

**Nikola Popović**

Ministry of Foreign Affairs RS, Belgrade, Serbia

**Julijana Mirčevski**

Court Expert Witness, Belgrade, Serbia

## **ANALYSIS OF THE PROBLEM OF DOCUMENT DIGITALIZATION IN STATE ADMINISTRATION Centralization vs. Decentralization**

**Abstract.** This paper analyzes the quality of digitalization within state bodies themselves. The functions of three document management systems were examined: an existing classic application, an internal Document Management System, and an external Document Management System. The relationships and specificities of these systems' interaction with their environment, as well as their impact on the efficiency of public administration services, were investigated. The paper presents experiences and suggestions for improving efficiency in the domain of managing documents in digital form. It highlights the advantage of dynamic data and information exchange compared to the concept of static information provision offered by a Document Management System. The need for introducing new methods of information search and consolidation is elaborated. The process by which purposefully structured information is generated from a digitized mass of documents is examined.

**Keywords:** state administration, digitalization, DMS, ePisarnica, digital security

### **1. Introduction**

The digitalization of documents in public administration is today a technologically stable process. A multitude of applications use the digital document as a basic element for creating information. However, viewed from the perspective of an individual institution, the digitalization process still generates certain anomalies. Indeed, after several years of monitoring the interaction of work processes in a given institution and the organization, i.e., the spectrum of functionalities of the various software systems on which the digitalization process is based, the projected digitalization systems have shown limitations. Disruptions manifest particularly in internal information processes within the institution itself. The problems were pointed out in a timely manner, but the designers suggested they would be resolved on the go. It turned out that the nature of the "disruptions" was at the level of the digitalization system's conception. The conception of the digitalization process was set up so that "everything is done to make it easier for the citizen to use the spectrum of services," i.e., "citizen first," "simplifying the realization of citizens' rights," and of course "savings." The manifested contradictions indicate that the applied methodology was to some extent misaligned with reality [1]. It is difficult to expect efficient service delivery to the citizen if the digitalization system has not appropriately transformed the work of the observed institution or group of institutions.

Logical and technological aspects of institutions' work with documents are generally prescribed by the *Regulation on Office Operations* [2] and a set of relevant laws and other acts covering areas from the concept and realization of the digital document to the protection of systems using digitalization technology. With due respect

to the efforts of a large number of people who designed and organize the digitalization process in the Republic of Serbia, it is necessary to state that the term "office operations" is not only outdated but represents an anomaly or misinformation. In the context of a computer network, local, and especially the Cloud, the concept of an "office" as a classic physical space simply disappears. A citizen communicating with an institution communicates with a function, not an "office." Similarly, offices within an institution are turning into virtual spaces of given functions. Officials are generally not tied to a physical space but to the hardware device through which they communicate with internal "functions." The physical space containing paper documentation, which arises during the execution of "functions," has long significantly overlapped with the virtual (and physical) space of digital documents. Even the duties and work organization of the observed institution's officials change significantly in a digitalization environment. However, the authors of the digitalization process took the path of least resistance, retaining a mass of concepts from previous rules of operation of state institutions ("job systematization", etc.). The adopted conception of digitalization firmly adheres to the principle of centralizing digitized functions. For decades, there have been comparative analyses of the advantages, disadvantages, and contradictions of choices between centralization and decentralization at all levels, from large systems to small ones, for example at the level of small shops and their logistics. The authors of digitalization adopted the centralized document management model as preferred.

The digitalization process in public administration objectively takes place in the context of public-private partnerships (so-called outsourcing). Digitalization tasks, according to the adopted conception at the state level, are regularly operations of very large scale, which cannot be realized by the internal IT services of state bodies. Digitalization requirements are projected to be realized in the shortest possible time, which also incurs disproportionate financial costs. A minimal participation of the academic community, i.e., domestic scientific institutions, in the realization of the digitalization process is noticeable. It is apparently considered that there is no time for specific professional or scientific research in the domain of public administration. In the authors' opinion, the stance that the existing knowledge of the corps of state officials and a mass of non-governmental sector organizations is sufficient for the realization of a large-scale digitalization project of the state administration system is inappropriate.

