# **ELECTRONIC PUBLISHING AS A GRAPHIC PRODUCT**

## Željko Zeljković (), Dragoljub Novaković (), Saša Petrović () Neda Milić Keresteš (), Stefan Đurđević () University of Novi Sad, Faculty of Technical Sciences, Department of Graphic Engineering and Design, Novi Sad, Serbia

**Abstract:** Electronic publishing has a particularly significant application as a graphic product in the publication of some products such as books and newspapers that have traditionally been essential products of newspaper printing, book printing and the like. In the electronic publication of newspapers, there are certain templates that are related to the identity of the publisher. There is a number of published and printed editions of newspapers that were published before the introduction of electronic editions. Newspapers stick to tradition and try to transfer printed editions to electronic ones, that is, to do the so-called digitalization. The process of digitization is a complex and extensive work consisting of analysis and proper preparation of the transformation into a bit of a mapped image with the help of reproduction devices, rearranging and arranging data with OCR software, creating metadata, database management and web application development. As current products of graphic production, multimedia appears as a set of different media that affect the human senses. Procedures for digitalization of graphic products with essential elements are necessary, and various software tools are needed for the realization of the process. This paper points out the most important segments of previously mentioned procedures.

Keywords: digitalization, printed media, electronic media, multimedia, metadata, electronic publishing

#### 1. INTRODUCTION

Digitalization, by definition, is the conversion of an analogue input into a digital form. In the case of graphic products, the goal is not to lose the quality of the input material by digitalizing the archive. Digitalization is a planned and expensive process, which should be conducted following known principles and priorities and according to established rules. The paper analyzes multimedia, modern software tools, metadata, based on which the concept of metadata is set. Metadata is "data about data" with a description of the content of electronic publications used during the search. Effective access to the required content means that these contents are well described using the metadata. When publishing some graphic products, such as newspapers and books, the problem is not in the use of technology, but in the organization of work. The advantage of digitalized media is that they can be easily and quickly reproduced without damaging the original, a large number of users have access to the material regardless of time and geographical limitations, the content is not damaged during the use, digital content can be searched and modified. Disadvantages include digital protection, digitalization converts images into bits that must then be reconstructed in order for images to be displayed, and shaded transitions in images can be transformed into sharp contrasts. In order to read an electronic document, it is necessary to have additional equipment and considerable financial resources needed for digitalization. Publishing graphic products such as newspapers is a complex process where certain attributes of advantages and disadvantages of electronic media in relation to print media get highlighted. The main difference between printed and electronic newspaper media is in accessibility and coverage. These differences can be compared using the basis of comparison (IFLA, 2002). In terms of the characteristics of the transfer of newspaper data, the printed media is a material form that delivers news and information through printed newspaper publications in relation to the electronic medium that creates, delivers and accesses news and information through electronic devices. To understand the print media, one needs to be literate in order to read certain content. In contrast, in the case of electronic media, this is not the primary requirement because it can be watched and listened to on presentation devices. The lifetime of the data in printed media is unchanged because what is printed remains on the substrate, while in electronic media, it can be changed and updated. It is not possible to exchange comments using printed media, while there are no restrictions with electronic media. The distribution of printed media has many limitations, while electronic media have no limitations. The language of the printed media is related to the publication, while the electronic media has the possibility of rapid change. E.g. newspaper print is available in all places where the reader brings it, while electronic editions of newspapers are dependent on the Internet. Printed newspapers can be archived and available

in the future, while the electronic ones are tied to server capacity and data storage. Visualization is on the side of electronic media through video content (Lukačević et al, 2015).

Printed newspaper media can be described as a mean of mass communication used to inform the general public through printed publications - newspapers, magazines, books, etc. Electronic news media is a newer form of mass media in which electronic devices are used to create and spread news and information. In both printed and electronic media, there is a connection between the generation and transmission of information, which depends on the current specific content of the media. Presentations in electronic media have certain specifics of displaying that must be taken to account in their design. A limitation of the display format can be noticed when compared to the printed media. For the printed media, paper formats act as a display, while for electronic media, this depends on the size of the device display. The quality of both is related to the display resolution, for printed ones it is the resolution in lines per inch and for electronic ones it is the screen resolutions. Printed and electronic media have common elements in data archiving devices. A unique feature is their recording capacity. Recording capacity is characterized by data compression procedures, which allow the reduction of a certain amount of data without a visible reduction in quality. An essential characterization of compressed data is its ability to be transferred to different media. Modern ways of data transfer are realized using various techniques such as the Internet, satellite transmission, high-speed cable connections to receiving devices that can reproduce them. These devices are various systems for the presentation of transferred data such as computer monitors, projection devices, audio playback systems and other devices. The control systems of these devices are various hardware elements and software applications. The use and distribution of printed and electronic media is regulated by the copyright protection of the media authors. A significant segment of digital media development is the field of cryptography, which deals with the protection of digital data. Another important area of development is data compression. Both of these areas will occupy a significant place in the future development of digital media.

