text{Dissemination}, T, C\} \] (11)

where:
- \( X \) – set of input data \( x_i \in X \) from different information resources or moderators;
- \( \text{Gathering} \) – operator of content collecting/creating from various sources;
- \( \text{Formatting} \) – content formatting operator;
- \( \text{KeyWords} \) – the operator of content keywords and concepts identification;
- \( \text{Caterization} \) – the operator of automatic categorization of content;
- \( \text{Backup} \) – the operator of content duplication identification;
Content duplication identification is described by the operator
\[ C = Backup(\text{Gathering}(X, U_G, T), U_B) \] (12)
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 searching, 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, and/or editors), direct scanning of commercial content and his addiction to a common XML format:

\[ C = \text{Formatting}(Backup(\text{Gathering}(X, U_G, T), U_B), U_{FR}) \] (13)

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 is based on the principles of the keywords finding in content (terms). This is based on Zipf’s law and reduced to the words choice with 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:

\[ C = \text{KeyWords}(\text{Formatting}(Backup(\text{Gathering}(X, U_G, T), U_B), U_{FR}), U_K) \] (14)

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 for categorization operator by the operator

\[ C_{Ci} = \text{Categorization}(\text{KeyWords}(C, U_K), U_{Ci}) \] (15)

for compliance with thematic requests, where:

\[ C_D = \text{BuDigest}(\text{Categorization}(\text{KeyWords}(C, U_K), U_{Ci}), U_D) \] (16)
Relevant content is send to users and are loaded into the database. Selective distribution of commercial content is described by operator

$$C_{Dn} = Dissemination(C_D, U_{Dn})$$  \hspace{1cm} (17)

7 METHOD OF COMMERCIAL WEB CONTENT MANAGEMENT

The process of commercial Web content management is presented by this general scheme of links:

$$User(q_d) \rightarrow q_d \rightarrow Q \rightarrow H(c_{i+1},q_d) \rightarrow Management(q_d,c_r,h_k,t_p) \rightarrow z_w \rightarrow User(z_w)$$

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

$$Management_Q = \{H,C,Q,\text{Presentation}, Edit, Weight, T, Z\}$$  \hspace{1cm} (18)

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 $c_j(t_{r+1} + \Delta t) = Edit(c_j(t_r), t_{r+1}, \Delta t)$ Stage of pages set forming is described by 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, i.e.

$$z_i = \left\{ \bigcup c_j(q_i,t_r) \big| (\forall c_j \in C_q) \land (\exists q_i \in C_q) \land \land Weight(c_j), C = C_q \lor C_q, j = 1, m, i = 1, n \right\}$$  \hspace{1cm} (19)

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

$$Management_E = \{C, H, Edit, Weight, T, Z\}$$  \hspace{1cm} (20)

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

$$Z(t) = Edit(C, Weight, H, t)$$  \hspace{1cm} (21)

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

$$Management_M = \{C, Q, H, \text{Presentation}, Edit, Weight, Caching, T, Z\}$$  \hspace{1cm} (22)

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

$$Cache = Caching(C, Weight, t, \Delta t)$$  \hspace{1cm} (23)

or

$$Cache = Caching(Z, Weight, t, \Delta t)$$  \hspace{1cm} (24)

i.e.

$$Cache = \left\{ \bigcup c_i \big| (\forall c_i \in C) \land \land Weight(c_i), t = t + \Delta t, i = 1, n \right\}$$  \hspace{1cm} (25)
\[
\text{Cache} = \bigcup_i \left\{ \sum_j \left( \exists c_j \in Z \wedge (\forall c_j \in C) \wedge \left( f = t + \Delta t, i = \frac{1}{n} \right) \right) \right\}
\]  
(26)

8 METHOD OF COMMERCIAL WEB CONTENT SUPPORTING

Process of commercial Web content support is presented as follows scheme links

\[
\text{User}(q_d, z_w) \rightarrow q_d \rightarrow z_w \rightarrow V(q_d, z_w, t_p) \rightarrow \text{Support}(v_i, h_c, r_c, z_w, t_p, \Delta t) \rightarrow y_j \rightarrow \text{Moderator}(y_j).
\]

The method of commercial Web content support looks like

\[
\text{Support} = \{ Q, C, H, V, T, \text{BuInfPort}, \text{IdThemTop}, \text{ConCorrTabConc}, \text{CalRankConc}, Z, Y \}
\]  
(27)

where:
- \( Q \) – set of user queries,
- \( C \) – the Web content set,
- \( \text{BuInfPort} \) – the operator of the information portraits formation,
- \( \text{IdThemTop} \) – the operator of thematic storyline identification,
- \( \text{ConCorrTabConc} \) – operator content relationship tables building,
- \( \text{CalRankConc} \) – 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.

