Content Management System in Educational Environment

prof. dr. sc. V. Mateljan[1], V. Juričić[1], mr. sc. G.Đambić[2]
[1] Faculty of humanities and social sciences, Zagreb, Croatia
[2] Keito, Zagreb, Croatia

***Abstract*—This paper points out and describes the characteristics that are shared between the most of today's content management systems. The intention was to find what makes the core of today's content management systems, i.e. what defines them. To provide this, authors analyzed great number of systems on a great number of attributes and features. Systems were compared and analyzed based on their approximate cost, supported operating systems and popularity. The result of those analyses indicates that the most represented and the most appreciated characteristics are support, interoperability with other systems and existence of additional applications that can be integrated into the web site, like blogs, newsletters and guest books.**

***Index Terms*—CMS, content management systems, education, analysis, comparison, characteristics**

I. INTRODUCTION

Increasing quantity of digitized information in education led to the need to collect and store information, in a way that provides quick access and usage of this information and that protects information from unauthorized changes. These needs are dynamic, they change in time, they are becoming more and more complex, and as a consequence, there is no permanent solution that satisfies all needs. The solution is a system that can be easily customized and enhanced without modifications of the existing data, but which also supports integration of new types and formats of digital information, and that satisfy the needs of modern education community.

Content management system proves to be extremely suitable for usage in online education, because not only it satisfies all fundamental needs of information activities, but it provides possibility to implement new methods of communication between members of education community. System is scalable, it is constantly developing and improving, utilizes new web technologies, and therefore ensures that its users always have a quality system to rely on.

There are numerous definitions of content management system, depending on the time of their creation and the author himself. The general definition of the web content management system says that it is software, usually implemented as a web application that is used to manage, create and control the content of a web site. To provide this functionality, but also to ensure the quality and accessibility of information, there are several common characteristics of every CMS:

1. *automated templates* – creation of standard templates that can be changed afterwards , in one central place, and applied to all content
2. *easily editable content* – content is separated from its presentation, and that way easier to modify
3. *scalable feature set* – additional modules can be installed to extend existing site's functionality
4. *web standard upgrades* – system can be upgraded, so that it supports current web standards
5. *workflow management* – possible creation and modification of content workflow
6. *document management* - tracking the changes in document lifecycle

New web technologies that were developed by experts, and new ideas and requirements that were created by the users of content management systems, resulted in growth of available CMS features. Thereby, the definition of CMS has also expanded, and the question that is raised now days is what makes the core of every modern CMS, i.e. what characteristics are the most and the least represented in today's content management systems.

II. RETRIEVING INPUT DATA

To make an answer to the above question, authors made an analysis on the great number of content management systems and their features. Input data for analysis was acquired from the web page CMSMatrix, which contains and maintains information for more than 850 content management systems. The characteristics are categorized, and, for the purpose of the analysis, divided into 10 sections, which are listed in the paragraphs below.

- General information

Category contains data about supported operating systems, programming languages, databases, web servers and application servers. Also, section describes what rights and privileges are required for installation and initial preparation of content management system.

- Security

Very important part of content management system is ensuring that stored information is protected from the users who are not allowed to access it. To prevent unauthorized access, system must implement authentication mechanism. This category analyses which authentication methods are supported, and the system's capability to track users actions; initial authentication, creation, modification and deletion of content. Depending on the importance of the stored information, system is exposed to the different types of attacks, so it is analyzed does the system have mechanisms to notify

administrators, or people who are in charge for this kind of situations, about forbidden actions.

- Support

Due to the changes and improvements of content management systems, there is a need to inform and to educate people that use them. This not only eases their work with certain parts of the system, but enables them to fully utilize various features of their system and to accommodate system to their specific needs. Category shows information about available types of support for the users, from online help and forums, to annual conferences.

- Ease of use

One of the most important characteristics of content management systems is the simplicity of their usage. It enables people with very little knowledge of web technologies to modify content of the web site, and, dependant on the system itself, administer the whole web site. Category analyses implemented tools and features that provide pleasant and quick work with the site's content, like WYSIWYG editor, spell checker, image resizing and multiple upload.

- Management

Depending on the size of the web site, it is imperative that content management system has a quality support for site's maintenance and administration. Category shows which types of administration are available, is there possibility to schedule and automate certain actions and to define certain rules that must be satisfied during the content publishing.

- Interoperability

Interoperability of the content management system refers to its ability to work together with diverse systems; the ability to exchange information with them and to use it. To achieve this, system must support different protocols and standards. This category analyses which protocols can be used to interact with the system, and to what standards is system compliant with.

- Flexibility

Flexibility refers to the system's ability to accommodate to individual user's needs. For example, site that is intended to be used by users from different countries should provide multilingual content and the localization of the user interface.

- Performance

Some content management systems are intended to be used by many users simultaneously and that could result in degradation of system's performance, depending on the platform and hardware. Category shows which actions can be made to improve system's performance. For example, supported types of caching, load balancing and database replication.

- Built-in applications

Build-in applications are mostly optional applications that do not come with the initial installation, but can be integrated into the existing site, if necessary. Applications can be roughly divided into three categories:

1. applications that are used to *present* different kind of information to users (chart application, newsletters)
2. applications that ease the *search* for certain information (search engines, site map)
3. application that are used to *retrieve* information from users (polls and quizzes application, guest book)

- Commerce

Category analyzes features that can be used in e-commerce, from putting items to shopping card, inventory management and prices, to pluggable payments and shipping. This category is the least relevant for analysis of currently available content management systems, because it covers completely different sphere of World Wide Web, which is not related to education.

III. PREPARING THE ANALYSIS

As said in the preceding section, authors retrieved data describing content management systems from the web site CMS Matrix. Because of large quantity of data on the web site, authors have made an application that parses the HTML content of each page, and that imports data into the database.

Initially collected data had certain disadvantages and was not suitable for analysis. First of all, not every CMS was entirely described with all the features. Problem was solved in a way that missing features were added, and it was presumed that those features were not implemented, i.e. CMS does not support them.

Also, the features were described with different values, what would affect the results of analysis, so there was a need to convert them to the more adequate ones. The values for which it was assumed that the system supports them are as follows:

- *Costs extra* – the feature is available, but does not come with initial installation, and one must pay in order to use it
- *Free add on* – the feature is available, and can be integrated into the content management system for free
- *Limited* – the feature is not entirely available, i.e. the feature is partially implemented
- *Yes* – the feature is available and comes with initial installation

For all other values it is taken that the system does not implement the feature.

Content management systems that implement less than 10 percent of the total number of features were removed and were not taken in further analysis. This was made to remove systems that are in the beginning of their development, that are currently developing or whose data was not updated on the web site for a long time.

IV. THE ANALYSIS OF COMMON CHARACTERISTICS

The purpose of this analysis is to find what characterizes today's content management systems, i.e. which characteristics are common to the majority of those systems.

First step of analysis was to calculate the number of content management systems that implement or support each feature. Then, authors have extracted the features that are supported by 70% of content management systems. The final result of this analysis is shown in the Table 1.

Table 1 shows for each feature its name, the number and the percentage of CMS that implement or support it. Also, data in the table is ordered, so that the most represented features are listed first. For example, online administration is the most represented feature, and has the ordinal 1.

Online administration means that system can be completely managed and administrated through a web

browser. This means that administrators and common users do not need to use any additional client software, but all the components can be managed online. This includes management of site's content, administration of site's options and templates and management of security related matters.

What You See Is What You Get editor is a web based editor that allows users to create rich text content without knowing web technologies and languages, like HTML, CSS and XML. It allows users to visualize what the article or any other content will look like when finished, without need to use tags or learn markup languages.

TABLE I.
COMMON CHARACTERISTICS OF CMS

Ordinal	Feature	Number	Percentage
1	Online Administration	772	91
2	WYSIWYG Editor	766	90
3	Friendly URLs	728	86
4	Granular Privileges	719	84
5	Commercial Support	687	81
6	Asset Management	670	79
7	Search Engine	667	78
8	Photo Gallery	657	77
9	Professional Services	640	75
10	Online Help	628	74
11	Content Approval	627	74
12	SSL Compatible	621	73
13	Commercial Training	616	72
14	Mail Form	615	72
15	Link Management	611	72
16	Server Page Language	611	72
17	Web-based Template	611	72
18	Content Syndication	611	72
19	Professional Hosting	610	72
20	Themes / Skins	606	71
21	Template Language	600	70
22	Page Caching	594	70

Friendly URLs refers to the systems capability in using short, but descriptive, human-readable URLs. Because most of the pages in CMS are generated dynamically, there is a need for certain data to be passed to the web server, and that often results in long and non-readable URLs, with lots of symbols and numbers. Also, search engines find it easier to index pages, when the URLs are shorter and descriptive.

Granular privileges refer to the lowest level of rights that can be assigned to the user. For example, one user has read and write privileges that apply to the all pages on the web site, and another user has read and write privileges limited to only some pages. It means that privileges are allowed on a per page basis.

Commercial support means that support can be purchased from a commercial organization that has trained staff members.

Asset management means that there is a central repository that content management system's users use to upload their files, and that those files can be reused anywhere on the site.

Search engine means that CMS implements its own search engine that can index content of the web site and that allow users to search indexed content.

Photo gallery is an application that is used for displaying thumbnails and images, and that can be also used to upload new images and delete existing ones (image repository).

Professional services refers to the commercially available services that provide customization and administration of CMS.

Online help means that CMS implements its own online help system that users can browse and query.

Content approval means that system has capability of defining workflow lifecycle that ensures that the work of an editor or publisher needs to be approved by reviewer before it is shown on the site.

SSL compatible means that a system, or one part of the system, can be used with the SSL certificate installed on the web server.

Commercial training means that training can be purchased from a commercial organization.

Mail form means that a system has an application for creating and customizing forms for contacting site's administrators or owners.

Link management means that a system has an application that allows users to browse site structure, to create, modify and delete links, and to change links order and hierarchy.

Server page language means that at least one server page language (like PHP or ASP) can be used to add custom functionality.

Web-based template means that there is a web-based interface that can be used to add templates and styles to the system to change design and layout of web pages.

Content syndication means that system can export content to one of the standard formats (like RSS or XML) so that content can be used by various applications and event republished on other sites.

Professional hosting means there is a service provider, with trained staff members, that offers creation and maintenance of web sites whose content is managed through content management system.

Themes / Skins means that system has mechanism that can transport themes and styles from one site to another, so that one theme or style can be created on one site and then reused on some other.

Template language is a fully HTML compliant language that is used for powerful layout controls.

Page caching is a mechanism that caches web pages so that if pages are requested more than one time, most of the processing required to create them is skipped.

A. *The results of analysis*

Analysis of the common characteristics shows that the most important characteristics are ones that provide complete control of the site's properties and content from the distance. There is no need for user to be physically present on the computer located in the local network of organization or institution, but he can fully utilize the features of the CMS remotely.

User interface for content creation is friendly and accommodated to users without knowledge of web technologies and markup languages. Many content management systems support server page languages and template languages. This enables more experienced users to extend system's current functionality and to create more powerful layout controls.

Support is an important element of today's content management systems. There are commercial organizations that provide creation and customization of sites, site's maintenance and administration. Some organizations also offer training and education.

Generally supported optional applications are ones that are used for searching content on the site, offering faster access to target information.

V. THE ANALYSIS OF CONTENT MANAGEMENT SYSTEMS AND OPERATING SYSTEMS

The analysis determines the differences and similarities between content management systems that support different operating systems.

First step of analysis was to calculate the number of content management systems that implement or support each feature. After that, CMSs are grouped by the supported operating systems and the features are analyzed as a group. Not all operating systems were included in analysis, but only the ones that are the most often: Microsoft Windows, Linux, Sun Solaris and MacOS. Authors also analyzed the features of the content management systems that are platform (operating systems) independent. The results of the analysis are shown in the Table 2.

For each analyzed operating system it is shown the number of content management systems that support it, and the percentage of this number in total number of content management systems. The sum of percentages exceeds 100% because many content management systems support more than one operating system. Common characteristics are also analyzed, but they are grouped by the operating system they support. Last column (Feature number) shows the number of features that are supported or implemented by more than 70% of content management systems.

TABLE II.
ANALYSIS BY SUPPORTED OPERATING SYSTEM

OS	CMS number	CMS percent	Feature number
Independent	376	46	19
Windows	335	41	47
Linux	199	25	37
Solaris	136	17	40
Mac OS	106	13	38
Non-common	41	5	34

The analysis shows that majority of content management systems are platform (operating system) independent, but also, those systems have minimum number of common characteristics. The most common characteristics have content management systems that support Microsoft Windows operating system.

Common characteristics are mostly the same for every operating system, although certain differences exist. CMS that support Windows operating system have one characteristic more represented that the rest of systems. Feature is called *sandbox* and means that there exists a private area for administrators and content management in which they can test and try new settings and ideas without affecting the work of the main site.

Content management systems that support Linux, Solaris and MacOS put more accents on the security of the web site and on the extendibility. Two characteristics related to security that are found common to those systems and that are more represented than in the rest of the operating systems are login history and audit trail. *Login history* means system keeps track of who and when logged in, from which IP address, which browser was used and so on. System also records the unsuccessful login attempts, which is useful in discovery of attacks on the web site. *Audit trail* means system keeps track of who and when created, modified or deleted content, and many other things, like user roles and themes. Majority of the system has also feature called *pluggable API*, which assumes that there is an open and well documented application programming interface that allows the system to be extended. Using documented and available API, developers and experts can write their own plug-ins and that way extend initial functionality of the CMS.

Platform independent content management systems have one characteristic that only few others have; they are XHTML compliant, meaning systems follow W3C specification for XHTML compliance. The number of those systems is 272, which makes more than 72 percent of the total analyzed content management systems.

VI. THE ANALYSIS OF APPROXIMATE COST

The analysis shows the relation between commercial and free content management systems. Free content management system is presumed to be a system that has GNU General Public License or that has free installation and basic features. That means that some additional features do not need to be free, but this analysis considers only capabilities of content management systems.

Comparison of those two types of systems showed that there is a very little difference between features they implement, with exception of features in 3 categories: flexibility, security and support. Categories with features in which content management systems mostly differ are shown in Table 3.

Table 3 is split into three parts: flexibility, security and support. For each feature is shown what percent of content management systems implement it, first column for commercial and second column for free content management systems.

Commercial CMS have slightly weaker percentage of implemented interface localization feature. This features means that system is internationalized; that it can be translated into other languages and it can take local preferences into account, like date and time formats.

Analysis shows that commercial CMSs, in average, have much more features related with security and support. Certain organizations exist that provide technical support, hosting and training. Free content management systems have better extensibility, because they are mostly open source software. Their code is available and well documented, and that allows developers to make their own add-ons to extend system's functionality and usage. Public

forum is also one feature that free CMSs have more often than commercial ones. Public forum means that it exist publicly available forum or message board, which allows users to comment, ask for help and discuss about content management system.

In average, security features are more often implemented in commercial CMSs than in free ones. Although granular privileges are approximately the same, safety of the content, information and privacy is more often considered in commercial CMSs. Versioning means that system provides some way of tracking and discarding the changes that are made to the content.

TABLE III.
THE ANALYSIS OF COMMERCIAL AND FREE CMSS

Feature	Commercial CMS percentage	Free CMS percentage
Flexibility		
Interface Localization	62	70
Security		
Granular Privileges	88	83
Content Approval	80	65
SSL Compatible	80	63
Audit Trail	76	61
Sandbox	75	58
Login History	74	63
Versioning	71	50
Support		
Commercial Support	90	62
Commercial Training	84	50
Professional Services	84	52
Professional Hosting	83	51
Online Help	80	66
Pluggable API	65	72
Public Forum	48	82

VII. THE ANALYSIS OF THE MOST POPULAR CONTENT MANAGEMENT SYSTEMS

The analysis compares the characteristics of the most popular content management systems. Popularity is determined by the data provided on the Blogsweek site, and the ranking is based on the three months of aggregated traffic data. Site tracks user's interest for certain content management system, in a way it calculates the total number of times users demanded content management system's data. The rankings are based on the limited set of users, and do not represent the opinion of the entire Internet population.

Top 10 CMS by popularity for 2007, according to Blogsweek site, are shown in the table 4.

TABLE IV.
MOST POPULAR CMSS IN 2007.

Ranking	CMS
1	WordPress
2	Joomla!
3	Drupal
4	ExpressionEngine
5	Mambo
6	Xoops
7	MODx
8	DotNetNuke
9	CMS Made Simple
10	b2evolution

Content management systems listed in Table 4 have very similar system requirements, with exception of DotNetNuke. They are platform independent content management systems that work with MySQL database and support any application server that supports MySQL and PHP. They are free and distributed under GNU General Public License. Drupal has similar requirements to ones listed above but also supports PostgreSQL database. ExpressionEngine is also a free content management system, but license is commercial, and includes full source code.

DotNetNuke is a Web Application Framework that supports Windows operating system and Microsoft SQL server database. System is free and under BSD license, which means that application framework has no limitations in commercial and non-commercial usage, with requirement of giving credit back to the DotNetNuke project community. It requires IIS application server and web server, and supports .Net Framework programming languages.

The analysis of the characteristics of the most popular content management systems is shown in the Table 5. Table shows common features of those systems, separated into categories. As the analysis shows, most popular content management systems have many optional applications that a common user frequently uses, like blogs, frequently asked questions applications and newsletters. They also have multilingual support and the capability of customizing user interface, to accommodate date, time and number formats to those that are common in user's country.

Such systems are capable of using RSS technology (Really Simple Syndication), not only to retrieve and display data from other sites, but also to export their own data so that other sites and applications can use it.

TABLE V.
ANALYSIS OF THE MOST POPULAR CMSS

Category	Feature
Built-in Applications	Blog
Built-in Applications	Events Calendar
Built-in Applications	FAQ Management
Built-in Applications	File Distribution
Built-in Applications	Guest Book
Built-in Applications	Newsletter
Built-in Applications	Syndicated Content (RSS)
Flexibility	Metadata
Flexibility	Multi-lingual Content Itegration
Flexibility	Interface Localization
Interoperability	Content Syndication (RSS)
Interoperability	UTF-8 Support
Management	Themes / Skins
Management	Web-based Template Management
Management	Content Scheduling
Security	Login History
Support	Developer Community
Support	Pluggable API
Support	Public Forum

VIII. CONCLUSION

The growing needs of content management system's users and development of new technologies are one of the factors that influenced expansion and enhancement of those systems. Systems include many features, they are flexible and easy to use which makes them adequate for numerous solutions.

Content management systems are also suitable for educational needs because not only that they provide an easy and instant way of publishing information, but also enable various kinds of communication and interaction between members of educational community. This paper provides a brief overview of the currently available content management systems and gives the current status of their development. The analyses that were made showed that the systems are accommodated to users, who can accomplish all their tasks with minimum knowledge of the web technologies. On the other hand, web experts do not have limitations in expanding site's functionality and layout, by using server and template languages.

It is shown that today's content management systems take great care of support, flexibility and interoperability related matters, with a tension to improve security and extend built-in applications. This gives them excellent conditions for their further usage and expansion.

REFERENCES

[1] The content management comparison tool, CMS Matrix, CMS Matrix, 06.02.2008, http://www.cmsmatrix.org

[2] Blogs platforms and content management systems, BlogsWeek.com, BlogsWeek, 05.02.2008, http://www.blogsweek.com/en/

[3] Typo3 CMS: Feature list, Typo3, Typo3, 06.03.2007, http://typo3.com/Feature_list.1243.0.html

[4] Online administration, Solutions Open Source, Open-net, 2007, http://open-net.ch/fre/eZ-publish/CMS-eZ-publish/Online-administration

[5] Web application framework, DotNetNuke, DotNetNuke, 2008, http://www.dotnetnuke.com/Products/WebApplicationFramework/tabid/777/Default.aspx

[6] W. Powel, C. Gill, Web content management systems in higher education, *Educause quarterly*, 2003.

AUTHORS

V. Mateljan is with the Faculty of humanities and social sciences, Department of Information Sciences, I. Lučića 3, 10 000 Zagreb, Croatia. (e-mail: vladimir.mateljan@gmail.com)

V. Juričić is with the Faculty of humanities and social sciences, Department of Information Sciences, I. Lučića 3, 10 000 Zagreb, Croatia. (e-mail: vedran.juricic@gmail.com)

G. Đambić is with the Keito, XII Podbrežje 3, 10 000 Zagreb, Croatia. (e-mail: goran.djambic@gmail.com)

This article was modified from a presentation at the 31st International Convention MIPRO 2008 in Opatija, Croatia, May 2008. Manuscript received 8 June 2008. Published as submitted by the authors.