The initial phase of media creation is media preparation. Further phases include tools and software for creating printed or electronic media. These media can be distributed in the created form as electronic or printed media. Their integration results in multimedia.

### 2. ELECTRONIC MEDIA

Electronic media can be characterized as time-dependent in terms of their monitoring, while printed media can be characterized as static. Ideas, content and their form from information sources are the basis for output. This data belongs to the field of media preparation. Data from the preparation are further directed, depending on the need, to the printed media or the electronic media in order to process or create products intended for end-users. Processing in the production process of electronic or printed media results in media products, electronic information from electronic media and printed products from printing media. Together they make up multimedia. The structure for the production of electronic media, print media, and multimedia is given in Figure 1.



Figure 1: Flow for the production of electronic media, print media and multimedia

People encounter multimedia every day. It is all around, on the Internet, computing, education, telecommunications, the world of games, advertising, e-commerce. Figure 2 shows the possible media combining which the multimedia gets created.



Figure 2: The possible media combining which the multimedia gets created

Planning and development of multimedia content consist of three main steps, them beeing:

- analysis,
- planning,
- realization.

Each of these steps can be further broken down into smaller steps as shown in Figure 3.



Figure 3: Content planning and development steps

Multimedia affects different senses of a person. Figure 4 shows all the human senses that are affected by multimedia content and the senses which will be affected in the near future.



Figure 4: Multimedia and its influence on the human senses

There is a tendency for the sense of smell and touch to be included in multimedia applications. It may still seem unthinkable from today's point of view, but it will soon become generally accepted. In the future, these devices will be able to connect to a PC and will have one special chip for scents.

Typography also has a long tradition in the printed media. When designing educational materials, their influence on the emotions of the recipient must not be neglected. Every sensory detail of the project needs to be carefully planned. The first impression of a multimedia content depends on the creativity of the designer, i.e. from the harmony of letters, colours and graphic elements.

When creating multimedia, in addition to hardware and software knowledge, it is also necessary to be familiar with design, didactics, psychology, communications, marketing. Media is an integral part of communication. They are an essential part of communication in society and among people. The media are used to make communication successful, both in individual and public communication. The media are always closely connected with the issue of communication. Every communication requires a medium, a mean by which a message is exchanged between communication partners and which is the basis for mediation between partners. The same goes for direct communication. The fundamental importance of media lies in the fact that they achieve an effect that contributes to the success of communication.

A successful presentation or performance is often labelled multimedia, thus emphasizing the robust, multifaceted and simultaneous flow of information between performers and participants in events. The multimedia information system should, in interactive communication with the user, simultaneously use various forms of information, such as text, graphics, animation, static or moving images, music and speech. Therefore, multimedia technology includes new input/output units for automated data acquisition from the environment. Such units are television camera, microphone, scanner. They generate multimedia entities: video, audio, images that are only partially interpreted via programs because they are insufficiently structured. Figure 5 shows the concept of multimedia as an integration of different media.



Figure 5: Multimedia as an integration of different media

Distributed multimedia information systems should enable the transfer of multimedia data from one or more sources located inside or outside of the workstation. Multimedia technology, which is changing rapidly, includes machines that transmit experience not only through text and equations but also by acting on human senses, through sound as well as moving images. Besides, multimedia still do not have an adequate way of management, and it is difficult to predict their development. What makes all this possible is a digital and optical technology that is evolving rapidly, enabling efficiency in the processing, storage, maintenance and addition of new information.

Multimedia combines text, image, sound, animation and video, and very different means have been used to play them before, although lately a multimedia computer has been used to play multimedia, and a CD-ROM or the Internet to store data.

In the interpretation of the media, various software can be used to interpret the solutions. The suitability of such an interpretation belongs to (Poliščuk, 2004):

- artificial intelligence,
- expert systems,
- neural networks,
- fuzzy logic,
- genetic algorithms,
- hybrid systems,
- distance learning,
- multimedia,
- general-purpose software systems,
- other application systems.

Artificial intelligence - artificial intelligence can be defined as a computer technique. The technique of artificial intelligence consists of the development of software applications, the execution of which can be characterized by an intelligent process similar to the decision-making process of humans. Research in the field of artificial intelligence begins and follows the development of computer systems. Expert systems are one of the parts of development in artificial intelligence that are based on the logic of expert thinking in a particular area where expert decisions are realized in the form of programming systems with special software tools.

**Expert systems** are intelligent computer programs that emulate problem-solving in the way that experts do and represent one of the most important areas of artificial intelligence research. The systems solve real problems from different areas, which would otherwise require human expertise. The goal is that the computer program always gives correct answers, in a given area, not worse than an expert, but that is difficult to achieve. An expert is an expert in a field who possesses and effectively uses certain knowledge, understanding of problems and tasks, skills and experience. Experts also have the ability to recognize a typical task in a specific problem they solve. Figure 6 shows the areas that define expert systems.



Figure 6: Areas that determine expert systems (Kukolj, 2007)

It is challenging to classify the tools for the development of ES into certain groups since the boundary between them is not strictly defined. It is often the case that individual tools can be classified into several groups according to their characteristics. Programming languages had a specific developmental course. Figure 7 shows the evolution of programming languages.

Artificial neural networks are a form of artificial intelligence that tries to mimic the function of the human brain and nervous system. Neural networks are being used to solve an increasing number of everyday problems of significant complexity. In programming, they can be used as a "generator" that is able to perform various recognizations and classifications, which also has the ability to generalize when deciding on unstructured input data.

**Fuzzy logic** is multivalued logic developed from phase set theory, as an extension of set theory. Fuzzy logic does not mean that it is imprecise logic. The term "fuzzy" means that logic describes vague concepts, such as human knowledge. In the logic phase, the degree of truth or the degree of membership represents the degree to which an element belongs to a particular phase of a set.

Genetic algorithms and genetic programming belong to the class of automatic methods for creating a computer program or other complex problem-solving structure. The idea of genetic algorithms is based on the processes of imitation of natural selection.



*Figure 7: The evolution of programming languages(Lisjak, 2004)* 

**Hybrid systems** - most of the traditional information systems, which are based on knowledge, are developed as independent systems, with minimal interconnection. Increasing the amount of information requires the development of more complex systems, which integrate knowledge and traditional processing. One generation of intelligent systems is being developed with the help of hybrid techniques. Hybrid systems are suitable for solving various application problems

**Distance learning** - in the conditions of rapid technological changes and changing market conditions, the educational system must provide increased educational opportunities without increasing material resources. Many educational institutions are responding to this challenge with the development of distance learning programs, which have become a vital need today due to the pandemic. Content delivery methods used in e-learning are shown in Figure 8.



Figure 8: Content delivery methods in e-learning

**Multimedia** in the process of education is primarily set on influencing sight and hearing. The learning process is realized through the media transfer of information in the relation of knowledge transfer "knowledge giver - user of knowledge" or "lecturer-user", "professor-student", "teacher-student", etc. or

different executors as providers of knowledge and different categories of recipients of knowledge. Significant conditions and contribution to the progress of distance learning have been made possible by the development of information and communication technologies.

General-purpose software systems are software applications of a universal character.

Application systems are dedicated applications developed to solve specific problems.

### 3. DIGITALIZATION OF PRINTED MEDIA

Over time, news outlets have realized the need for modern digitalization of earlier print editions. This process is preceded by technical requirements and the implementation of digitalization. A digital image is an electronic photograph mapped in the form of a series of image elements, pixels and arranged following a previously defined ratio of horizontal and vertical. In the process of digitalization, the original documents are translated into bitmapped images with the help of a scanner or camera. The scanned document is stored as a digital image record without data compression (e.g. TIFF format) or with compression (JPEG format). Such images can be subsequently processed on a computer, where their brightness, contrast, colour intensity, sharpness, etc. can be increased or decreased. Reasons behind image compression are high memory space requirements, slow devices that are unable to display an uncompressed image in real-time, network bandwidth that does not allow real-time image transfer.

Media microfilming was a commonly used archiving process. Microfilming is one of the most efficient, economical and most straightforward ways to store large amounts of information on different types of films: silver halide, visual and diazo film. Advantages of microfilming: the document on film is an identical copy of the original, it is possible to save large format documents in the overall perspective of the image, connected to compatible output devices, can be printed, faxed and e-mailed, the cheapest way to obtain archival information, availability is not limited to electronic devices, a lot of information can be stored in a smaller space than needed for the printed material, it can be scanned and thus provide an identical copy that can be stored permanently. Disadvantages of microfilm are: the use of film is limited to one person, users must have reading devices, it is necessary to have a suitable printer or copier in order to reproduce it, pages must be read in order, there is no possibility to search specific information, microfilm records cannot be updated without the use of microfiche technologies, durability and usability of the film are regulated by many factors (film quality, proper processing control, appropriate storage conditions and handling) (Lukačević et al, 2015).

OCR software is needed to allow extracting, editing and rearranging data from scanned documents, camera images, PDFs, or in general - original documents. The OCR software must have dictionary support in the language in which the source document is written. How the OCR software works: the program analyzes the structure of the document image, after that the program divides the page into elements such as blocks of texts, tables, images, etc. The lines are divided into words and then into signs. After the extraction of the scanned images, the program compares them with a set of sample images. It advances several hypotheses about the possibilities of scanned images. Based on these hypotheses, the program analyzes different variants of breaking lines into words and words into signs. After processing a vast number of such probability hypotheses, the program finally makes a decision, presenting a recognizable text. When converting materials, it is very important to:

- Choose scanners that can be adjusted to the physical dimensions of the original document, the type of media, the range of details, tones and colours on the documents, as well as the physical condition of the documents
- Choose a resolution, bit depth that is sufficient to convey the finest essential details in the newspaper which are to be scanned and to apply newspaper improvement procedures very carefully.
- Use standard compression techniques "lossless" for the parent, i.e. master files
- Try to strike a balance between the visual quality acceptable to users and the file size that the computer can access with an acceptable delay.

**Metadata** is structured information that describes, explains, locates, or otherwise makes it easier to find, use, or manage a source of information. Metadata is often defined as data about data or information about information.

Metadata can be descriptive, structural, and administrative. Metadata can be embedded in a digital object and can be stored separately. Metadata is often embedded in HTML documents and image file headers. Storing metadata together with an object ensures that it will not be lost, avoids the problems of linking data and metadata, and provides security that metadata and the object will be updated together. Flaws of metadata: creating metadata is expensive and time-consuming, metadata is complicated (especially for end-users), metadata is subjective and context-dependent, there is no exact limit to the level of detail up to which metadata is useful, metadata is useless (given the fulltext of the browser such as Google) (Mirosavljević et al, 2018).

**Format Metadata in librarianship** - Formats for machine-readable bibliographic data recording: MARC21 and UNIMARC. The two largest library systems in Serbia, COBISS and BISIS, operate on the UNIMARC platform. The role of metadata in librarianship is to find the source of the data according to the relevant criteria; to organize and manage various archives; to enable the exchange of information and content between different systems, data structures and interfaces with minimal loss of content and functionality; then digital identification (via permanent identifiers such as URLs or DOI numbers); and finally archiving and protection of digital information. Metadata schemes: Dublin Core, TEI (Text Encoding Initiative), METS (Metadata Encoding and Transmission Standard), MODS (Metadata Object Description Schema), LOM (Learning Object Metadata) (Mirosavljević et al, 2018).

**Dublin Core -** The basic characteristics of Dublin Core are: simple, adaptable, international, expandable, the elements themselves are optional, repeatable, can appear in any order, contains a codebook and its application is recommended.

**Data models -** Many different data models have been developed over the last fifty years. Some of them were just an attempt or a step in the development of other data models, while others came to life and left a mark both in the theory and in the practice of databases. Some of the models that have been used, or are still used in practice today, are:

- network data model,
- hierarchical data model,
- relational data model,
- model of entity and link types,
- object-oriented data model and
- XML data model.