An analysis result of the \( S \) e-business systems \( S \) functioning and \( C \) commercial Web content support has formed set \( Y = \{ Y_p, Y_T, Y_C, Y_R \} \) under the conditions \( V = \{ V_p, V_T, V_C, V_R \} \),

where:
- \( Y_p = Y_{p_C} \lor Y_{p_R} \) – a subset of the information portraits of \( Y_{p_C} \) content and \( Y_{p_R} \) 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,
- \( V_p = V_{p_C} \lor V_{p_R} \) – 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, presented as \( Y_{p_C} = \text{BuInfPort}(V_{p_C}, C, H, Q, T) \), and \( Y_{p_R} \) set of the users’ portraits are given as:

\[
Y_{p_R} = \text{BuInfPort}(V_{p_R}, Q, H, Z, T)
\]  
(28)

where:
- \( V_p = V_{p_C} \lor V_{p_R} \) – set of the portraits formation conditions,
- \( \text{BuInfPort} \) – the operator of the portraits formation as \( Y_p = Y_{p_C} \lor Y_{p_R} \).

The thematic storyline set for the \( Y_T \) content is presented as

\[
Y_T = \text{IdThemTop}(C, H, Q, V_T, T)
\]  
(29)

where:
- \( V_T \) – conditions set for content storyline identification,
- \( \text{IdThemTop} \) – the operator of \( Y_T \) thematic storyline definition.

The \( Y_C \) set of relationship content tables is presented as \( Y_C = \text{ConCorrTabConc}(C, V_C, T) \),

where:
- \( V_C \) – the conditions set of the relationship tables building,
- \( \text{ConCorrTabConc} \) – tables constructing operator.

The \( Y_{Re} \) set of the content ratings is presented as:

\[
Y_{Re} = \text{CalRankConc}(C, Q, H, Y_C, Y_{Re}, \text{Spam}, \text{Tonality}, T)
\]  
(30)

and a \( Y_{Rm} \) set of a moderators ratings is presented as

\[
Y_{Rm} = \text{CalRankConc}(C, Q, H, Y_C, Y_{Rm}, T)
\]  
(31)
where:

- \( V_R = V_{Rc} \vee V_{Rm} \) – the parameters set for the content ratings calculation,
- \( T_{onality}(Q^+, Q^0, Q^-, T, H) \) – the tonality criteria for commercial content,
- \( Spam(Q, T) \) – operator of comments filtering definition,
- \( CalRankConc \) – operator of content and moderators ratings identification

\[ Y_R = Y_{Rc} \vee Y_{Rm} \]

The \( Y \) set of output statistical data is presented as

\[ Y = \{ Y_P, Y_T, Y_C, Y_R \} = \]
\[ = \text{Support}(V, C, Q, H, Z, T, \Delta T) \]
\[ \quad \text{(32)} \]

\[ Y = \{ Y_P, Y_T, Y_C, Y_R \} = \]
\[ = \text{Support}(V_P, V_T, V_C, V_R, C, Q, H, Z, T, \Delta T) \]
\[ \quad \text{(33)} \]

where:

\( Y_P = Y_{Pc} \vee Y_{Pq} \) – 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,
- \( Y_R = Y_{Rc} \vee Y_{Rm} \) – a subset of the content and moderators ratings,
- \( \text{Support} \) – 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. 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**

