# Content Analysis of Text-based Information in E-commerce Systems

Vasyl Lytvyn[1], Victoria Vysotska[2], Lyubomyr Chyrun[3],
Mariya Hrendus[4], Oleh Naum[5]

[1-4]Lviv Polytechnic National University, Lviv, Ukraine
[1]Silesian University of Technologyty, Gliwice, Poland
[5]Drohobych Ivan Franko State Pedagogical University, Drohobych, Ukraine

[1]`Vasyl.V.Lytvyn@lpnu.ua,`
[2]`Victoria.A.Vysotska@lpnu.ua,`
[3]`Lyubomyr.V.Chyrun@lpnu.ua,`
[4]`mariya.h.hirnyak@lpnu.ua,`
[5]`oleh.naum@gmail.com`

**Abstract.** Instrumental means, information technologies and software for constructing such systems have been analyzed and summarized in the article based on the analysis of the basic tasks of electronic content commerce system (ECCS). ECCS functional diagram with information resources processing subsystems has been developed. The overall architecture, objectives and principles of ECCS realization were described in details. The functional elements of the system were described according to GOST 24.204.80, GOST 24.201-79, GOST 19.201-78, GOST 34.602-89, IEEE Std 1233, 1998 Edition, IEEE Std 830-1998. Software creation tools as well as management and maintenance of content were developed. The software realizations of developed ECCSs with resources processing subsystems to set up e-commerce in online newspapers and online journals were presented.

**Keywords:** content, commercial content, information resource, business-process, content management system, content lifecycle, Internetnewspaper, electronic content commerce system.

## 1     Introduction

The lack of common standardized approach to the overall ECCS design as well as process of information resources elaboration causes a number of issues while developing appropriate systems typical architecture. Due to the lack of common and detailed classification of ECCS, it becomes problematic to define and form the unified methods of information resources processing in these systems. This creates problems for the implementation of the appropriate information resources processing subsystems in ECCS such as the formation, management and maintenance of content.

## 2    ECCS development project

The existing ECCSs work by unknown algorithms for a wide range of programmers/specialists in the field e-business [1-4]. While creating of new ECCS the teams of specialists have to re-develop methods/information resource processing tools and content life cycle support [5-9]. Teaching and learning materials for specialists in the field are missing [10-19]. The studies concerning patterns and level of impact on the ECCS functioning relative to implement of all or some stages of commercial content life cycle for information resources processing are missing [20-28]. The analysis of ECCS functioning results aren't available because of inability to access administrative units of existing ECCS which are already known, as they are commercial projects [29-37]. The novelty of project development lies in generalized typical architecture designing as well as methods, tools and technologies for ECCS creating, and implementation of commercial content life cycle stages. Implementation of formation subsystems, management and maintenance of commercial content in ECCS leads to a reduction of production cycle and time saving while distributing commercial content, increasing of potential/constant audience and number of participants in e-business, which promotes its active development and ECCS functionality extension. The developed recommendations concerning ECCS overall typical architecture designing which differ from existing by detailed elaboration of steps and presence of sub-processing information resources that make it possible to effectively maintain content life cycle at the level of systems developer (reducing the time and resources on developing, improving the quality of system operations) [38-45]. There were developed and implemented software tools for creation, management and maintenance of content in order to reach a greater effect of operation at the level of owner (increasing profitability, growth of users interest) and user (comprehensibility, interface simplification, unification of information resources elaboration process and wider choice functional capabilities) of ECCS.

## 3    Expected effects of ECCS introduction

*Expected economic effect* from ECCS application is predicted by reducing expenses for project development and architecture of the system, for additional resources and personnel usage in the absence of clearly regulated plan. The factors of economic impact are following.

1. The presence of commercial content forming subsystem reduces expenses for informational resource moderators extra staff, as it carries out the authors' and moderators' work, namely: collecting data from various sources, formatting and categorization of content, keywords and duplication identifying, digest formation as well as content distribution.

2. Content management subsystem reduces expenses for personnel responsible for information resource updating.

3. Content support subsystem reduces expenses for personnel responsible for collecting and system functioning analyzing.

4. The presence of information resources processing subsystems allows reducing time on actual unique commercial content prompt getting, which leads to ECCS target audience increase, and, respectively, increases ECCS implementation economic effect on several positions.

ECCS application leads to the growth of productivity of the project participants, improving the quality of information resources processing, reducing time spent on ECCS implementation and actual unique commercial content prompt receipt. The reasons for effect growth are following.

1. Increase of labor efficiency occurs through the usage of the results of additional professional resources work, such as Google Analytics, moderators, administrators, programmers and analysts.

2. Analysis and increase of labor efficiency leads to improving the quality of formatting, management and maintenance of the content as well as reducing time spent on system implementation and prompt content receipt.

3. Improving the quality of information resources processing is based on the analysis of statistics and main characteristics of ECCS functioning, such as number of visits, average time spent on visit, failures indicator, achievement of search goal, content dynamics, number of page views, number of page views per visit, new visits, absolutely new visitors, traffic source etc.

4. Reducing time spent on ECCS implementation and actual unique commercial content prompt receipt improves the decision-making quality for e-business participants:

a) The authors in order to create actual content according to the established, selected and distributed the list of digests related to the actual subject;

b) The moderators in order to form rules as well as operational data collecting source lists;

c) The moderators in order to form rules for categorization, duplication, formatting and content management, resource formation;

d) The administrators in order to perform resource and system management;

e) The analysts in order to perform the research system functioning statistics, formation of rules for identifying new stories and personalization of cooperation with user, content ranking.

*Organizing effect* lies in reducing the number of staff (moderator 1-3, administrators 1-2, authors 1-10, analysts 1-3, programmers 1-2), which is launched to preparation, formation and decision-making and changing its functions, which are performed by formation, management and content support subsystems (data preparation data for authors, tracking the results of staff performance, data collecting for analysts and moderators); changing the organizational structure (a clear division of functions between project participants, i. e. moderator won't be analyst, and vice versa); reducing the number of operations performed by personnel (a part of operations should be performed by ECCS via information resources processing).

*Technological effect* through the release or reducing such resources as staff, and more efficient usage of information resources processing subprograms in ECCS as well as clear allocation of responsibilities among participants, launched to certain project. The development of the new technologies such as formation, management and maintenance process of the content, and search engines peculiarities.

*Ergonomic effect* lies into influence of the results of ECCS operation and information resources processing through the formation, management and maintenance of content on number of visits, average time spent on information resources (min.: s), failures indicator (%,) achievement of searching purpose, content dynamics (% ), total number of page views, number of page views per one visit, new visits (%), absolutely new visitors, traffic source in %, etc.

*Psychological effect* lies in friendly interactive interface for each participant of the project, which simplifies work for authors, moderators, administrators, analysts, and improves psychological indicators for visitors and regular ECCS users through individualization of work.

*Advertising effec*t lies in unique content, content template and information resource usage, which improve searching results from search engines, and serve as self-promotion of ECCS, information resource and content. Cooperation with Google advertisement improves ECCS information resource advertisement rates as well as e-business.

*Social effect* lies in increasing the number of information resource visitors (permanent and unique), increasing the limits of target audience, resource availability and commercial content support, coverage of a wider social audience, ability to change the boundaries of the target audience through social regulation topics and information resource filling. The support of topically similar and relevant content, its uniqueness and efficiency of creation, formation, representation through information resource and support allows adjusting the limits of constant target social audience for ECCS and forecasting/adjusting these limits changes.

## 4    ECCS input data

List of input data types, their characteristics, description, classification allow them be conditionally divided into seven groups of incoming content, depending on the project participants' class: visitors, users, authors, administrators, moderators, analysts and other resources. The input data elements structure allows you to create the requirements to ECCS and its components as well as clearly limit their functionalities, describe the source, their incoming frequency and additional conditions/limitations which are imposed by input data source. ECCS input data are following:

1) a content from different sources specified by moderator predefined list (subscription, free and open content from information sources predefined list, author content, content as the result of the search engines operations etc.) in the form of data without predefined structure in HTML/XML-format in order to create commercial content [1-4];

2) information requests from users/visitors ECCS information resource as a text message in the appropriate fields (keywords/phrases in order to find commercial content, messages and/or requests to participants of the project) [1];

3) actual data (set of source address, words/phrases, and/or user logins with disabling printing/access option) and/or rules (set of operations like IF (fact i) THEN ban ELSE check OR permit/print) from ECCS information resource moderators;

4) actual data (ECCS and Google Analytics parallel cooperation statistics as XML-tables and/or fixed customized/personalized user actions such as viewing, downloading, storing and/or content searching) and/or statistical analysis rules of user action (formation of associative lists of popular, topical, outdated, recently revised and/or similar content/subjects or author works, and forecasting of thematic content demand) from analysts of ECCS operation;

5) information resources URL-address from moderators for information filters databases such as ECCS data source [1];

6) language dictionaries replenishment by moderators as a list of words, phrases with the defined characteristics (part of language and gender, number, cases etc.) as well as additional set of morphological features to each of them [3-4];

7) ECCS information resource operation statistical data, collected at specified intervals from Google Analytics as XML-tables [3-4];

8) comments and user feedbacks as a text data array completed in specially designated places of information resources;

9) members voting results regarding content and quality of the commercial content according to the defined numerical evaluation scale with the possibility to support linguistically inaccurate voting, e. g. "good" etc.;

10) individual statistics/personalized user actions (viewing of content, viewing time, content downloading/uploading, searching);

11) database content components, content collected from various sources, duplicate content, registered users, project participants, linguistic dictionaries in order to determine keywords and headings, keywords and subjects, keywords in order to determine subjects and their replenishment;

12) external advertisement in the form of banners from Google and partner sites;

13) thematic stickers of information and/or entertainment content (weather, exchange rates, anecdotes, announcements, etc.);

14) ECCS settings from administrators in the form of change/creation and liquidations of additional options and system configurations and/or information resource through the administrative part of the system.

## 5    ECCS output data description

ECCS output data are following [1-4]:

1) the final information product of ECCS operation in the form of content (article, announcement, digest, e-book, audio, video etc.);

2) answers on user's information requests (list and set of a similar subject content according to the content searching results using keywords and concepts which contained in the information request);

3) commercial content digests (short thematic announcements of content according to their set of keywords and concepts);

4) information resource visits' statistics according to the number of clicks on the link, time spent on visit, number of page downloads and additional investments;

5) user actions' statistics and/or resource visitors in order to form individual portrait of user/audience according to the number of clicks on the link, time spent on visit, number of page downloads and additional investments, switching between pages etc.;

6) forming/filling in of information resource page individually for user according to its statistical action history (selection of related content, current recommendations for thematic content etc.);

7) new rubrics/topics of commercial content (creation and formation of a new category according to the input data sources analysis, authors operations, comments and requests from users);

8) ranking the results of commercial content in the form of evaluation within a scale [0;5], [0;10], [0;12] or [0;100];

9) relationship tables of similar, recently revised, popular, outdated, authorial and/or topical content according to the keyword list of this content, user pages viewing analysis, sequence and time spent on thematic content viewing;

10) comments evaluation as the result of the content user's comments in the form of permission/ prohibition of printing on resources, and, if needed, with the prohibitive recommendations for certain user to post the following comments.

The list of output data, messages, their description, characterization, classification, method of formation and transfer allows to create precise functional requirements for the development and implementation of ECCS.

## 6      ECCS functional requirements

ECCS interface basic requirements: scalability/ performance while working with a large number of users, sessions, transactions and database connections; productive browser connection and back-end data storage; rapid development and deployment of Web OLTP-applications support; support of synchronous/asynchronous transaction management via servers. ECCS transactions servers characterize the following features: built-in transaction management services; the mechanism of starting and management of servlets; calls of distributed objects in order to ensure communication in multilevel applications; facilities for rapid software development for intermediate level, including component development. ECCS provides support to six interfaces: with limited access for visitors and users; without restrictions for the administrator and moderator; with free access to author and analyst. The access for users is implemented using login and password. Such services as a choice of content for a

certain period of time starting from the beginning of content filling through a calendar are additionally implemented there. The convenient rubricator allows you to select content by certain category. Searching is carried out by using keywords in the database.

System administration is carried out through administrator interface access to which is limited and implemented using login and password. The adjustments to the structure of the system/resources are made here as well as user access rights are added, edited or deleted, and content distribution rule change. Creating/editing content is carried out through author interface access to which is limited and implemented using login and password. The development of functional requirements for formation subsystems, management and maintenance of content in ECCS promotes the development of such systems typical architecture. On the one hand, ECCSs facilitate work of moderators, authors, analysts and administrators of these systems, and increase system functionality to their users on the other.

ECCSs select a range of topical issues in the form of content plurality from other sources for moderators and authors according to their rating through commercial content formation subsystem. The author creates his own commercial content according to ECCS analysis chosen from various sources of actual content. If needed, the moderator creates new rules for filtering content from various sources and updates the addresses of other sources in commercial content formation subsystem. Analyst analyzes the activity of the target audience and ECCS operation for the development of new rules and statistical analysis of the dynamics of commercial content lifecycle stages through commercial content support subsystem. These rules should increase the range of the target audience, number of visits, number of unique visitors, number of repeated visits, number of visits from search engines, number of direct visits, number of regional visits, number of visits etc. for thematic information resources in ECCS. The purpose of work is also to determine the functional requirements for information resources processing subsystems in ECCS as formation, management and maintenance of commercial content. Commercial content formation subsystem facilitates the work of the authors and moderators of ECCS. Content management subsystem facilitates the work of the authors and moderators of ECCS as well as supports different functionalities for users of these systems. Commercial content support subsystem facilitates the work of the analysts of ECCS. Information resource moderation is carried out through moderator's interface access to which is limited and implemented using login and password. Here are introduced the rules and parameters for monitoring content from different sources; commercial content is being added, edited or deleted; the content of the day is set (for publishing ECCSs); the content of static pages as well as rules of sending letters with the content is changed. There was implemented a convenient service for adding new clients with a possibility to group them into certain categories and set the time limits of access that is blocked automatically after a certain period. Statistic formation and its analysis is carried out through author interface access to which is limited and implemented using login and password. Here are made the amendments in the rules of information resource operating statistics; this statistics analysis rules are being added, edited or deleted; the

rules of commercial content ranking as well as commercial content rubrics and commercial content authors are established. There was implemented a convenient service for automated adding of information resource operating statistics and this resource users activity as well as the rules of this statistics analysis.

Software tools created for content version control ensure that the online portals content will not be lost or accidentally overwritten. Moderators and administrators have the opportunity to easily find the required version of content and information resource. Building of business processes based on roles and user groups means the independence from delays while execution by individual persons.

Structural elements designed for constructing interfaces and information resources processing software tools [1-4]:

1) Communication protocols (HTTP, FTP, IIOP).

2) Integration of HTTP-servers with information sources (CGI, Perl, PHP and specialized API).

3) Implementation of hypertext capabilities (HTML, WML, XML, XHTML, JavaScript).

4) Implementation of multimedia capabilities (Flash, formats for audio/video, VRML).

5) Communication implementation (POP, UDP, SMTP).

6) Calculation support (PHP, Java).

7) Systems and service of content management (CMS, CMIS).

8) Network management protocol (CMIP).

9) Organization of mobile access and calculations (GPRS, EDGE,UMTS, WAP).

10) Implementation and development of distributed objects (CORBA, COM,DCOM, EML, ORB).

11) Saving and processing of data (File systems, OS, database management systems).

ECCS technology is automation (full or partial) of the business process by which content, documents, information or tasks are transferred for the appropriate action from one participant to another according to the set of procedural rules. ECCS describes, creates and manages workflow (business process) using the software that interprets the process description, cooperates with the participants of the workflow and, if needed, elicits corresponding software applications and instrumental tools. ECCS automates a business process (not a function), and implements the rules of interaction between participants, as these aspects are the main centres of losses because of their ambiguity. The result of ECCS formation is such a system like online newspaper, online magazine, online editions, online publishing, distance learning, online store created for selling content in the form of e-books, photo, video, audio etc.

Standardization and implementation of functional requirements for ECCS formation provides the creation of a generalized approach for such systems developers in order to reduce the time for formation and implementation of such systems with avoiding of appropriate project development phase. Requirements for ECCS operation results, operation regulation, ways of displaying, transfer and store depends on the implementation of major information resources processing subsystems

such as formation, management and maintenance of commercial content. Requirements for compatibility and ways of interacting and communication with other systems lie in support of text arrays processing in HTML and/or XML-format. ECCS ergonomic requirements lie in convenience of maintenance and system maintenance, rational configuration of program and interface elements, convenience of system management tools, aesthetic design. ECCSs support security/protection of data and other system components from unauthorized access, loss, destruction, damage. Support of organizational and procedural requirements for personnel, its composition and qualifications, system operation charts, rights and powers to operate the system etc. allows to implement/introduce ECCS, maintain the system functioning process on high level and analyze the results of its approbation and its major subsystems of information resources processing.

The commercial content formation subsystem is based on the multilevel model of processes organization. Such an organization provides the separation of subsystem structure on such individual modules: gathering/creation of content from various sources, formatting, identifying keywords and concepts, headings, identifying duplication, formation of digests and selective dissemination of content between ECCS users. Content formation subsystem is implemented in accordance with the algorithm 1-2.

Algorithm 1. Formation of content analysis set

$1^{st}$. Gathering content from various data sources and saving it in DB.

$2^{st}$. Filtering content by a set of rules defined by system moderator.

$3^{st}$. Formatting content in XML-format and saving it in the database.

$4^{st}$. Determination of content duplication and duplicates filtering.

$5^{st}$. Determination of keywords and concepts of the content and saving them in the database.

$6^{st}$. Formation of content digest and saving it in the database.

$7^{st}$. Content formation and making an entry in the annotated DB.

$8^{st}$. Commercial content categorization.

Step 1. Identification of concepts from content using data from definitions database.

$2^{sp}$. In case experts weights are presented in the definition of the concept, then it should be calculated the weight of concept emergence in the text, taking into account the frequency of the phrase appearing in content.

$3^{sp}$. Adoption of decision on content belonging to a particular category based on the rules of categorization and set of concepts found in the text of the calculated weights.

$9^{st}$. Selective content distribution among moderators and authors of commercial content according to the direction of their work and their estimated rating of popularity and quality of work.

The principle of identifying keywords within the meaning (terms) based on Zipf's law and it comes down to choice of words with average occurrence frequency (the most used words are ignored due to "stop dictionaries", and the text rare words aren't taken into account). Then the terms for new keyword formation are synthesized, using

structural parts of speech base. The process of categorization using automatic indexing of commercial content components is divided into consecutive blocks: morphological analysis, syntactic analysis, semantic and syntactic analysis of linguistic structures and text content substantial writing variation. Based on analysis of the given set of content, authors create commercial content, which then passes the following stages of processing (alg. 2.)

Algorithm 2. Content formation for information resource in ECCS

**1st.** Author analysis of content plurality obtained from various sources and filtered.

**2st.** Author content formation as a result of the analysis of content plurality obtained from various sources and filtered.

**3st.** Author content formatting in XML-format and saving it in DB.

**4st.** Checking for commercial content duplication and doubles prohibition. In case there are duplicates, the author's bugs should be fixed, and move to the 10st, otherwise - move to the 5st.

**5st.** Verification and validation of commercial content. In case there is unique content $> w$, move to the 6st, otherwise - the author's bugs should be fixed, and move to the 10st.

**6st.** Determination of keywords and concepts of the content and saving them in the database.

**7st.** Author commercial content categorization.

**8st.** Formation of content digest and saving it in the database.

**9st.** Content formation and making an entry in annotated database.

**10st.** Recalculation of commercial content author rating.

Content distribution subsystem selectively sends digests among the authors according to their work quality rating. From the very beginning the subsystem receives ready digests from sources via RSS. Then digests are distributed among authors according to their rating: The author with the highest rating is the first person who receives digests for revision. In case all operations are ready, the subsystem goes into standby mode before the appearance of new content. Authors' rating indicates the performance/effectiveness of work of each of them individually. It is affected by the next criteria: the percentage of unique content (quality of the author work), the number of content views (weight of search and direct conversions), user evaluation (users' activity) and time spent on visit (a measure of users' interest in the content). Static representation of the content support model that describes the attributes and behavior of subsystems is presented on a class diagram. Principles of content analysis are the basis for the implementation of the subsystem. This allows automating of the various stages of formation of information product of such a type without loss of content and reducing quality. The results of its application while developing a number of commercial content projects confirm the effectiveness of subsystem functioning. The developed automation means for content formation allow accelerating of content formation process and increasing the usage rate of information resources which were created by their help. Commercial content management subsystem supports interactive communication between user and electronic content commerce system through information resource. The subsystem must form information resource to the

needs user and respond to his/her requests. Information resource consists of a clearly defined set of components (nodes) that are semi ordered (some of the components are clearly defined order, and their location in the graph is not necessarily defined). The number of information resource components and the corresponding volume of components set is precisely defined/evaluated. Construction of graph of navigation (navigation graph) on information resource is going on the basis of given ratio of adherence to the set of relations of resource components. Navigation is carried out without interruption and transition to a new node of the navigation graph is logical. The units are submitted by the random types of relationships that do not have significant limitations. The minimum number of project nodes determines the graph that provides the most meaningful information about information resource. Each arc in the navigation graph displays the elementary relationship between components of resource and must have an orientation. In order to receive content with certain parameters the user must pass a number of steps. Content management process by user or moderator is implemented through alg. 3.

Algorithm 3. Content management.
**1$^{st}$.** Authorization in electronic content commercial system.
**2$^{st}$.** Selecting a template content from the dialog box.
**3$^{st}$.** Selecting and setting various parameters.
**4$^{st}$.** Selecting the required functional under the protocol template.
**5$^{st}$.** Content generation out of template by the 3$^{st}$ defined criteria.

The basis of content management subsystem is a core - a subsystem that connects all parts into a single application and is responsible for loading and configuring modules, connecting general dependencies and providing information resources integration points. The most important task is to ensure that the information needs of system problem-oriented elements, maintaining access to data of different categories of users, compliance integrity and consistency of data, minimization and control under data excess, ability to develop and change the internal organization of information resource, compliance with the requirements of quality and data efficiency. ECCS provides modification of information resources through submission methods, formats, and content internal organization; content storing environment, physical storage units, technical means; user requirements, the emergence of new requirements and types of users; procedure for distribution of content and methods of user access.

# 7 Conclusion

Content analysis enables formation of conclusions regarding the possibility of influence on target audience growth, tendencies of information resource annual volume of visits increase over time. It is possible to investigate the dynamics of changes in global, national, regional, industrial, thematic content flow or dynamics of growth of homogeneous flow part on any basis. The basis for the obtaining data on dynamics of change and target audience growth is secondary sources, such as, Google

Analytics. In order to improve the accuracy of content analysis dynamic results there should be implemented the following steps: users' information portraits and content analysis, analysis of content thematic subjects, formation of content relationship tables, comments tonality, rating/content history and information resource statistics analysis.

## References

1. Vysotska, V.: Tekhnolohiyi elektronnoyi komertsiyi ta Internet-marketynhu. Saarbrücken, Germany: LAP LAMBERT Academic Publishing (2018)
2. Content Management Interoperability Services, Version V. 0.5, Hopkinton: EMC (2008)
3. Lande, D., Andrushchenko, V., Balagura, I.: An index of authors' popularity for Internet encyclopedia. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 47–55 (2017)
4. Lande, D.: Creation of subject domain models on the basis of monitoring of network information resources. In: 1st Inter. Conference Computational Linguistics and Intelligent Systems, COLINS, http://colins.in.ua/wp-content/uploads/2017/04/Lande.pdf (2017)
5. Vysotska, V., Chyrun, L., Lytvyn, V.: Methods based on ontologies for information resources processing. Germany: LAP LAMBERT Academic Publishing (2016).
6. Vysotska, V., Lytvyn, V.: Web resources processing based on ontologies. Saarbrücken, Germany: LAP LAMBERT Academic Publishing (2018)
7. Vysotska, V., Shakhovska, N.: Information technologies of gamification for training and recruitment. Saarbrücken, Germany: LAP LAMBERT Academic Publishing (2018)
8. Vysotska, V.: Internet systems design and development based on Web Mining and NLP. Saarbrücken, Germany: LAP LAMBERT Academic Publishing (2018)
9. CM Lifecycle Poster, Content Management Professionals, Access:http://www.cmprosold.org/resources/poster/ (2010)
10. Hackos, J.: Content Management for Dynamic Web Delivery. Hoboken: Wiley (2002)
11. Halvorson, K.: Content Strategy for the Web. Reading: New Riders Press (2009)
12. McGovern, G., Norton, R.: Content Critical. Upper Saddle River: FT Press (2001)
13. McKeever, S.: Understanding Web content management systems: evolution, lifecycle and market. In: Industrial Management & Data Systems (MCB UP), 103(9), 686–692 (2003)
14. Nakano, R.: Web content management: a collaborative approach. Boston: Addison Wesley Prof. (2002)
15. Papka, R.: On-line News Event Detection, Clustering, and Tracking: thesis for the degree doctor of philosophy. Amherst: Massachusetts University (1999)
16. Vysotska, V.: Computer linguistics for online marketing in information technology : Monograph. Saarbrücken, Germany: LAP LAMBERT Academic Publishing (2018)
17. Rockley, A.: Managing Enterprise Content: A Unified Content Strategy. New Riders Press (2002)
18. Stone, W. Plagiarism, Duplicate Publication and Duplicate Submission: They Are All Wrong! In: IEEE Antennas and Propagation, 45(4) 47-49 (2003)
19. Lytvyn, V., Vysotska, V., Chyrun, L., Smolarz, A., Naum O.: Intelligent System Structure for Web Resources Processing and Analysis. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 56-74 (2017)

20. Lytvyn, V., Vysotska, V., Wojcik, W., Dosyn, D.: A Method of Construction of Automated Basic Ontology. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 75-83 (2017)
21. The Content Management Possibilities Poster. http://metatorial.com/pagea.asp?id=poster.
22. Lytvynenko, V., Lurie, I., Radetska, S., Voronenko, M., Kornilovska, N., Partenjucha, D.: Content analysis of some social media of the occupied territories of Ukraine. In: 1st Inter. Conference Computational Linguistics and Intelligent Systems, COLINS, 84–94 (2017)
23. Shepelev, G., Khairova, N.: Methods of comparing interval objects in intelligent computer systems. . In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 100–109 (2017)
24. Orobinska, O., Chauchat, J.-H., Sharonova, N.: Methods and models of automatic ontology construction for specialized domains (case of the Radiation Security). In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 95–99 (2017)
25. Hamon, T., Grabar, N.: Unsupervised acquisition of morphological resources for Ukrainian. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 20–30 (2017)
26. Grabar, N., Hamon, T.: Creation of a multilingual aligned corpus with Ukrainian as the target language and its exploitation. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 10–19 (2017)
27. Hamon, T.: Biomedical text mining. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, http://colins.in.ua/wp-content/uploads/2017/04/2017COLINS-THAMON-keynote.pdf (2017)
28. Protsenko, Y.: Intuition on modern deep learning approaches in computer vision. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, http://colins.in.ua/wp-content/uploads/2017/04/protsenko.pdf (2017)
29. Kolbasin, V.: AI trends, or brief highlights of NIPS 2016. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, http://colins.in.ua/wp-content/uploads/2017/04/CoLlnS_TuS.pdf (2017)
30. Kersten, W.: The Digital Transformation of the Industry – the Logistics Example. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, http://colins.in.ua/wp-content/uploads/2017/04/CoLlnS_TuS.pdf (2017)
31. Shalimov, V.: Big Data – Revolution in Data Storage and Processing. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, http://colins.in.ua/wp-content/uploads/2017/04/BigData_eng.pdf (2017)
32. Hnot, T.: Qualitative content analysis: expertise and case study. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, http://colins.in.ua/wp-content/uploads/2017/04/Qualitative-content-analysis_expertise-and-case-study.pdf (2017)
33. Romanyshyn, M.: Grammatical Error Correction: why commas matter. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, http://colins.in.ua/wp-content/uploads/2017/04/Grammatical-Error-Correction-why-commas-matter.pdf. (2017)
34. Yukhno, K., Chubar, E.: Gamification: today and tomorrow. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 139–140 (2017)
35. Pidpruzhnikov, V., Ilchenko, M.: Search optimization and localization of the website of Department of Applied Linguistics. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 137–138 (2017)

36. Olifenko, I., Borysova, N.: Analysis of existing German Corpora. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 135–136 (2017)

37. Kolesnik, A., Khairova, N.: Use of linguistic criteria for estimating of wikipedia articles quality. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 133–134 (2017)

38. Kirkin, S., Melnyk, K.: Intelligent data processing in creating targeted advertising. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 131–132 (2017)

39. Hordienko, H., Ilchenko, M.: Development and computerization of an English term system in the fields of drilling and drilling rigs. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 129–130 (2017)

40. Gorbachov, V., Cherednichenko, O.: Improving communication in enterprise solutions: challenges and opportunities. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 127–128 (2017)

41. Didusov, V., Kochueva, Z.: Statistical methods usage of descriptive statistics in corpus linguistic. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 125–126 (2017)

42. Verbinenko, Yu.: Discursive units in scientific texts. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 120–123 (2017)

43. Titova, V., Gnatchuk, I.: Evaluation ofa formalized model for classification of emergency situations. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 110–119 (2017)

44. Kotov, M.: NLP resources for a rare language morphological analyzer: danish case. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 31–36 (2017)

45. Kuprianov, Ye.: Semantic state superpositions and their treatment in virtual lexicographic laboratory for spanish language dictionary. In: 1st International Conference Computational Linguistics and Intelligent Systems, COLINS, 37–46 (2017)