**Hierarchical and networked data models** emerged in the second half of the 1960s. Database management systems based on a hierarchical and networked data model were soon introduced into everyday use. However, these database management systems did not lead to the expected increase in productivity of either programmers or users. This is due to the fact that the logical and physical aspects of the database were not sufficiently separated, data structures were complex, and procedural languages were used as operational components.

The relational data model originated in the 1970s, and the main reasons for its introduction were the introduction of a clear distinction between the logical and physical aspects of the database, structural simplicity and the existence of declarative language for defining and using the database. In previous data models, information about the physical structure was built into the programs themselves. If the physical structure of the database were to change, the programs affected by that change would also have to change. In the relational data model, the largest organizational unit of data is the n-array relation, which contains a set of n-torques.

XML data model - XML (eXtensible Markup Language) is a standard set of rules for defining data in electronic form. It is prescribed by the W3C organization. Its main advantage is that software for working with XML documents exists on all computer platforms (PC / Mac, Windows / Linux / Unix / macOS). It gives full results only when used with related technologies, some of them being: DTD, XML schema, XPath, XQuery, XSL transformations.

The XML data model has been in use for more than a decade. Although not new, it has not been sufficiently researched, especially compared to the relational data model. The XML data model, like other data models, has a structural, integrity, and operational component. The structural component of the XML data model is based on the concept of an XML document, at the extension level. At the intension level, an XML Schema document or DTD document is used. The XML data model is specific compared to other data models in that the XML extension can exist without the XML intension. This is due to the self-explanatory nature of the XML extension. However, if there is no schema according to which the XML document was formed, one cannot speak of an XML database, in the real sense of the term database) (Mirosavljević et al, 2018).

The XML language has made a real revolution in the field of data storage and transfer in the past decade, with a tendency to spread further. The language is human and machine-readable at the same time, with

an emphasis on scalability and simplicity. A large number of formally defined languages have been developed on its basis.

XML is a meta-language for tagging text documents. The language is standardized by the international organization W3C (World Wide Web Consortium) for the development of Internet standards. Many formally defined markup languages are based on XML syntax. An XML document is a text document that can contain only textual content, while binary-type data cannot appear in it.

XML notation is used to describe structured documents. Extensible Markup Language (XML) is a family of document specification languages, which gained great popularity soon after its introduction in 1998. XML is easy to understand and easy to learn, standardized and open to further extensions. XML is used to store, publish, and share information, and allows communicating parties to compile their language with well-defined semantics, syntax, and structuring rules.

Using the XML language to form documents helps to avoid keeping documents in a private format. XML documents are said to be self-explanatory because they contain both data and metadata. In addition to exchanging XML documents, it is necessary to define how to store, search and update XML documents. In order for all parties to the communication to use the same type of XML documents, it is necessary to define the semantics, syntax and rules of document structuring. Intensive data exchange via XML documents has created the need to store, search and update them. This has led to the development of database management systems based on XML technology. They are used in building XML database systems. XML has languages and concepts that can be used in the role of data models with its structural, integrity and operational component, thus providing all the necessary conditions for expressing the XML database schema in a formal way. The XML data model is in intensive development.

The Database Management System (DBMS) is software introduced as a link between users (user programs, applications) and database records on disk. User programs do not access the data directly, but communicate with this program, through an interface close to the user. Figure 9 provides an overview of software that currently represents the state of the art and have a significant share of the world market. Databases are usually implemented using one of the tested software packages (Zeljković, 2016).

Producer	Product	Operating System	Languages
IBM Corporation	DB2	Linux, UNIX (razni), MS Windows NT/2000/XP, VMS, MVS, VM, OS/400	SQL, COBOL, Java, etc.
Oracle Corporation	Oracle	MS Windows (various), Mac OS, UNIX (various), Linux, etc.	SQL, Java, etc.
IBM Corporation (pre: Informix Software Inc.)	Informix	UNIX (various), Linux, MS Windows NT/2000/XP	SQL, Java, etc.
Microsoft	MS SQL Server	MS Windows NT/2000/XP	SQL, C++, etc.
MySQL AB	MySQL	Linux, UNIX (various), MS Windows (various), Mac OS	SQL, C, PHP, etc.
Sybase Inc.	Sybase SQL Server	MS Windows NT/2000, OS/2, Mac, UNIX (various), UNIXWare	SQL, COBOL, etc.
Hewlett Packard Co.	Allbase/SQL	UNIX (HP-UX)	SQL,4GL, C, etc.
Cicom System Inc.	Supra	MS Windows NT/2000, Linux, UNIX (various), VMS, MVS, VM	SQL, COBOL, etc.
Microsoft Corporation	MS Access	MS Windows (various)	Access, Basic, SQL

Figure 9: Software tools for working with databases

#### 4. CONCLUSION

As a consequence of the growing automation of all phases of the printing process, strong innovation potential in the fields of machines, devices and processes will be released in the coming years. This can be best seen on the basis of technical development in the preparation of printing with short periods of change of the device on which the preparation is realized. The control electronics enable high-quality retention and higher productivity.

Today, the field of press preparation is, as a rule, in digital form. Only in this way is it possible to shorten the delivery time of printed products and meet the high demands of customers related to quality.

Basically, all printing plants are expected to be more flexible in the production of various printed products. In accidental printing, more than half of the orders are already in digital form. The trend towards multicolour as well as towards decorative and special colours is growing. The volume of editions is decreasing, which opens the perspective for small short-term editions. According to expert opinion, such publications have the greatest chance of growth.

Positive impulses for the entire printing industry are expected from the information and communication sector.

In addition to the classic areas of work, media production will gain in importance in the future. Thus, the needs for the design of printed media and the creation of multimedia products, Internet sites, printed media in combination with electronic media, etc. will grow more intensively. The additional capability of electronic media, especially the Internet, will increase the competitive pressure on printed media and will partially replace printed products, but will also create new requirements for printing.

The market for printed products remains, despite the growth of electronic media, as a further attractive and an extensively strong market, as a large part of the ever-growing advertising fees flows into the print media with the widespread growth of the needs for printed and electronic media.

With the spreading of the Internet, the sale of goods and services over the Internet by many companies is expanding as a new form of sales. The Internet is inevitably changing traditional media.

The digitalization of our cultural heritage connects different sectors of the global community in a whole new way. Due to the great use of the Internet, user groups of library structures have been transformed. Newspaper publishers integrate print and online publications to reach a broader range of readers.

Interactive Internet techniques provide libraries and archives with a new opportunity to develop a global community of users, so the newspapers also need to follow this trend in order to survive. The old archive needs to be digitalized, and the new editions need to be properly preserved and organized.

### 5. ACKNOWLEDGEMENTS

This research (paper) has been supported by the Ministry of Education, Science and Technological Development through the project no. 451-03-68/2020-14/200156: "Innovative scientific and artistic research from the FTS (activity) domain".

#### 6. REFERENCES

- [1] IFLA, "Smjernice za projekte digitalizacije", (Nacionalna biblioteka Crne Gore Đurđe Crnojević, Cetinje, 2002.), page 29.
- [2] Kukolj, D.: "Sistemi zasnovani na racunarskoj inteligenciji", (Fakultet tehničkih nauka, Novi Sad, 2007.)
- [3] Lisjak, D.: "Primjena metoda umjetne inteligencije pri izboru materijala", PhD thesis, Faculty of Mechanical Engineering and Naval Architecture, 2004.
- [4] Lukačević, S., Buljević, J., Mokriš, S.: "Formati digitalnih sadržaja za čuvanje i zaštitu novina u knjižnicama", Knjižničarstvo: glasnik Društva knjižničara Slavonije, Baranje i Srijema 19 (1-2),104, 2015.
- [5] Milosavljević, B., Zaric, M., Zeljković, Ž.: "Elektronsko izdavaštvo materijal sa predavanja", URL: https://www.grid.uns.ac.rs/predmet2.html?predmet=88 (last request: 2020-08-10), 2018.
- [6] Poliščuk, J.: "Ekspertni sistemi", (ETF Podgorica, Podgorica, 2004.)
- [7] Zeljković, Ž.: "System for identification of printing process parameters", PhD thesis, University of Novi Sad, Faculty of Technical Sciences, 2016.



© 2020 Authors. Published by the University of Novi Sad, Faculty of Technical Sciences, Department of Graphic Engineering and Design. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license 3.0 Serbia (http://creativecommons.org/licenses/by/3.0/rs/).