<table>
<thead>
<tr>
<th>№</th>
<th>Information Resource Name</th>
<th>System Appointment</th>
<th>Address</th>
<th>Subsystem availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fotoghalereja-vysoczkykh</td>
<td>Online shop</td>
<td>fotoghalereja-vysoczkykh.com</td>
<td>+/- + +/-</td>
</tr>
<tr>
<td>2</td>
<td>Vgolos</td>
<td>Online Newspaper</td>
<td>vgos.com.ua</td>
<td>+ + +</td>
</tr>
<tr>
<td>3</td>
<td>Tatjana</td>
<td>Online Newspaper</td>
<td>tatjana.in.ua</td>
<td>- +/- +/-</td>
</tr>
<tr>
<td>4</td>
<td>Presstime</td>
<td>Online Newspaper</td>
<td>pressetime.com.ua</td>
<td>+/- + +</td>
</tr>
<tr>
<td>5</td>
<td>AutoChip</td>
<td>Online shop</td>
<td>autochip.vn.ua</td>
<td>- + +/-</td>
</tr>
<tr>
<td>6</td>
<td>Kursyvalyut</td>
<td>Online Journal</td>
<td>kursyvalyut.com</td>
<td>+ + -</td>
</tr>
<tr>
<td>7</td>
<td>Good morning, accountant!</td>
<td>Online Newspaper</td>
<td>dobryjranok.com</td>
<td>+ + +/-</td>
</tr>
<tr>
<td>8</td>
<td>Victana</td>
<td>Online Journal</td>
<td>victana.liv.ua</td>
<td>+/- +/- +/-</td>
</tr>
<tr>
<td>9</td>
<td>Information for Businesses</td>
<td>Online Journal</td>
<td>goodmorningua.com</td>
<td>- - -</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>The developed e-business systems</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Visiting</td>
<td>4144</td>
</tr>
<tr>
<td>Visits time</td>
<td>4:50</td>
</tr>
<tr>
<td>Failures/refusals</td>
<td>55.62</td>
</tr>
<tr>
<td>Goal Conversion</td>
<td>4.35</td>
</tr>
<tr>
<td>Unique visitors</td>
<td>2674</td>
</tr>
<tr>
<td>Repeat visits</td>
<td>1470</td>
</tr>
<tr>
<td>Page-views</td>
<td>19082</td>
</tr>
<tr>
<td>Pages/visit</td>
<td>4.60</td>
</tr>
<tr>
<td>% New Visits</td>
<td>64.48</td>
</tr>
<tr>
<td>New Visitor</td>
<td>64.53</td>
</tr>
<tr>
<td>Returning Visit</td>
<td>35.47</td>
</tr>
<tr>
<td>Search traffic</td>
<td>56.47</td>
</tr>
<tr>
<td>Direct Traffic</td>
<td>19.50</td>
</tr>
<tr>
<td>Transitions (%)</td>
<td>24.03</td>
</tr>
<tr>
<td>Language – RU</td>
<td>45.15</td>
</tr>
<tr>
<td>Language – EN</td>
<td>36.20</td>
</tr>
<tr>
<td>Ukraine</td>
<td>81.25</td>
</tr>
<tr>
<td>Russia</td>
<td>9.10</td>
</tr>
<tr>
<td>Poland</td>
<td>0.19</td>
</tr>
<tr>
<td>United States</td>
<td>1.21</td>
</tr>
<tr>
<td>Germany</td>
<td>0.34</td>
</tr>
<tr>
<td>City – Kyiv</td>
<td>17.25</td>
</tr>
<tr>
<td>City – L’viv</td>
<td>38.92</td>
</tr>
<tr>
<td>Dnipropetrovsk</td>
<td>2.82</td>
</tr>
<tr>
<td>City – Kharkiv</td>
<td>1.50</td>
</tr>
<tr>
<td>City – Ternopil</td>
<td>2.24</td>
</tr>
</tbody>
</table>

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. 2-3 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. 2a. Distribution of time of arrival at the site

Fig. 2b. Distribution of traffic

Fig. 3a. Distribution of new visits and failures
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. 4).

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 which provides a content database creation 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 categories. And he is credited descriptors with keywords. This facilitates the process implementation of content management.

Tasks 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 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.

9 CONCLUSIONS

In this paper 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 is solved. It is substantiated by 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. E-business system model is developed, which helped to develop the commercial content lifecycle, the generalized system architecture and standardized methods for content processing in these systems. Overall system architecture for e-business systems is improved, which is different from the existing architectures of availability of commercial content processing modules. This enabled the stage implementation of the content lifecycle.

In this paper 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) is developed. The general guidelines for the architecture designing of e-business systems is developed, which different from the existing systems more detailed steps and the presence of commercial Web content processing modules. They allow effectively implement an 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 the modules architecture of e-business systems for the stages implementation of the commercial Web content lifecycle is developed. 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 is developed and implemented.

WORKS CITED


<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Year</th>
<th>Pages</th>
<th>Publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stohasticheskie fraktaly v informatike. Nauchno-tehnicheskaya informatsiya,</td>
<td>Ivanov, S.</td>
<td>2002</td>
<td>7-18</td>
<td></td>
</tr>
<tr>
<td>Nekotorye metody analiza novostnyh informatsionnyh potokov (pp. 277-287). Donetsk, Ukraine: IKVT, Donetsk National Technical University Publ., № 93.</td>
<td>Lande, D.</td>
<td>2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nekotorye metody analiza novostnyh informatsionnyh potokov (pp.</td>
<td>Lande, D.</td>
<td>2006</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Analytical methods for commercial web content processing


Received for publication: 14.04.2014
Revision received: 17.06.2014
Accepted for publication: 03.07.2014

How to cite this article?

Style – APA Sixth Edition

Style – Chicago Fifteenth Edition:

Style – GOST Name Sort:

Style – Harvard Anglia:

Style – ISO 690 Numerical Reference: