The Content Analysis Method for the Information Resources Formation in Electronic Content Commerce Systems

Andrii Berko^{*a*}, Vasyl Andrunyk^{*a*}, Lyubomyr Chyrun^{*b*}, Mykhailo Sorokovskyy^{*a*}, Oksana Oborska^{*a*}, Oksana Oryshchyn^{*a*}, Mykhailo Luchkevych^{*c*} and Oksana Brodovska^{*d*}

^a Lviv Polytechnic National University, S. Bandera Street, 12, Lviv, 79013, Ukraine

^b Ivan Franko National University of Lviv, University Street, 1, Lviv, 79000, Ukraine

^c Drohobych Ivan Franko State Pedagogical University, Ivan Franko Street, 24, Drohobych, 82100, Ukraine

^d Limited Liability Company Financial Company MAGNAT, E. Konovaltsya Street, 36-V, Kyiv, 01133, Ukraine

Abstract

The article on developing the unified method of processing information resources in electronic content commerce systems (ECCS) is devoted. A model and a typical generalized architecture of an ECCS are developed. Strategies for designing and implementing ECCS using an on-line journal and newspaper that reflects the results of theoretical studies are developed. The general approaches for creating these systems, typical architecture and standardized methods for content formation/management/support. The on-line journal and newspaper is a typical example of an ECCS, which refers to managed cybernetic systems and has a certain properties set. Articles content is based on the on-line newspaper/journal, according to which the user searches to obtain the necessary data. However, texts filled with keywords are not always enough to get the required content by the user. Highlighting keywords in each article is a laborious and lengthy process. Using a formalized method of content analysis, the process is fully automated and occurs when the author adds a new article. The method defines similar articles to topics viewed by the user. The using content analysis advantage is to determine the content availability for a specific request, for example, to determine the lack of content on a specific topic and the authors' focus on this issue development.

Keywords 1

Electronic content commerce system, information resource, information technology, intelligent system, content commerce, commercial content, content analysis, processing information resource, text analysis, on-line magazine, natural language, textual content

1. Introduction

The problem of processing information resources in ECCSs is relevant through the active development of research in e-business and the lack of theoretical justification for methods and software tools unification for content management, formation, and support [1-3]. New approaches/ways to solve this problem appear, but there is a mismatch between the well-known technology of processing information resources and the principles of building ECCSs. There are no general approaches for creating these systems, typical architecture and standardized methods for content formation/management/support [2-6].

The relevance of the topic is due to the globalization of e-business;

- Increased demand for commercial content and quick access to it;
- Uneven functioning of business processes following the regions;

ORCID: (A. Berko); 0000-0003-0697-7384 (V. Andrunyk); 0000-0002-9448-1751 (L. Chyrun); 0000-0002-9564-791X (M. Sorokovskyy); 0000-0001-9606-0267 (O. Oborska); 0000-0002-8969-1891 (O. Oryshchyn); 0000-0002021960252X (M. Luchkevych); 0000-0003-4828-528X (O. Brodovska)



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

COLINS-2021: 5th International Conference on Computational Linguistics and Intelligent Systems, April 22–23, 2021, Kharkiv, Ukraine EMAIL: 0000-0001-6756-5661 (A. Berko); vasyl.a.andrunyk@lpnu.ua (V. Andrunyk); Lyubomyr.Chyrun@lnu.edu.ua (L. Chyrun); Mykhailo.V.Sorokovskyy@lpnu.ua (M. Sorokovskyy); oksana.v.oborska@lpnu.ua (O. Oborska); oksana.orushchun@gmail (O. Oryshchyn); luchkevychmm@gmail.com (M. Luchkevych); ksenyabrodovska@gmail.com (O. Brodovska)

CEUR Workshop Proceedings (CEUR-WS.org)

- The need to promptly / regularly / periodically receive the necessary content;
- Saving time/resources in obtaining the essential content;
- Personalization in the provision of services and the integration of ECCSs.

The advantages of introducing these systems are to increase the efficiency of receiving content; shortening the cycle of production and sale of commercial content; lower costs associated with the exchange of content; openness towards users; automatic informing users on-line; creating alternative sales channels [1]. The peculiarity of using ECCSs is [7-9]:

• Direct interaction with the user (reduction of distribution channels and elimination of intermediate links),

- Openness (access for all companies/users and low barrier to enter the market) of ECCS processes,
- Infinity in time (access at any time),
- Globalism (access from anywhere in the world),
- Automatic processing of requests and tracking information about users reduces the cost of operating the e-business and provides SRI additional information interactively [9-13].

2. Related works

Active development of the Internet contributes to the growth of demand for commercial content and new forms for Web services [1-2]. Textual content, prepared according to users' needs and designed to meet them, is a commercial content and the main object of e-business. The ECCSs implementation /development is one of the strategic areas of e-business development. A characteristic feature of such ECCSs is the Web resources processing to increase content sales to regular users actively attract potential users and expand the boundaries of the target audience [14-21]. In particular, the e-business technologies and principles are actively used in the software creation for on-line/off-line content sale in such systems as the Internet store, Cloud storage/computing, analysis, transfer and storage of content (Table 1).

Table 1

Directions for the development of electronic content commerce

Specialization	Specialization characteristic
On-line	Content distribution through on-line newspapers, on-line magazines, distance
selling	learning, on-line publications in the form of dictionaries/directories, on-line
	publishing, entertainment/information/children's portals.
Off-line	Sales of content through systems such as RSS Subscription Extension, Marketing
	Services Shop, copywriting services
On-line store	Content systems such as files, forms, certificates, articles, manuals, digital art,
	picture, movies, music, video, software, eBooks, and more
Systems	Cloud storage systems (Windows Azure Storage, Ubuntuone, iCloud by Apple,
cloud storage	Google Cloud Storage, FilesAnywhere, EMC Atmos, Amazon S3) and cloud
/computing	computing (Micorsoft, Palm Pre, Android, iPhone, Linux, Mac, Windows, Apple,
	Google) to save different types of content or just movies, music, media / images.
	Cloud computing service delivery models include infrastructure such as IaaS (Red
	Hat, Rackspace, VMWare, Microsoft, and Amazon), platform as PaaS (Google
	Apps), and software as SaaS (Google docs, Gmail).
Systems	Site analysis services, content verification/validation, search frequency
analysis	determination, keyword positioning. There are special programs, services and
content	search engines for images to test the uniqueness of the content. Google / Yandex
	/ Rambler search engines allow you to check for plagiarism by typing \leq 100
	characters in quotation marks in the search bar. Web analytics systems are the
	tracking, gathering, and measuring of quantitative/qualitative data on the
	attendance of an information resource, followed by their analysis. The main task

	of Web analytics is to optimize your site and Internet marketing initiatives. Google
	Analytics is a powerful tool for tracking sites of any size
Systems	CMS (Drupal, Promodo, Kasseler, Aurus, LiteDiary, SiMan, VaM Shop,
management	Impresspages_CMS, OpenCms, Santafox, e107, PHP-Nuke, PHP Fusion, Xaraya,
content	TYPO3, Mambo, Plone, osCSS, OpenCart, PrestaShop, Magento, Business Catalyst,
(CMS)	Joomla) is divided into Open source software for Python, Proprietary software,
	software as a service, Microsoft ASP.NET, Java, PHP, Ruby on Rails, CFML /
	ColdFusion Markup Language, File / flat file, Perl, Java packages, other. Content
	management framework (CMF) technologies that allow you to create CMS.
Programs-	BitTorrent file sharing. Allow organizing networks with the same privileges of
customers	participants (at the same time are clients and servers). Share on cross-platform
	intended for Android, Mac OS, <i>Window</i> s or GNU / Linux / UNIX.

The practical factor of the problem of processing information resources in ECCSs is associated with solving the problems [22-27]:

- The fast growth rate of Internet content and the spread of access to it,
- Expanding the recruitment and growth of demand for content/services,
- The active development of e-business,
- Creating technologies / tools and expanding the scope of the methods of formation,
- Content management and support [2]

The leading world manufacturers of information processing tools actively work in this direction, notably Oracle, Open Text, Microsoft Alfresco, IBM, SAP, EMC, CM Professionals organization, AIIM, and Google. The theoretical factor of the problem of processing Web resources in ECCSs involves the development of unified methods for the formation/management/ support of content based on research in the field of e-business and scientific works of leading scientists (Table 2) [28-42].

Table 2

The main theoretical results in the field of electronic content commerce

Method	Characteristic
Analysis of	In the works of D.V. Land, S.M. Braichevsky, A.N. Grigoryeva, V. N. Furasheva
information	investigated and developed mathematical models of electronic information
flows	flows and their estimation [9].
Keyword	George Kingsley Zipf proposed an empirical pattern of the frequency words
Analysis	distribution in a natural language [2, 9].
Capture events	G. Salton and R. Papka pronounced the idea of the emergence of new trends in informational flows [9].
Content	Halverson, Woods Randy, Bob Doyle, Russell Nakano, Ann Rockley, JoAnn
lifecycle	Hackos, Gerry McGovern, Bob Boiko, McKeever Susan described life cycle
A A A	models of content [2].
Content	They started the methodology of D. Robertson, Strauss, Glaser, A. Lindesmith, J.
analysis	Kaiser, D. Wilcox, D. Spiida, A. Tenni, B. Matthews and actively developed
	F. Joubish, E. Babbie, O. Holsti, H. Lasswell, Stemler, J. Naisbitt proposed a
	methodology for exploring texts to determine authorship, authenticity, or
	meaning. K. Krippendorff and K. Neuendorf have developed methods for
	quantitative and qualitative text analysis. D. McKeone found a difference
A A A	between prescriptive analysis and open text analysis [2].
Content	Oracle, Open Text, SAP, Microsoft Alfresco, IBM and EMC have developed
management	Content Management Interoperability Services (CMIS) specifications for Web
	Services interfaces designed to provide interoperability between ECM systems.
	The Interoperable Content Application toolkit interacts with content from

	different repositories through a service interface and a custom CMIS
	Implementation, developed by each CMIS participant independently [2].
Computer	Natural-language texts analysis consists of sequential processes such as
linguistics	semantic, syntactic and morphological analysis. For each analysis, appropriate
	models and algorithms are created for such natural language processing tools
	development as educational didactic, semantic/syntactic/morphological
	analysis, annotation, machine translation, content retrieval systems, etc. [2-6].

The paper presents solutions to this problem in the form of a theoretically sound concept and description of the corresponding unified methods by processing Web resources based on the principles of construction and functional capabilities of the respective systems.

3. Material and methods

A Web resource in ECCSs is a set of data with a set of properties (Table 3), which are the object of the action of the technology for converting them into commercial content. The result of applying one technology is an information resource of another [2].

 Table 3

 Properties of Web resources in ECCSs [43-47]

Name	Property
Heterogeneity	The presence of content of different origin, content, and format of presentation.
Consistency	Lack of conflicting/opposing content values.
Accessibility	Accessibility for all users based on standardized methods /tools/interfaces.
Openness	The ability to interact, exchange values and share with external resources.
Dynamics	Quick update in accordance with the conditions of the system or the external
	environment.
Scalability	The ability to change the logical/physical volume of content (the number of
	quantities/concepts and their notation).
Controllability	Change control/use of content and its impact on processes by elements/users of
	the system.

Content - in the field of information technology (IT) is formalized content and knowledge in the system without a detailed specification of their properties, formalization, and ordering [2]. Converting data that is heterogeneous in nature /content/origin into a coherent centralized information resource is essential for building and functioning ECCSs. Important tasks are ensuring the information needs of problem-oriented elements of the system, maintaining access to content of various categories of users, observing data integrity and consistency rules, minimizing / controlling excess content, the ability to develop / change the internal organization of an information resource, observing the quality and effectiveness of content [1]. It is necessary to ensure the invariance of the environment of ECCSs to modify information resources in such changes [1-2, 48-52]:

- Presentation methods, formats and internal organization of content;
- Storage environments for content, physical units, and hardware;
- User requirements for content;
- The emergence of new requirements and categories of users;
- The order of distribution of content and user access methods.

There is a problem of creating a single conceptual description of an information resource to maintain external/internal designations of content in accordance with their tasks/requirements/ changes. Therefore, we classify information resources for the study of their natural / technological / consumer qualities in order to identify characteristic / specific properties, as well as patterns / features of their formation / application. The classification is based on the basic properties of content (syntax, structure, and semantics), based on which the main classification factors are selected [1-2]: ways of presenting content; resource structuring methods; resource access paths; resource assignment. The general

principles of Web resource formation in ECCSs (Fig. 1) determine the procedure and methods for selecting information from primary sources, filtering, fixation, converting to a specific format for generating the content and placing it in a database. There are plural *N* predefined primary sources of fixed / variable content. Each source of information x_i ($i = \overline{1, N}$) forms a set of values of information/knowledge/facts from the subject area of the ECCS [2, 53-60].

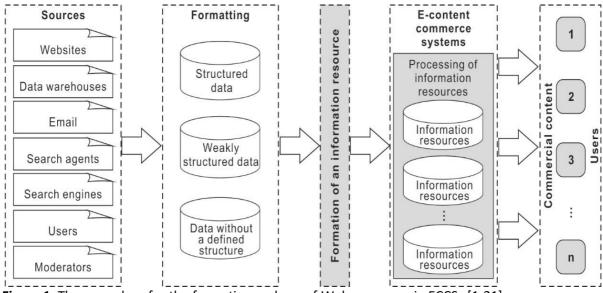


Figure 1: The procedure for the formation and use of Web resources in ECCSs [1-21]

The result of the appeal of technological means to the source x_i is the generation of a variety of values $X(x_i)$, which is perceived, fixed and served in a certain form. The set of values generated by each information source is converted into an input content set of a specific format X_i $(i = \overline{1, N})$. Each set of content is presented in structured, loosely structured data or without a detailed description of the structure [61-69]. Structuring the content involves forming each set of reports its composition, ways of combining elements and their ordering (plural of conditions U_i , where U_i) [70-76]. Each set of content is a combination of many values in a given format and plural of conditions $\langle X_i, U_i \rangle$. In the case of the formation of an input content set without a description of the structure $U_i = \emptyset$. The received content, before formatting and saving, goes through the verification / validation procedure to confirm formal / substantive correctness / relevance according to the requirements of the system. In case of non-compliance with the specified criteria is withdrawn from further use [77-90].

Formally, we represent such a sequence of technological processes as a chain:

$$\begin{aligned} x_{i} \to X(x_{i}) \to X_{i} \to \langle X_{i}, U_{i} \rangle \to Verification(\langle X_{i}, U_{i} \rangle) \to Qualification(\langle X_{i}, U_{i} \rangle) \to \\ \to Conversion(\langle X_{i}, U_{i} \rangle) \to Downloading(\langle X_{i}, U_{i} \rangle) \to \langle X, U \rangle, \quad npu \ i = \overline{1, N}, \end{aligned}$$
(1)

where N is the number of sources,

 x_i is *i* the source, $X(x_i)$ is the set of generated values,

 X_i is the set of selected content,

 $\langle X_i, U_i \rangle$ is the set of content with many conditions,

 $\langle X, U \rangle$ is system resource,

Downloading $(\langle X_i, U_i \rangle)$ is content loading,

 $Conversion(\langle X_i, U_i \rangle)$ is content conversion,

Qualification $(\langle X_i, U_i \rangle)$ is qualification,

$Verification(\langle X_i, U_i \rangle)$ is verification

The Table 4 presents the architectural principles of building information resources in ECCSs, [1-2, 5-6, 9-10]. The formation of the information resource is carried out by the method of homogenization/distribution/ integration (Table 5) [1-27].

Table 4

Architectural principles for building information resources in e-commerce systems

Principle	Information Resource Property
System	Building a resource requires consideration and coordination with the
	requirements/functions of other components of the system.
Complexity	In its composition, presentation formats, and content, the resource is
	heterogeneous, therefore, when it is formed, they provide a combination of
	heterogeneous elements into a single coherent set.
Fullness	The composition of the information resource ensures the fulfillment of all tasks
	and functions of the system.
Integrity	Compliance with the structure (provides an isomorphism of the states of data
	units and the relationships between them when they are updated), semantics
	(forms a relevant interpretation of each unit / aggregate of data in all states of
	the resource) and functionality (ensures the performance of all functions of the
	system when changing the states of its resource).
Openness	The ability of a resource to interact and exchange content with resources of
	other systems/sources.
Interoperability	The content, methods of perception, application, interpretation, and
	functioning of the resource do not depend on platforms and implementation
	technologies and are unchanged when the system functioning environment
	changes.
Continuity	The new state of an information resource when updating directly depends on
	the previous one (it is a function of it and a set of operations on it).

Table 5

Methods of forming Web resources in ECCS

Title	Characteristic	
Homogenization	 The formation of a single homogeneous set of content, namely: obtaining knowledge and information from various previously defined sources in an arbitrary informal form; formalization / standardization of received from sources and presentation in the form of thematic content; verification of content (verification of compliance with syntax rules, content, completeness, lack of repetition, etc.); content downloads (manual / automated placement of content in a database with additional conversion from an intermediate format to a specific one). 	
Distribution	Autonomous formation and use of local components (sets of the content of certain content, structure, and purpose with a set of technologies for creating /maintaining/managing/ access) of a single global content set.	
Integration	An integrated resource forms a consistent set of values/information /knowledge/facts suitable for sharing. All components of an integrated resource are available for use within a single set of methods, tools and technologies. The effect of using a global integrated information resource exceeds the total effect of using all local resources (components of the integration process).	

Performing the sequence of steps of forming the resource by the method of homogenization, a lot of content as a result is formed/replenished. The objective of a set of distributed resources management tools is to organize the coordinated work and interaction of local technological tools to ensure the sharing of all / part of local resources. The first two methods have a number of disadvantages (Table 6) [1-27]. The main factors in the development of the integration method are the increase in the share of poorly structured data and without a preliminary description of the structure with an increase in the volume of resources.

Table 6

Comparison of the methods of generating information resources in e-commerce system	vstems
--	--------

Title	Advantages	Limitations
Homogenization	They are used in the following cases: there is only one way of presenting content and a set of tools/technologies for its support; sources are predefined, identified and specified; the number of sources is small; information from sources is suitable for converting to uniform, unified formats.	Resource formation requirements are stringent; their implementation is problematic and often impossible.
Distribution	The main area of application of distributed resources is the Internet, although the approach is often used in the construction of large-scale corporate systems (distributed databases).	The share of distributed resources of a weakly structured type or without a predefined structure is growing. Therefore, special systems are used.
Integration	 Allows you to: Combine content of different formats, contents and origin in a single agreed set; Combine content without generalized formatting Build virtual custom images of content that are not dependent on their actual appearance; Operate with various content in its combination; Dynamically supplement, modify and transform content and its descriptions; Provide unified technologies for the perception and application of a large number of different content. 	No limitations.

A feature of the integration method is a homogeneous/heterogeneous result that requires physical movement of the data or uses their virtual images, a single global description of the data, or many local descriptions. Thus, it is possible to form an information resource as a single coherent consistent set of data suitable for use in solving a wide range of diverse tasks/problems in various directions of development of the electronic content commerce industry (Table 7) [1-27].

Table 7

A set of urgent tasks in various areas of development of the electronic content commerce industry

Title	Explanation
Internet as a	The greatest success is achieved and actively developed projects that use
platform	the Internet as a platform for e-commerce [6, 12-19].

Provision of services	It is the provision of the services on the Internet and not the sale of software with the installation on the user's computer [6]. The highest user activity is concentrated here, and the purchase or free download of local software is on the periphery. The rapid success of Google (provides services) compared to Netscape (sells software) indicates a correctly chosen strategy on the eve of the dominance of Web 2.0 [6, 12-19].	
Multilingualism	When serving users from all over the world, the need for prompt translation of the interface and content/site content is growing. Popular/common Internet services support multilingualism [12-19].	
Independence from the user	Removes the need for regular official releases / updates [16]. Free open source software is being developed for users to independently create sites / portals according to the requirements of Web 2.0 [6, 12-19].	
Linguistic components of the	Active development of applications for editing the interface in real time by editors / users [6]. There is a growing demand for electronic dictionaries,	
system	systems for automatic spelling / grammar checking, automatic translation with mounting in the admin / editorial part of systems [2-9].	
Non-rigid	Grouping / structuring content according to semantics / content and key	
categorization Content search	facilitate the implementation of content search algorithms [2-3, 6-9]. New approaches are introduced to present search results for users who rely	
Content search	more on semantics. Various linguistic filters are created to filter out random pages and spam, search by synonyms, check for viruses, etc. [2, 6-9, 11].	
Use of databases	Access to unique databases (terminological / spelling, paradigmatic, for automatic translation, corpus of texts, etc.) for payment [6, 16].	
Cooperation	RSS technologies allow you to receive data from a site without visiting it,	
between network	display this data on other sites or locally in the program [1-3]. Mutual	
projects	integration of projects requires applied linguistic programs, in particular, semantic analysis (for example, in the field of automatic or automated	
	selection and grouping of news from various Internet resources) [1-9].	
User as an active	Popularity and / or commercial success directly depends on the number of	
participant in the	users involved in the project, for example, as the author of texts (Wikipedia,	
project	LiveJournal) or media (YouTube, Flickr, Photo.net), comments / product	
	reviews (Amazon), seller / buyer (eBay), contributors to the project	
	database in a specific area of knowledge (Open Library), content owner (BitTorront, Wikipadia), atc. [1,2,5,6]. Attention to the user contributes to	
	(BitTorrent, Wikipedia), etc. [1-3, 5-6]. Attention to the user contributes to the weight gain of linguistic content on the site. User participation in the	
	project provides for the provision of certain linguistic tools (programs for	
	editing text with automatic spelling, terminology and translation	
	dictionaries of the corresponding knowledge industry, etc.) [13].	
User interest	They use	
	 various psychological / psycholinguistic techniques, 	
	• the maximum simplification of the interface and convenience [15],	
	the lack of advertising,	
	 acceleration of page loading speed (AJAX), technologies for converting a site into an analogue of a negular. 	
	 technologies for converting a site into an analogue of a regular program with a large set of functions and ease of use, for example, 	
	mail systems (Gmail),	
	Web-editors of texts / images,	
	electronic dictionaries and programs for automatic translation of	
	text, office programs, etc. [1, 2, 6].	

Cooperation with users	They build a specific hierarchy where users with trust are moderators / administrators of the project. They use the achievements of psycholinguistics and other disciplines related to linguistics [2, 6].
Electronic libraries	Google Books and the Open Library contain full-text corrugation books [6], which use automatic / automated text processing / recognition systems, bibliographic and cataloging systems.
Distance Learning	Applied linguistic programs related to linguistics, for example, distance learning of natural languages (Mova.info) [1-3, 6].
Blogs	Written communicative aspect (users comment on notes, know, communicate, etc.) [6], which led to a geometric growth of various text on the Internet and the need for new principles for its automatic processing by search engines and the development of text editing programs [2].
Social networks	The community is closed to outsiders and protected (Facebook, Vkontakte, Odnoklassniki) by professional or other interests [6, 19]. There are projects where a user's virtual workstation is created, which allows him to keep the sites he needs in one place without the need to visit them [2].
Computational linguistics	The communicative aspect of the Internet is growing, which is directly related to the language; therefore, the weight of research and development on computer linguistics in the functioning of ECCSs is also growing [6, 11].

Commercial content is a purchase / sale object between e-commerce participants [2], for example:

- Announcements of materials from other sections of the site or other sites (with a link),
- Information blocks that are divided into syndicates (for example, a weather block),
- Reference information (for example, holiday dates, event announcement, train schedule),
- Entertainment information (for example, a joke of the day),
- Advertising, buttons and links of information partners, statistics buttons.

The formal model of the commercial content life cycle is six elements [1-27, 77-90]:

 $S = \langle X, Creation, C, Processing, Distribution, Y \rangle$,

(1)

where $C = \{c_1, c_2, \dots, c_{n_c}\}$ is a lot of content,

 $X = \left\{ x_1, x_2, \dots, x_{n_x} \right\} \text{ is a lot of input information,}$

Creation is a function for creating content,

Processing is a function for processing content,

Distribution is a function for distributing content;

 $Y = \{y_1, y_2, \dots, y_{n_v}\}$ is a lot of source information (Fig. 2).

→ Content Creation → Content Processing → Content Distribution → Figure 2: Model of a content distribution system in ECCSs

One type of ECCS is the Internet magazine [5], that is, a lot of structured content (for example, electronic publications in the form of articles, announcements, digests, books, reports, blogs, comments, etc.) in the information resource, is designed to meet the needs of the target audience and perception using appropriate software over the Internet (Fig. 3) [1-27]. Advantages of implementing an on-line magazine: compact; cost-effective; instantly delivers any content anywhere in the world and at any time; not tied to the problem of replication; reduced to a minimum the distance between the author and the reader; new ways of conveying information to the reader. Automatic processing of electronic publications reduces the search time for the necessary content only by entering the appropriate address or keywords in the search system [1-27].

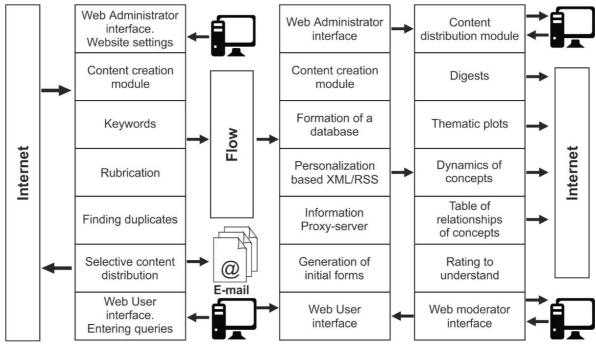


Figure 3: The interaction scheme of the modules of ECCSs developed according to [9]

4. Experiments, results and discussion

The process of creating content is described by a function of the form $\vec{c}(x_i, t) = Creation(\overline{u_c}, x_i, t)$, where $\vec{u_c}(x_i)$ is the set of conditions for creating content, that is

$$\overline{u_{C}}(x_{i}) = \left(u_{c_{1}}(x_{i}), u_{c_{2}}(x_{i}), \dots, u_{c_{n_{U_{c}}}}(x_{i})\right).$$
Content is created as $c_{j} = \left\{\bigcup u_{c_{k}} \mid (x_{i} \in X) \land (\exists u_{c_{k}} \in U_{F}), U_{C} = U_{C_{X}} \lor U_{C_{X}}, i = \overline{1, m}, k = \overline{1, n}\right\}, \text{ i.e.}$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(3)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1)$$

$$(1$$

Content processing is described as

$$\vec{c}(q_i,t) = Processing(\overrightarrow{u_P},q_i,t),$$

where $Q = \{q_1, q_2, \dots, q_{n_Q}\}$ is the set of user requests [2],

 $\overrightarrow{u_P}(q_i)$ is the conditions set for content processing, that is

$$\overrightarrow{u_P}(q_i) = \left(u_{p_1}(q_i), u_{p_2}(q_i), \dots, u_{f_{n_{U_P}}}(q_i)\right).$$

Content processing occurs as $c_j = \left\{ \bigcup u_{p_k} | (q_i \in Q) \land (\exists u_{p_k} \in U_P), U_P = U_{Mq} \lor U_{P\bar{q}}, i = \overline{1, m}, k = \overline{1, n} \right\},\$

that is
$$q(1): \frac{1}{q_{1}} + \frac{1}{q_{2}} + \frac{1}{q_{3}} + \frac{1}{q_{4}} +$$

The content distribution process is described as

$$\vec{y}(t + \Delta t) = Distribution(\vec{u_D}, \vec{c}, q_i, t, \Delta t),$$

where $\overrightarrow{u_D}(q_i, \vec{c})$ is the set of conditions for the distribution of content

$$\overrightarrow{u_D}\left(q_i, \vec{c}\right) = \left(u_{d_1}\left(q_i, \vec{c}\right), u_{d_2}\left(q_i, \vec{c}\right), \dots, u_{d_{n_{Ud}}}\left(q_i, \vec{c}\right)\right)$$

when

$$y_{j} = \left\{ \bigcup u_{d_{k}} \left| (q_{i} \in Q) \land (\vec{c} \in C) \land (\exists u_{r_{k}} \in U_{R}), U_{D} = U_{Dc} \lor U_{D\overline{c}}, i = \overline{1, m}, k = \overline{1, n} \right\}$$
(5)

A well-known method of text processing is content analysis, that is, the study of the contents of text arrays / products of communicative correspondence (for example, comments, forums, electronic correspondence, articles, etc.). Content analysis consists in searching for information [7-8, 11] by

semantic units (phrase, topic, idea, author, character, social situation, and the text part, clustered by the content of the category analysis) and interpretation of the result (Table 8).

Table 8

Basic procedures for a formalized method of content analysis

Title	Features of the formalized method of content analysis
Identification	Depending on the content / goals / objectives / hypotheses, studies form many
of units	substantial units.
Allocation of	The units of the account may coincide (counting the frequency of mentioning
units	the selected semantic unit) or not coincide (the researcher forms based on the
	analyzed material) with the units of analysis.
Counting	Classifications by distinguished groupings using special formulas (for example,
procedure	estimating the specific gravity of meaningful categories in a text), statistical
	calculations of text comprehensibility / attractiveness.
Protocol	Classifier general table with categories / units of analysis, which extremely
development	clearly captures the units of expression of each category. For example, the
	categories of analysis in the questionnaire play the question and the units of
	analysis the answers. Content analysis contains information about the content
	(author, publication time, volume, etc.); the study results (the amount of use in
	specific units of analysis and conclusions by analysis category). The protocol is
	filled in encoded form for compact presentation of content and quick
	comparison of different content analysis results based on counting data of all
	registration cards (coding matrices, which indicate the number of units of
	account and the units of analysis are characterized).
Table	The type of table is determined by the research stage, for example, in the form
development	of a system of coordinated and subordinated analysis categories, which looks
	like a questionnaire: each category (questions) provides a number of features
	(answers) by which the text content is quantified.
Matrix	If the sample size \geq 100 units, then analyze the set of matrix sheets. Otherwise, a
development	two-dimensional / multidimensional analysis is carried out, where for each text a
	coding matrix is built

Analysis units are considered against the background of wider linguistic/informative structures, indicating the nature of the segmentation of the text, within which the presence/absence of contextual units is identified. For example, for an analysis unit, the word contextual unit is a sentence. The unit of account is a quantitative measure of the analysis unit, allowing you to register the frequency (regularity) of the appearance of a characteristic of the analysis category for the content (the number of specific words or their combinations, lines, printed characters, pages, paragraphs, copyright sheets, text area, etc.). The sample parameters are determined by the objectives and scope of the study. Content is the backbone of the on-line magazine, by which the user searches to obtain the necessary information. However, texts filled with keywords are not always enough for the user to get the necessary information. In addition, the selection of keywords for each article is a long-term and time-consuming process. Using a formalized method of content analysis, the process is fully automated and occurs when the author adds a new article. Identify similar articles to those viewed by the user, and display them not. When using keywords, the user gets not quite similar, or completely different from the one that interests him (to increase the popularity of the article, authors add keywords on various topics). The advantage of using content analysis is to determine the availability of content for a specific request, for example, to determine the lack of content on a specific topic and the focus of the authors on the development of this issue. The on-line magazine has implemented the distribution of articles. Raw articles are sent to the input, sent to the article-processing unit. According to the rules, in search parameters, keywords and dates of the viewed articles, the articles are divided into categories (similar, popular, selected search or stop looking) and their placement in the database. The administrator controls the process. The developed articles are submitted to the blocks for searching for articles popular (with the most views), stay overlooked (by publication date within a period from a given to the current date), parametric (by user-entered parameters) and similar (by keywords of previously viewed articles by a user) and output result. The elaborated reports fall into the search unit for matches with keywords viewed by the user earlier. The number of matches for further selection, the most similar to those viewed by the author sort the generated list of articles. The first articles are selected, the number of which is set in the ECCS settings. Further, the article author work is currently being revised, moved to the top of the list and displayed for viewing by the user. Fig. 4 shows a class diagram describing the totality of the Internet journal objects (user interface, Web page and Server class group, which includes the DBMS class, Web server and search components for parameters similar to new and popular articles) and their relationship from the point of view of the project manager.

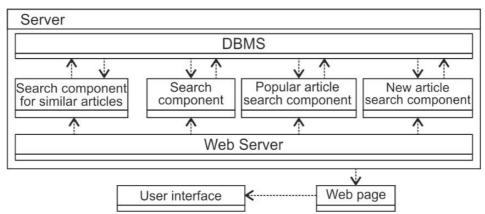


Figure 4: Class diagram of the system of electronic content commerce type Internet magazine

Figure 5 shows a use case diagram and a scenario for systematic execution of the on-line journal tasks. The author's article is worked out and in a database is stored. The user's search parameters fall into the article search block by parameters, which generates a request to the database and the corresponding articles-list to the parameters search. It transfers them to the user and the articles list viewed by the user block. The latter sends the generated list to the block for generating a list of similar articles, which gives the result to the user. Blocks for forming a new and / or popular articles list form queries into the database and display formed lists.

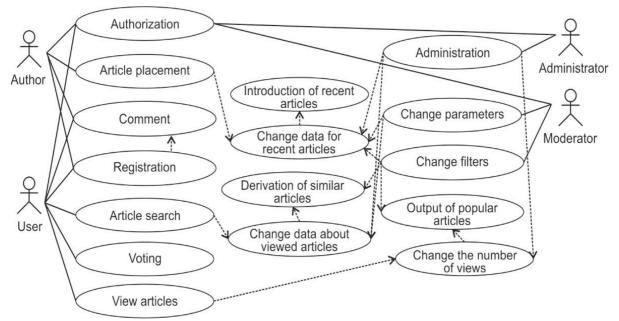


Figure 5: Use case diagram of an ECCS such as an on-line magazine

The actor is registered/authorized as the author in the ECCS, comments/posts articles. Actor A user is registered / authorized in the ECCS, evaluates, views, comments and searches on articles. Viewing articles by the user changes the data on the number of views of the article, including the output of an updated list of popular articles. Also, viewing and searching by parameters leads to a change in the data on viewed articles, including the output of an updated list of similar articles. Actor and Moderator generate/modify article search parameters. The ECCS generates and displays an articles list viewed by the user. Fig. 6a shows a weblog deployment diagram.

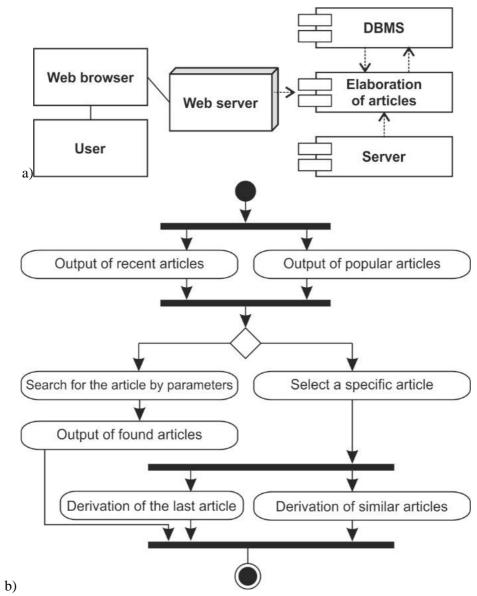


Figure 6: Diagram a) deployment and b) the activity of the ECCS such as Internet journal

After authorization, the user/author sends for processing the article. From time to time, information is exchanged with the database to form a new/popular articles list that meets the parameters search and the like reviewed by the user. All four of these article lists are submitted for viewing to the author/ user. Based on the lists received, the user makes a new request and the mechanism repeats. After performing the operations, the unprocessed author article is transformed into a developed article and stored in the database. It is used to search by keywords, popularity, size, and author. Fig. 6b shows a diagram of the activity of an on-line journal. The user is provided with lists of recent or popular articles that have the most views. The user can select a specific article to view, or search by parameters, leads to the conclusion of the selected / similar / found articles. In fig. 7 is a sequence diagram for an on-line journal and in Fig. 8 is a diagram of system components and dependencies between them.

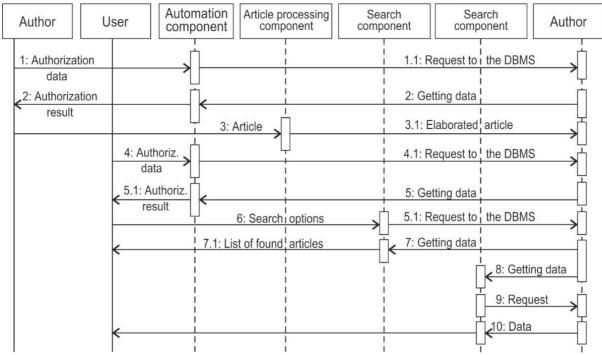


Figure 7: Sequence diagram of an e-commerce system such as an on-line journal

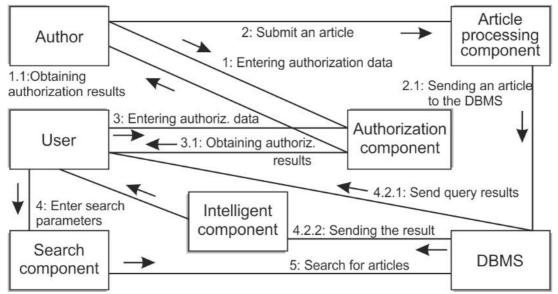


Figure 8: Diagram of components of an ECCS such as an on-line journal

5. Conclusions

The on-line journal and newspaper is a typical example of an ECCS, which refers to managed cybernetic systems and has a certain properties set. Articles content is based on the on-line newspaper/journal, according to which the user searches to obtain the necessary data. Nevertheless, texts filled with keywords are not always enough to get the required content by the user. Highlighting keywords in each article is a laborious and lengthy process. Using a formalized method of content analysis, the process is fully automated and occurs when the author adds a new article. The method defines similar articles to topics viewed by the user. The using content analysis advantage is to determine the content availability for a specific request, for example, to determine the lack of content on a specific topic and the authors' focus on this issue development.

6. References

- V. Lytvyn, V. Vysotska, Designing architecture of electronic content commerce system. In: Computer Science and Information Technologies, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2015, pp. 115-119.
- [2] V. Vysotska, L. Chyrun, L. Chyrun, The Commercial Content Digest Formation and Distributional Process, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2016, pp. 186-189.
- [3] A. Gozhyj, L. Chyrun, A. Kowalska-Styczen, O. Lozynska, Uniform Method of Operative Content Management in Web Systems, volume 2136 of CEUR Workshop Proceedings, 2018, pp. 62-77.
- [4] O. Naum, L. Chyrun, O. Kanishcheva, V. Vysotska, Intellectual System Design for Content Formation, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2017, pp. 131-138.
- [5] O. Bisikalo, V. Vysotska, V. Lytvyn, O. Brodyak, S. Vyshemyrska, Y. Rozov, Experimental Investigation of Significant Keywords Search in Ukrainian Content, volume 1293 of Advances in Intelligent Systems and Computing V, 2021, pp. 3-29.
- [6] O. Pavlenko, I. Tymofieieva, Search Query Data Analysis: Challenges and Opportunities, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 452-461.
- [7] N. Antonyuk, M. Medykovskyy, L. Chyrun, M. Dverii, O. Oborska, M. Krylyshyn, A. Vysotsky, N. Tsiura, O. Naum, Online Tourism System Development for Searching and Planning Trips with User's Requirements, volume 1080 of Advances in Intelligent Systems and Computing, 2020, pp. 831-863.
- [8] T. Basyuk, A. Vasyliuk, V. Lytvyn, Mathematical Model of Semantic Search and Search Optimization, volume Vol-2362 of CEUR Workshop Proceedings, 2019, pp. 96-105.
- [9] L. Chyrun, I. Kis, V. Vysotska, L. Chyrun, Content monitoring method for cut formation of person psychological state in social scoring, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2018, pp. 106-112.
- [10] O. Bisikalo, V. Vysotska, Linguistic analysis method of Ukrainian commercial textual content for data mining, volume Vol-2608 of CEUR Workshop Proceedings, 2020, pp. 224-244.
- [11] Critical Perspectives of Web 2.0. Special issue of First Monday. Vol. 13, #3, 2008. Режим доступу: http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/issue/view/263/ showToc.
- [12] P. Graham, Web 2.0. URL: http://www.paulgraham.com/web20.html.
- [13] R. MacManus, J. Porter, Web 2.0 for Designers. URL: http://www.digitalweb.com/articles/web_2_for_designers.
- [14] D. Hinchcliffe, The State of Web 2.0. URL: http://web2.wsj2.com/the_state_of_web_20.htm.
- [15] T. Scholz, Market Ideology and the Myths of Web 2.0. URL: http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2138/1945.
- [16] T. O'Reilly, What Is Web 2.0. URL: http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/ 30/what-is-web-20.html.
- [17] G. Vickery, S. Wunsch-Vincent, Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking. URL: http://www.oecd.org/ document/40/0,3343,en_2649_201185_ 39428648_1_1_100.html.
- [18] R. Singel, Are You Ready for Web 2.0? URL: http://www.wired.com/news/technology/ 0,1282,69114,00.html.
- [19] O. Veres, B. Rusyn, A. Sachenko, I. Rishnyak, Choosing the Method of Finding Similar Images in the Reverse Search System, volume 2136 of CEUR Workshop Proceedings, 2018, pp. 99-107.
- [20] P. Radiuk, N. Hrypynska, A Framework for Exploring and Modelling Neural Architecture Search Methods, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 1060-1074.
- [21] L. Chyrun, A. Gozhyj, I. Yevseyeva, D. Dosyn, V. Tyhonov, M. Zakharchuk, Web Content Monitoring System Development, volume 2362 of CEUR Workshop Proceedings, 2019, 126-142.
- [22] B. Rusyn, V. Lytvyn, V. Vysotska, M. Emmerich, L. Pohreliuk, The Virtual Library System Design and Development, volume 871 of Advances in Intelligent Systems and Computing, 2019, pp. 328-349.

- [23] O. Kliuiev, N. Vnukova, S. Hlibko, N. Brynza, D. Davydenko, Estimation of the Level of Interest and Modeling of the Topic of Innovation Through Search in Google, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 523-535.
- [24] D. Hand, N. Adams, Data mining, in: Wiley StatsRef: Statistics Reference Online, 2014, pp. 1-7.
- [25] M. W. Berry, J. Kogan, Text mining. Applications and Theory, in: West Sussex, PO19 8SQ, UK: John Wiley & Sons, 2010.
- [26] C. C. Aggarwal, C. Zhai, Mining text data. Springer Science & Business Media, 2012.
- [27] V. Lytvyn, N. Sharonova, T. Hamon, V. Vysotska, N. Grabar, A. Kowalska-Styczen, Computational linguistics and intelligent systems, volume of Vol-2136 CEUR Workshop Proceedings, 2018.
- [28] M. Emmerich, V. Lytvyn, I. Yevseyeva, V. B. Fernandes, D. Dosyn, V. Vysotska, Preface: Modern Machine Learning Technologies and Data Science (MoMLeT&DS-2019), volume Vol-2386 of CEUR Workshop Proceedings, 2019.
- [29] V. Lytvyn, N. Sharonova, T. Hamon, O. Cherednichenko, N. Grabar, A. Kowalska-Styczen, V. Vysotska, Preface: Computational Linguistics and Intelligent Systems (COLINS-2019), volume of Vol-2362 CEUR Workshop Proceedings, 2019.
- [30] M. Emmerich, V. Lytvyn, V. Vysotska, V. Basto-Fernandes, V. Lytvynenko, Preface: Modern Machine Learning Technologies and Data Science (MoMLeT+DS 2020), volume Vol-2631 of CEUR Workshop Proceedings, 2020.
- [31] V. Lytvyn, V. Vysotska, P. Pukach, Z. Nytrebych, I. Demkiv, A. Senyk, O. Malanchuk, S. Sachenko, R. Kovalchuk, N. Huzyk, Analysis of the developed quantitative method for automatic attribution of scientific and technical text content written in Ukrainian, volume 6(2-96) of Eastern-European Journal of Enterprise Technologies, 2018, pp. 19-31.
- [32] V. Lytvyn, V. Vysotska, T. Hamon, N. Grabar, N. Sharonova, O. Cherednichenko, O. Kanishcheva, Preface: Computational Linguistics and Intelligent Systems (COLINS-2020), volume Vol-2604 of CEUR Workshop Proceedings, 2020.
- [33] V. Lytvyn, V. Vysotska, P. Pukach, Z. Nytrebych, I. Demkiv, R. Kovalchuk, N. Huzyk, Development of the linguometric method for automatic identification of the author of text content based on statistical analysis of language diversity coefficients, volume 5(2) of Eastern-European Journal of Enterprise Technologies, 2018, pp. 16-28.
- [34] V. Lytvyn, V. Vysotska, I. Budz, Y. Pelekh, N. Sokulska, R. Kovalchuk, L. Dzyubyk, O. Tereshchuk, M. Komar, Development of the quantitative method for automated text content authorship attribution based on the statistical analysis of N-grams distribution, volume 6(2-102) of Eastern-European Journal of Enterprise Technologies, 2019, pp. 28-51.
- [35] V. Vysotska, O. Kanishcheva, Y. Hlavcheva, Authorship Identification of the Scientific Text in Ukrainian with Using the Lingvometry Methods, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2018, pp. 34-38.
- [36] V. Vysotska, V.B. Fernandes, V. Lytvyn, M. Emmerich, M. Hrendus, Method for Determining Linguometric Coefficient Dynamics of Ukrainian Text Content Authorship, volume 871 of Advances in Intelligent Systems and Computing, 2019, pp. 132-151.
- [37] V. Lytvyn, V. Vysotska, Y. Burov, O. Veres, I. Rishnyak, The Contextual Search Method Based on Domain Thesaurus, volume 689 of Advances in Intelligent Systems and Computing, 2018, pp. 310-319.
- [38] V. Starko, Semantic Annotation for Ukrainian: Categorization Scheme, Principles, and Tools, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 239-248.
- [39] Y. Burov, V. Lytvyn, V. Vysotska, I. Shakleina, The Basic Ontology Development Process Automation Based on Text Resources Analysis, in: 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT, 2020, pp. 280-284.
- [40] O. Kuropiatnyk, V. Shynkarenko, Text Borrowings Detection System for Natural Language Structured Digital Documents, volume 2604 of CEUR workshop proceedings, 2020, pp. 294-305.
- [41] V.-A. Oliinyk, V. Vysotska, Y. Burov, K. Mykich, V. Basto-Fernandes, Propaganda Detection in Text Data Based on NLP and Machine Learning, volume Vol-2631 of CEUR workshop proceedings, 2020, pp. 132-144.

- [42] I. Gruzdo, I. Kyrychenko, G. Tereshchenko, O. Cherednichenko, Application of Paragraphs Vectors Model for Semantic Text Analysis, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 283-293.
- [43] V. Vasyliuk, Y. Shyika, T. Shestakevych, Information System of Psycholinguistic Text Analysis, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 178-188.
- [44] Y. Yusyn, T. Zabolotnia, Methods of Acceleration of Term Correlation Matrix Calculation in the Island Text Clustering Method, volume 2604 of CEUR workshop proceedings, 2020, 140-150.
- [45] O. Artemenko, V. Pasichnyk, N. Kunanets, K. Shunevych, Using sentiment text analysis of user reviews in social media for e-tourism mobile recommender systems, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 259-271.
- [46] N.B. Shakhovska, R.Yu. Noha, Methods and tools for text analysis of publications to study the functioning of scientific schools, volume 47(12) of Journal of Automation and Information Sciences, 2015, pp. 29-43.
- [47] R. Bekesh, L. Chyrun, P. Kravets, A. Demchuk, Y. Matseliukh, T. Batiuk, I. Peleshchak, R. Bigun, I. Maiba, Structural Modeling of Technical Text Analysis and Synthesis Processes, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 562-589.
- [48] N. Shakhovska, O. Basystiuk, K. Shakhovska, Development of the Speech-to-Text Chatbot Interface Based on Google API, volume 2386 of CEUR Workshop Proceedings, 2019, 212-221.
- [49] M. Sazhok, V. Robeiko, R. Seliukh, D. Fedoryn, O. Yukhymenko, Written Form Extraction of Spoken Numeric Sequences in Speech-to-Text Conversion for Ukrainian, volume Vol-2604 of CEUR workshop proceedings, 2020, pp. 442-451.
- [50] B. Rusyn, V. Vysotska, L. Pohreliuk, Model and architecture for virtual library information system, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2018, pp. 37-41.
- [51] T. Batura, A. Bakiyeva, M. Charintseva, A method for automatic text summarization based on rhetorical analysis and topic modeling, volume 19(1) of International Journal of Computing, 2020, pp. 118-127.
- [52] V. Lytvyn, V. Vysotska, B. Rusyn, L. Pohreliuk, P. Berezin, O. Naum, Textual Content Categorizing Technology Development Based on Ontology, volume Vol-2386 of CEUR Workshop Proceedings, 2019, pp. 234-254.
- [53] V. Lytvyn, V. Vysotska, I. Peleshchak, T. Basyuk, V. Kovalchuk, S. Kubinska, L. Chyrun, B. Rusyn, L. Pohreliuk, T. Salo, Identifying Textual Content Based on Thematic Analysis of Similar Texts in Big Data, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2019, pp. 84-91.
- [54] J. Su, V. Vysotska, A. Sachenko, V. Lytvyn, Y. Burov, Information resources processing using linguistic analysis of textual content, in: Proceedings of the Intelligent Data Acquisition and Advanced Computing Systems Technology and Applications, Romania, 2017, pp. 573-578.
- [55] J. Su, A. Sachenko, V. Lytvyn, V. Vysotska, D. Dosyn, Model of Touristic Information Resources Integration According to User Needs, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2018, pp. 113-116.
- [56] V. Vysotska, V.B. Fernandes, M. Emmerich, Web content support method in electronic business systems, volume Vol-2136 of CEUR Workshop Proceedings, 2018, pp. 20-41.
- [57] V. Lytvyn, V. Vysotska, D. Dosyn, O. Lozynska, O. Oborska, Methods of Building Intelligent Decision Support Systems Based on Adaptive Ontology, in: Proceedings of the IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP, 2018, pp. 145-150.
- [58] V. Vysotska, V. Lytvyn, V. Kovalchuk, S. Kubinska, M. Dilai, B. Rusyn, L. Pohreliuk, L. Chyrun, S. Chyrun, O. Brodyak, Method of Similar Textual Content Selection Based on Thematic Information Retrieval, in: Computer Sciences and Information Technologies, 2019, pp. 1-6.
- [59] V. Vysotska, Linguistic Analysis of Textual Commercial Content for Information Resources Processing, in: Proceedings of the Modern Problems of Radio Engineering, Telecommunications and Computer Science, TCSET, 2016, pp. 709-713.
- [60] V. Lytvyn, V. Vysotska, D. Dosyn, Y. Burov, Method for ontology content and structure optimization, provided by a weighted conceptual graph, volume 15(2) of Webology, 2018, 66-85.
- [61] V. Vysotska, Y. Burov, V. Lytvyn, O. Oleshek, Automated Monitoring of Changes in Web Resources, volume 1020 of Advances in Intelligent Systems and Computing, 2020, pp. 348-363.

- [62] Y. Burov, V. Vysotska, P. Kravets, Ontological approach to plot analysis and modeling, volume Vol-2362 of CEUR Workshop Proceedings, 2019, pp. 22-31.
- [63] T. Basyuk, The Popularization Problem of Websites and Analysis of Competitors, volume 689 of Advances in Intelligent Systems and Computing, 2018, Springer, Cham, 54-65.
- [64] V. Andrunyk, L. Chyrun, V. Vysotska, Electronic content commerce system development, in: The Experience of Designing and Application of CAD Systems in Microelectronics, CADSM, 2015.
- [65] T. Basyuk, Popularization of website and without anchor promotion, in: Computer science and information technologies, 2016, pp. 193-195.
- [66] N., Melnykova, O.: Markiv, Semantic approach to personalization of medical data, in: Proceedings of the Computer Sciences and Information Technologies, CSIT, 2016, pp. 59-61.
- [67] E. Vasilevskis, I. Dubyak, T. Basyuk, V. Pasichnyk, A. Rzheuskyi, Mobile application for preliminary diagnosis of diseases, volume 2255 of CEUR Workshop Proceedings, 2018, 275-286.
- [68] T. Basyuk, Innerlinking website pages and weight of links, in: Computer science and information technologies (CSIT), 2017, pp. 12-15.
- [69] P. Zhezhnych, O. Markiv, Linguistic comparison quality evaluation of web-site content with tourism documentation objects, volume 689 of Advances in Intelligent Systems and Computing, 2018, pp. 656-667.
- [70] L. Vasyl, V. Victoria, D. Dmytro, H. Roman, R. Zoriana, Application of Sentence Parsing for Determining Keywords in Ukrainian Texts, in: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT, 2017, pp. 326-331.
- [71] P. Zhezhnych, O. Markiv, A linguistic method of web-site content comparison with tourism documentation objects, in: Computer Sciences and Information Technologies, 2017, pp. 340-343.
- [72] P. Zhezhnych, O. Markiv, Recognition of tourism documentation fragments from web-page posts, in: Proceedings of the 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET, 2018, pp. 948-951.
- [73] J. Chen, D. Dosyn, V. Lytvyn, A. Sachenko, Smart data integration by goal driven ontology learning, volume 529 of Advances in Intelligent Systems and Computing, 2017, pp. 283-292.
- [74] D. Dosyn, V. Lytvyn, V. Kovalevych, O. Oborska, R. Holoshchuk, Knowledge discovery as planning development in knowledgebase framework, in: Proceedings of the Modern Problems of Radio Engineering, Telecommunications and Computer Science, Proceedings of the 13th International Conference on TCSET, 2016, pp. 449-451.
- [75] D.G. Dosyn, O.L. Ivantyshyn, V.V. Koshovyy, I.M. Romanyshyn, S.O. Soroka, To a question on the mechanism of formation of ionospheric disturbances at groundbased artificial acoustic excitation, in: Proceedings of the International Seminar/Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory, DIPED, 2003, pp. 211-214.
- [76] V. Lytvyn, D. Dosyn, M. Emmerich, I. Yevseyeva, Content formation method in the web systems, volume 2136 of CEUR Workshop Proceedings, 2018, pp. 42-61.
- [77] M. Davydov, O. Lozynska, Information system for translation into Ukrainian sign language on mobile devices, in: Computer Sciences and Information Technologies, CSIT, 2017, pp. 48-51.
- [78] O.H. Lypak, V. Lytvyn, O. Lozynska, R. Vovnyanka, Y. Bolyubash, A. Rzheuskyi, D. Dosyn, Formation of Efficient Pipeline Operation Procedures Based on Ontological Approach, volume 871 of Advances in Intelligent Systems and Computing, 2019, pp. 571-581.
- [79] M. Davydov, O. Lozynska, Mathematical method of translation into ukrainian sign language based on ontologies, volume 871 of Advances in Intelligent Systems and Computing, 2018, pp. 89-100.
- [80] V. Lytvyn, The similarity metric of scientific papers summaries on the basis of adaptive ontologies, in: Proceedings of the 7th International Conference on Perspective Technologies and Methods in MEMS Design, MEMSTECH, 2011, p. 162.
- [81] V. Savchuk, O. Lozynska, V. Pasichnyk, Architecture of the Subsystem of the Tourist Profile Formation, volume 871 of Advances in Intelligent Systems and Computing, 2019, pp. 561-570.
- [82] V.V. Lytvyn, O.I. Tsmots, The process of managerial decision making support within the early warning system, volume 149(11) of Actual Problems of Economics, 2013, pp. 222-229.
- [83] P. Kravets, Game methods of construction of adaptive grid areas, in: The Experience of Designing and Application of CAD Systems in Microelectronics, 2003, pp. 513-516.

- [84] P. Kravets, Adaptive method of pursuit game problem solution, in: Proceedings of the Modern Problems of Radio Engineering, Telecommunications and Computer Science Proceedings of International Conference, TCSET, 2006, pp. 62-65.
- [85] P. Kravets, Game methods of the stochastic boundary problem solution, in: Proceedings of the Perspective Technologies and Methods in MEMS Design, MEMSTECH, 2007, pp. 71-74.
- [86] P. Kravets, O. Prodanyuk, Game task of resource allocation, in: Proceedings of the Experience of Designing and Application of CAD Systems in Microelectronics, CADSM, 2009, pp. 437-438.
- [87] P. Kravets, Game method for coalitions formation in multi-agent systems, in: Computer Sciences and Information Technologies, CSIT, 2018, pp. 1-4.
- [88] P. Kravets, R. Kyrkalo, Fuzzy logic controller for embedded systems, in: Proceedings of the International Conference on Perspective Technologies and Methods in MEMS Design, MEMSTECH, 2009, pp. 58-59.
- [89] P. Kravets, The game method for orthonormal systems construction, in: Proceedings of The Experience of Designing and Application of CAD Systems in Microelectronics, 2007, 296-298.
- [90] P. Kravets, The control agent with fuzzy logic, in: Proceedings of the Perspective Technologies and Methods in MEMS Design, MEMSTECH, 2010, pp. 40-41.