The concept "Citizen First" is promoted as the primary goal of digitalization. Designing technological systems important for society as a whole should not be based on just one important aspect. The digitalization process introduces very significant changes into the entire society and is therefore inherently multi-aspect.

For more than two decades, the application of so-called "best practices" has been promoted. This is particularly insisted upon in the domain of activities related to integration into the European Union. The term "*best practice*" is interpreted as the import of (best/selected) foreign experiences. However, those foreign experiences logically arose in significantly different contexts (cultural, historical, technological, etc.). There is a phenomenon of uncritical copying of others' experiences. It is noticeable that this term is regularly used by individuals or organizations that have only marginal knowledge in the field of project management, and even less knowledge about the specificities of the area related to the given project. Objectively, the phrase "best practice" is a synonym for concealing insufficient knowledge of the idea and methods for solving problems that the given entity deals with.

## 2. Problem of terminology and language

The terminology problem is of great importance for the digitalization process. The uncritical adoption of new/foreign terms and their introduction into widespread public use creates corresponding contradictions and ambiguities. The general term "document" has been introduced into use, where it is continuously necessary to specify whether it is a document in electronic form or a document in paper form. With the new *Regulation on Office Operations* [2], the use of the term "registry material" ceases, and the term "documentary material" is introduced. In this case as well, the word "material" creates doubt about whether an electronic document also falls under "material" and similar ambiguities.

Work on computer tools for processing natural languages, in the specific case of the Serbian language, is significantly behind other linguistic areas. The digitalization process typically requires a series of specialized types of Serbian language dictionaries. It is not known to us that there exists an "official" *spelling checker* for all thematic areas of the Serbian language. Experiments have been conducted, but texts from the press and similar sources were used as a basis. Since a "manual" *spelling checker* was done for the mentioned internal registry system, a preliminary check was performed on 30,000 registry records. Testing document databases established a percentage of errors during entry, at least 6%. The expected number of errors is even higher due to the specificities of texts recorded in registries (foreign and domestic names, acronyms, and many other terms.). Data entry operators, due to high workloads, often do not use text spelling checks. Incorrectly entered text in a digitized system means that a document in a certain context may not be found. In the other two observed document management systems, we did not notice the presence of a *spelling checker*.

It is characteristic that OCR (*Optical Character Recognition*) programs mostly contain capabilities for recognizing documents in the Serbian language. In the *TESSERACT v5* program, the size of the so-called *traineddata* for languages can be observed:

Language / Script	Size of traineddata file
Serbian - Cyrillic	11.6 MB
Serbian - Latin	13.7 MB
Slovenian	14.8 MB
Croatian	18.7 MB
Russian	26.8 MB

The size of these files clearly indicates how much work is actually being done on individual languages, i.e., how much importance is given to the Serbian language in the digitalization process. OCR technology will continue to be used for a long time in the work of state bodies as well as in relevant scientific fields and culture.

A particular problem is the lack of functions in the domain of "*Part-of-Speech (PoS) Tagging*". The presence of a large number of digital documents makes the task of mass manual assignment of keywords pointless. For the automatic assignment of keywords, document classification by topic, and other similar tasks that would replace human labor, there are neither the necessary tools nor methodologies for the Serbian language. In this text, we will not delve into the much more serious problem of applying artificial intelligence methods and tools in office operations and related areas. The need for applying artificial intelligence in the economy is often discussed. However, without

the necessary linguistic software tools, there is no application of artificial intelligence either. How the necessary LLMs (*Large Language Models*) for the Serbian language domain will be generated has no clear answer. Using foreign LLMs and corresponding APIs (API - *Application Programming Interface*) is something that must be analyzed and addressed from various aspects.

### 3. Problem of data protection

From the beginning of the general digitalization process, the need for data protection has been emphasized by all subjects of the digitalization process. From an engineering standpoint, observing the three mentioned systems, it can be concluded that data protection problems remain insufficiently resolved. A centralized data storage and protection system formally looks like a final solution. However, the example of the *Republic Geodetic Institute* in Belgrade is useful. This institution, which holds a truly large resource of data of state importance, was one of the first involved in the digitalization process. It is noticeable that, for at least three years, there have been problems with repeated "hacker" attacks on the cadastre system. Similar or related problems exist in the information system of the *Ministry of Justice* [3]. An assessment of available (formal and informal) information provides a basis for concluding that sufficiently detailed assessments were not made during the system realization process, nor at the level of the adopted digitalization conception. The conclusions of the *State Audit Institution* (SAI) report for the justice domain [4] were a sufficiently serious warning. Viewed from the information technology aspect, large systems as a rule require an evolutionary approach to development and implementation. In this case, the primary criterion of the digitalization project was the speed of implementing the digitalization system, based on strict requirements from the financiers and other project participants - organizers. Great effort has been invested in bringing about a set of laws covering areas from the very concept of the digital document to the methods of using digital documents and systems for their storage.

Text [3] states that the *cyber* attacks were of such nature and intensity that at one point the *Geodetic Institute* had to discontinue using the *Data Center*. It is justified to assume that the highest quality (usually also the most expensive) possible hardware and software equipment was applied in the digitalization project. Interesting views are presented in S. Janković's article [5], which among other things relates to monopolies in the field of hardware and software production intended for data protection and especially the data transfer process. It is useful to draw an analogy with the market for "artificial intelligence" systems. The largest companies in the information technology domain have already designed the market for selling artificial intelligence products, with product prices and stock market values. However, outsiders such as DeepSeek or Qwen suddenly appeared with product prices almost two orders of magnitude lower than the competition. Simultaneously, both media and professional literature point to an increasing problem of reliability in the operation of AI products from the largest market players (non-existent references, conclusions derived in a problematic manner, unverified claims, etc.). The logical question arises whether similar problems exist in the mentioned market of software systems for data protection.

In the given situation, when there is objectively no choice, nor alternatives in protection systems, it is reasonable to develop and apply additional own logical protection methods (encrypting documents made in docx or pdf format, etc.). And of course, a dynamic form of protection through frequent, planned changes of protection models and methods. It is necessary to keep in mind that aggregates of open data and

information can very often take on the character of confidential information. For example, the number of unresolved court cases in a targeted area, in a given time period, can become confidential information depending on the context. Therefore, we believe that when using *Data Center* services, no distinction should be made between open and confidential data, i.e., both classes of documents should be protected in the same way. It is rational that the previous protection [6] be performed on-site, more precisely before delivering documents to the *Data Center*.

From the aspect of applying the concept of centralized information services, it is necessary to keep in mind the event [7] of October 20, 2025, when there was a sudden outage of *Amazon Web Services*. According to later reports, the problem significantly expanded. The agency [7] lists the impact on financial platform operations in seventh place. Although the system outage during the day was regulated from *Amazon's* perspective, further reports do not mention the consequences of such an event or to what extent the consequences were regulated on the side of numerous users. The perimeter of the impact of the event on 20.10.2025 is enormous, but there are still no more precise data on the duration and success of the recovery process of the multitude of users of systems that were under its influence.

Analysis of the adopted regulations and program agendas related to digitalization so far shows that the term "information logistics" is not mentioned anywhere, and emphasis has been on access possibilities and data access speed as well as planned large savings. Data exchange reliability is "assumed" but without evaluation of the applied digitalization methods. It is necessary to adapt the ways of designing and implementing the information structure that supports the work of state body systems, especially in individual state bodies. Work methodologies, and especially data protection methods [6], can differ depending on the environment: interaction with citizens, with the economy, with foreign partners, etc. For example, ministries of foreign and internal affairs by definition must operate in a 24/7 regime. Also, they must have multiple data transmission channels realized by different technical systems.

From the perspective of data protection (protection of data on physical carriers), it is important to keep in mind the amount of digital data generated annually in one state body. For example, in a given state body employing about 1,000 associates, less than 4 TB of data is generated annually. These are data from the domain of office operations (incoming and outgoing documents and accompanying collections of digitized documents). The current price of an external disk of that capacity is around 200 euros. The *Data Center* is logically a solution for a large part of the problems of working with data. However, the information presented above indicates that it is necessary to organize and maintain in readiness a local backup system, located within the state body. Regardless of problems and events that may or may not be foreseen, the state body must be able to protect and use its information resources under all conditions and recover them in the shortest possible time.

#### **4. Overview of the current state and observed problems**

Three systems intended for document management were observed. One has the form of a classic application and is still in operational function since 1990. The second system is a local *Document Management System*, introduced in 2020. The third is the *Document Management System ePisarnica* maintained by the *Office for IT and eGovernment*, examined since 2024. The third system, *ePisarnica*, is still not in regular use because contradictions in usage possibilities have manifested.

In the introduction, it was pointed out that the application of any version of a *Document Management System* changes the content of the concept of "office

*operations*". External (citizens, economy, institutions) and internal users (officials) communicate with a distributed, virtual "office" that is no longer limited by physical space. A longer, planned observation of the two mentioned modern document management systems was conducted. By monitoring the use of available functionalities in practice, it was concluded that interaction with the environment functions significantly better than use within the medium in which they are implemented. By definition, the requirements set before a document management system are met. It was noted that both systems provide efficient document distribution (to external and internal users), as well as efficient document storage. However, analysis of the used functionalities indicates that these functionalities, in a logical sense, have long been present and used in email systems. Distribution of messages to an individual or group functioned correctly. Perceived shortcomings of email systems have always been in the domain of data and text search and especially in the domain of data storage. This relates to limited quotas for the number of messages, limited memory capacity, and similar problems. It is useful to note that the concept of document management appeared at the end of the eighties of the last century, and that the current digitalization process is being carried out on rather outdated foundations.

The dynamics of content and the dynamics of data structures change with the digitalization process. The virtual office becomes an interface with a function or a system of functions, which are executed at the user's request. Observing the work of system users within an administrative unit gives a picture that does not meet expectations regarding the "virtual office." Namely, the use of paper documents, theoretically, is reduced, but analysis of the organization of work processes, i.e., manipulation of digital documents, indicates new qualitative problems. The digitalization strategy is predominantly aimed at helping users and strongly strengthening information system protection. The *Regulation on Office Operations*, in an updated form harmonized with the digitalization process, fits into such an orientation. Research of documents related to digitalization and corresponding normative documents indicates that the concept of "*generating new information as a product of the digitalization process*" is literally not mentioned anywhere. Furthermore, analysis of the functionalities of the mentioned systems, as well as others available on the market, shows that they explicitly do not support processes of generating new information, i.e., specialization. The primary obstacle is the process of data protection at all levels of data access hierarchy. An elementary example is the process of identifying incorrectly entered words. Analysis of errors in the data entry process, or the formation of dictionaries of terms used in the linguistic corpus of a given administrative unit, for example, the domain of foreign affairs, is effectively impossible. Such analysis requires the abolition of all levels of access protection to each individual document, i.e., the annulment of the entire protection system. Viewed from the other side, the formation of term dictionaries and maintaining the correct status of documents is a prerequisite for any, even elementary, use of artificial intelligence methods. Even at the technical access level, there is no possibility for performing such operations. For example, correction or editing (harmonization) of incorrect transliteration of foreign names, etc. There is no talk about content dynamics and especially data structure dynamics, nor is this problem mentioned in the documentation of the observed systems. Generating new information typically requires additional, often auxiliary, but also completely new data structures, which are not foreseen by software projects. This observation points to an important contradiction in the existing, applied digitalization concept.

The question arises as to how much the internal work organization (job systematization, actual human resource structure) of state bodies is adapted to the digitized environment. The digitalization process requires re-examining concepts like

"Catalog of jobs" or "Register of administrative procedures", from the perspective of digitalization, not the other way around. Foreign *mega-companies*, including *Microsoft*, in the digitalization process, vigorously reduce the volume of human resources and require their significant IT education and skills. No published documents were found that would more precisely define jobs specifically adapted to work with digitized documents, nor work processes and sub-processes adapted to organization and work activities in a digitized environment. It must be kept in mind that part of the documents still cannot be introduced into the digitized environment, not only for technological reasons but primarily due to valid laws. The impression is gained that jobs in state bodies are still mostly performed in the same way as when paper was used. It could be concluded that it is not just a matter of people's habits, but limitations imposed by the very organization of software support, which simply replaced a paper document with a digital document. Officials consider a scanned document to be a digital document, but it is only an image of a document. Metadata are not entered into the PDF file itself, for the PDF to be a self-contained digital object. A digital document must have its defined location where it is stored to be a digital document and to be always found. A digital document must have its coordinates to be able to be found in the multi-dimensional space of documents. That space includes virtual folders organized by topic, by physical storage, etc. Each segment of the document's URL (*Universal Resource Locator*) represents one dimension.

In the context of the above analysis, the basic problem is the purpose and structure of software projects for document management that are present on the market or are being introduced for use in the state body system. Generally, the idea of a system for "document management" is older than three decades and as such is definitely outdated, primarily due to the development of information technologies themselves. A follow-up problem, especially in our environment, is that designers persistently try to make "universal software systems," primarily due to the idea of selling the software in as many copies as possible. However, "ideologists of universality" overlook the fact that state bodies differ formationally or thematically by jurisdiction, and that documents in state bodies primarily belong to the class of unstructured, atypical documents (i.e., they are not order forms, invoices, etc.). Local self-government bodies, although integrating most functions of state bodies, also have their specificities. Software producers and especially designers persistently neglect these diversities, most likely due to insufficient knowledge of the specifics of state body system operations. Also, a lack of a modular approach to software design can be observed, i.e., adapting the software structure to functions and the way they are realized in reality. A third problem that regularly manifests in the process of implementing such systems is the duality of data carriers. Despite the ever-increasing use of documents in digital form, by nature, documents on paper as carriers of data and information must still be used.

A special class of problems relates to work with documents that have a confidentiality mark. Confidentiality can relate to personal data, in banks, in registers of health institutions, in courts, up to confidential correspondence at the level of state bodies or correspondence with foreign entities. Generally, the importance of data protection is emphasized, and specialized protection methods [6] are designed. However, the previously mentioned reports on problems in the work of the *Ministry of Justice*, the *Geodetic Institute*, etc., indicate the fact that data protection systems and the system itself have still not reached the necessary level of reliability. Viewed from a practical angle, this conditions that confidential data, i.e., documents, should still be stored on isolated software systems, however much this is contrary to the applied idea of digitalization. Besides, such an approach generates another problem - which software

to use for confidential document corpora. Parallel, isolated systems create difficulties in using documents, and especially financial and organizational problems.

Digitalization processes regularly encounter the problem of existing ("old") data. A simpler case is when existing data is recorded in paper documents or in the form of documents on machine-readable media. Paper documents are scanned and placed in required locations in the computer system, or reorganization of existing documents on computers, doc, pdf, and other formats, is performed. A more complex case is when a state body or any public sector institution already uses some form of document management software. In that case, the problem of exporting data and documents from the existing system and importing them into the new system appears. Project implementers of digitalization often try to skip the problem of data format conversion by citing the obsolescence of data importance. However, in a significant number of cases, this is not possible. The problem is reflected not only in adapting data formats but also in different data system organizations. Special problems appear when working with lookup tables, for example, lists of organizational units of the institution, lists of document senders, lists of document recipients, lists of classification marks of thematic areas, etc. Over time, data in these tables have changed, the names of document senders or recipients have changed. *Lookup* tables in the new system being implemented must either contain references to old, changed data names, or data structures must be changed to contain old names within metadata tables not as *lookup* data but as data in static fields. Such requirements regularly create problems in implementation processes. Special problems are generated in institutions dealing with administrative procedures that last a long time, several years or longer. Changes in classification marks of thematic areas as well as changes in document retention periods are sensitive operations that require careful planning.

Marking acts, i.e., assigning an identification number under which a document is registered in an institution, is a procedure that requires precise formulation. Paper documents were most often registered with a number of simple structure adapted to paper registries or simpler computer registries. Documents generated in document management systems are usually of more complex structure, often containing a mark of the organizational unit in which the document is registered, a classification mark of the topic to which the document relates, and other specificities. We analyzed in detail the identification number of a document registered in the *ePisarnica* system, which consists of 30 or more numeric characters. For comparison, examples of act identifiers in domestic systems [8] and acts of foreign origin are given:

Document identifiers of domestic systems	Identifiers of foreign documents
<b>01 broj:</b> 337-1-158/24-13 <b>Broj:</b> 119-01-57/2024-01	<b>Our reference n<sup>o</sup>:</b> J4/2025/0060
Document number from ePisarnica: <b>Broj:</b> 003563455 2024 10527 001 003 000 001	72014L0040FRA_240395 (EU) ili A/C.1/80/1.56 (United Nation Gen. Ass.)

Detailed research on the internet, versions of numbers under which documents are registered, showed that only documents from the *EU headquarters*, which is itself a large organizational structure, have document numbers up to 20 characters in length. Identifiers of *United Nations* acts, as a large system with very well-organized documentation, were also analyzed. The document number from *ePisarnica* contains a minimum of 30 digits. It is noticeable that foreign document identification numbers also contain letters, thereby increasing the range of marking possibilities. For example, with 26 alphabetic characters and 10 numeric characters, it is possible to mark 36 classes

with one character. Using digits from 0 to 9, only up to 10 classes can be marked. The *Instruction on Office Operations of State Administration Bodies* from 2022 [8] defined such a structure of identification mark for matters or petitioners. The *ePisarnica* designers accepted this without analysis of best practices in other similar systems - administrations of the *European Union*, *United Nations*, and similar entities. When introducing new digitalization systems, little is said about the problem of linking documents with the old mark from the existing system. The old identification mark of a document, or the identification mark of an old document, is metadata like any other metadata. This remark also relates to the need to link digital documents with paper documents, which is not a rare case.

Scanning paper documents is a field in itself. Scanning is done on local desktop scanners, and the digital copy, a file in pdf format, is marked with some prescribed label. However, nowhere is it mentioned that it is useful to provide the name, description, and other metadata of the document in pdf format, as internal metadata of the document in pdf format. In this way, the pdf file containing the document image becomes a *self-contained* object that can be found with appropriate software tools even if misplaced within the system's hardware resources. Pdf file names are not unimportant data if the corpus of digital documents is observed over a longer period. Although pdf files and various digital image formats have long been a standard, it is necessary to regulate procedures by internal rules for the case that new standards appear. Similarly, handling paper documents must be defined in much more detail than is mentioned in the *Regulation on Office Operations of State Bodies* and respective regulations. Here, the conclusion imposes itself that legislators and the archival segment of public administration have still not sufficiently specified work procedures. The use of microfilm is rarely mentioned, although it is an information carrier with a long usage time and without major technological problems. There is visual readability without the use of specific, changing technologies.

The volume of data generated in the observed state body is on the order of 4 TB annually. If it is assumed that the data volume is even an order of magnitude larger, what should be the retention period for the digital documents themselves? Namely, the *Regulation on Office Operations of State Bodies* and the corresponding *Instruction* define retention periods for documents depending on the (current) class of document - 3 years, 10 years, and further. Under conditions of using *Data Center* services, according to the presented estimates of data volume growth, the technological aspect of storing digital data should not represent a more serious problem. Therefore, retention periods for digital data in a certain sense represent an anachronism. If there are links between digital and paper documents, different retention periods can cause confusion, so it is necessary to regulate handling by internal regulations.

Problems with using old software or transitioning to a new generation of software have long been observed. It is commonly believed that by switching to a new operating system and implementing a new set of software tools, most problems are solved. Situations where *all* available systems for working with digital documents are innovated simultaneously are rare. For hardware components, a common problem is drivers for different devices, especially when changing versions of operating systems and/or other necessary software tools, etc. Current virtual machine technology allows simultaneous use of different versions of operating systems, different other software systems, in the process of synchronization and consolidation of hardware and software platforms.

## 5. Information consolidation and search

Organizing information consolidation involves collecting, structuring, and presenting data from multiple sources into a unified coherent format. The primary purpose of information consolidation in state bodies is to support the decision-making process. The scope depends on the purpose of the state body and the required thematic depth. For example, in the domain of legislation or justice, literally everything is context. In the domain of foreign affairs, due to the dynamics of geopolitics, retrospective analyses on very different aspects are frequent. Data sources are both internal and even more often external. Analysis of the two observed Document Management Systems based on WEB technology indicates two types of deficiencies. Connection possibilities with other software systems are significantly limited so that in practice, only manual methods of information consolidation are possible. Due to difficult connection conditions, the application of restrictive APIs between different software systems and the entire data protection system, search possibilities are also limited. The primary possibility is finding relevant documents within one system. The concept of "thematic depth" requires abandoning thinking of the "one document = one record" class. The documentation fund of any state body, especially in the domain of foreign affairs, is an ecosystem. It is necessary to move away from applying basic metadata (*Dublin Core standard*) to encompass thematic relations. Consolidating means synthesizing information over a set of documents or a set of different document collections. It is necessary to enable deep content analysis without *full-text search* limitations as well as maintaining the origin of data from a set of documents, but with the possibility of forming aggregates. There is a possibility of forming thematic virtual collections of documents and other forms of information. In such cases, specialized search interfaces are necessary. A *Document Management System* (DMS) and an *Information Management System* (IMS) are related but serve different purposes. DMS is focused on storing, organizing, tracking, and searching digital documents. An *Information Management System* (IMS) represents a system of broader purpose, which manages all types of information, including structured data (databases), unstructured data (documents), and partially structured data (emails, XML). IMS can encompass document management as its subset. IMS supports manual entry, automatic data import, and integration with other systems, as well as advanced search functions (full-text search, by keywords, fuzzy search, etc.).

Centralization by definition requires a significant level of standardization of data access and the presence of hierarchical organization. System centralization is useful in terms of data access scope and their storage. Simultaneously, for technical protection reasons, hierarchical relations are established that hinder the development of data and information management systems created and circulating at the level of the state body. *State bodies*, by their specific functions, have needs for information at different levels of internal organization, but simultaneously can have needs for information from different levels of other state bodies. Interdepartmental cooperation of state bodies takes place according to hierarchical organization; dispatches with information requests move between *state bodies*. Digitalization should enable the development of thematic networks and adaptation of thematic depth within individual network components. For example, the interaction of ministries of foreign affairs and education in the function of educating foreign students in the country. That is the basic cause which already at the current stage of digitalization generates strong contradictions and, by nature, slows

down the digitalization process. The adopted digitalization concept at this moment does not have the methods and means that would simply resolve the mentioned contradictions.

In texts [9], [10], the mentioned problems were already emphasized, but current digitalization processes are guided by another concept.

An analysis of the possibility of applying the Open Source software tool *KNIME* [11] is underway. *KNIME* is an extensive set of software procedures for data analysis (statistical procedures, neural networks, machine learning procedures, etc.). It possesses a very useful set of procedures for editing and transforming data into the required format. Procedures for forming complex data processing flows and automating their execution are built-in. The graphical interface classifies the *KNIME* system into the domain of 'no-code' programming. The application of the *KNIME* system in "horizontal" connection of databases of different purposes is of interest.

## 6. Conclusion

Under conditions essentially determined by modern computer technology and the modern way of administration operations, the speed of information use is of key importance. The earlier striving that it is important for a document to be safely stored is put in the background. Information systems that only store documents but do not ensure flexible use of their content slow down the work of the state body and make it inefficient. This fact is not sufficiently emphasized in reform agendas.

For citizens to receive fast and adequate service from public administration, it is necessary for institutions and state bodies to be supplied with the necessary hardware equipment and software tools and to rationalize their operations in accordance with modern tendencies.

Digital technology for data management and protection enables the synthesis of information and the generation of new information from a set of stored data. Storing data on paper can be safe but does not allow fast and efficient searching, extraction, and synthesizing of new information. The logical and physical mobility of data obtained in computer-supported systems are elements that give new quality, i.e., synthesize new information. This stance is particularly important when observing the mass of documents accumulated in the digitalization process.

Changing methods for storing data requires more rational data management methods. Establishing these new methods requires previously performed detailed analyses of systemic contradictions that recording data on digital media causes over a longer period.

## References

- [1] Conceptualizing digitalization orientation as a strategic posture – How to boost the impact of digital technologies in businesses: A qualitative good practices approach, Alexander Kessler, Hermann Frank, Elena Fuetsch 2025. Elsevier, Technological Forecasting and Social Change Volume 213, April 2025, 123997
- [2] Uredba o kancelarijskom poslovanju organa državne uprave: 21/2020-41, 32/2021-9, 14/2023-23, "Službeni glasnik RS", br. 21 od 6. marta 2020, 32 od 2. aprila 2021, 14 od 20. februara 2023.
- [3] <https://www.mpravde.gov.rs/sr/vest/45934/data-centar-na-meti-hakerskih-napada.php>
- [4] <https://dri.rs/izvestaj/11429>
- [5] <https://www.pecat.co.rs/2025/08/it-kompanije-i-izraelski-medijsko-elektronski-rat>

- [6] Julijana Mirčevski, Nikola B. Popović, Mirjana Andrić and Branko Stanojev. Security of Financial Software to Support Cryptocurrency Trading. Proceedings of the 13th International Conference on Business Information Security (BISEC-2022)Belgrade, Serbia, December 02, 2022.
- [7] <https://apnews.com/article/amazon-east-internet-services-outage-654a12ac9aff0bf4b9dc0e22499d92d7>
- [8] Uputstvo o kancelarijskom poslovanju organa državne uprave: 20/2022-34, "Službeni glasnik RS", broj 20 od 14. februara 2022. 20/2022-34
- [9] Nikola Popović, Srbojub Ranković, Julijana Mirčevski. ANALIZA PROBLEMA DIGITALIZACIJE DOKUMENATA U DRŽAVNOJ UPRAVI, Pregled NCD 7 (2005), 26–33
- [10] Nikola Popović, Julijana Mirčevski. Aktuelni problemi digitalizacije u državnoj administraciji. ZBORNIK RADOVA, LXVI KONFERENCIJA ETRAN, ISBN 978-86-7466-930-3, Novi Pazar 6 - 9. juna 2022.
- [11] <https://www.knime.com/>

[nikola.popovic@mfa.rs](mailto:nikola.popovic@mfa.rs)

[julijana.mir@gmail.com](mailto:julijana.mir@gmail.com)