

REVIEW

by Prof. Tatiana Atanasova, PhD,
Institute of Information and Communication Technologies – Bulgarian Academy of Sciences
on the PhD dissertation for awarding the educational and scientific degree "Doctor" (PhD)
in the doctoral program "Informatics",
professional field 4.6. Informatics and Computer Science
on the topic:

METHODS AND ALGORITHMS FOR PERSONALIZATION AND ADAPTIVITY IN CONTENT MANAGEMENT ENVIRONMENTS,

presented by Emanuela Dimitrova Mitreva
Supervised by: Prof. Desislava Paneva-Marinova, PhD

By Order No. 57/17.06.2026 of the Director of IMI-BAS, I have been appointed as a member of the Scientific Jury regarding the dissertation defense of Emanuela Dimitrova Mitreva on the topic: "Methods and Algorithms for Personalization and Adaptivity in Content Management Environments", for awarding the educational and scientific degree "Doctor" in higher education area 4. Natural Sciences, Mathematics and Informatics; professional field 4.6. Informatics and Computer Science; doctoral program "Informatics".

1. General Characteristics of the Dissertation

The presented dissertation by Emanuela Mitreva consists of 167 pages, including an introduction, five chapters, dissertation contributions, lists of author's publications, citations, reported results, 26 figures, 15 tables, a glossary of terms used, and a bibliography containing 202 titles. The work addresses digital transformation in scientific and cultural heritage and examines the development of digital libraries and the necessity of deploying intelligent management systems for massive datasets. The study focuses on the transition from simple data storage to selective, user-oriented presentation of scientific and cultural heritage through Artificial Intelligence (AI) methods.

2. Content Analysis of the Dissertation and Related Publications

The introductory section outlines the general problem statement. The object, main goal, and subject of the research are defined. The object of the dissertation is defined as the process of content adaptation and personalization in digital libraries using artificial intelligence and machine learning methods and techniques. Five main research objectives are formulated.

The second chapter presents the theoretical foundations and modern concepts related to information content personalization. It notes the evolution of the modern digital library from a static repository into a flexible, AI-driven ecosystem combining secure long-term archiving with intelligent, personalized, and globally accessible user services. In this context, the role of the digital library is redefined—not merely as an archive, but as a dynamic knowledge exchange hub that adapts to user needs. As a high-tech and stable infrastructure, it stimulates academic innovation and ensures the longevity of cultural memory, occupying a central place in today's digital ecosystem. The need for intelligent filtering algorithms in large electronic resource repositories and the use of personalization methods in information retrieval are substantiated.

Table 2, titled "Comparison between Personalization Approaches and Methods," demonstrates the author's ability to systematize, classify, and compare a wide range of technological solutions in content personalization. Table 2 proves that the PhD candidate conducted a thorough preliminary analysis of the technological toolkit.

The dissertation systematizes the evolution of the personalization concept, tracking its trajectory from elementary static tools (surveys) and traditional content filtering to modern, complex, and adaptive approaches (hybrid models, fuzzy sets, support vector machines, and semi-supervised learning).

It is appropriate to emphasize the use of the term "collaborative," as the filtering and recommendation generation are based on shared actions, ratings, or behavior within a given user group. The algorithm characteristics are defined correctly from a scientific and applied perspective (for instance, the "cold start" problem in collaborative filtering and the feature independence constraint in the Naive Bayes classifier are correctly identified).

For personalization purposes, how text information is represented and how its similarity is measured are of key importance. In this context, leading text vectorization methods and their corresponding mathematical proximity measures are summarized.

The theoretical overview in this chapter relies on a detailed study of selected publications released between 2018 and 2025. It encompasses diverse empirical, theoretical, and applied research. The conducted review enables the candidate to convincingly justify the choice of her own methodological approach, upon which the architecture and hybrid model for adaptive personalization are developed in subsequent chapters.

The third chapter focuses on the methodological foundation and architectural organization of the proposed solution for personalized content presentation in a digital library. The software architecture and principles of a hybrid personalized recommendation system are described.

A modern data encoding model is proposed to replace classical mathematical dimensionality reduction, which carries a risk of losing semantic details. The chosen architecture (MiniLM) generates compact dense vectors by design, successfully replacing high-dimensional sparse matrices. This approach allows incremental addition of new documents without requiring re-indexing.

The dissertation presents a model combining "content" and "popularity" factors to address the "cold start" problem, placing emphasis on architectural transparency and result explainability. A unified similarity index is used, integrating behavioral data and content metrics. The primary assumption in

this study is that this hybrid model is more effective in critical scenarios—such as a lack of user history or the addition of new resources—compared to single-source systems.

The proposed model features built-in explainability, allowing clear visibility into the exact contribution of each component regarding thematic similarity and shared "named entities."

The user-centric module implements a hybrid approach to personalized recommendations by integrating semantic proximity, weighted behavioral data, and a global popularity vector, ensuring system robustness when user history is unavailable.

The fourth chapter covers the practical implementation and experimental validation of the proposed software architecture for personalization and the hybrid recommendation model. The architecture decouples heavy preprocessing from real-time serving to guarantee fast, accurate, and explainable recommendations via a content-popularity hybrid model. The recommendation system features asynchronous and interactive layers designed to prevent intensive calculations from affecting response speeds for user queries. Incremental updates are supported—when a document is added or edited, only its connections are recalculated without restarting the entire process.

Document similarity is computed through a weighted linear combination of four independent layers combining text content and information on the named entities present.

Model validation is carried out through a two-stage procedure combining grid search with component ablation to measure individual layer contributions, followed by a multi-criteria analysis of structural cohesion, separability, and ranking stability under incremental data.

The advantages of multi-component evaluations of documents/results/recommendations are demonstrated.

The architecture was verified on a synthetic user-profile dataset and a digital library sample, outlining the "boundaries within which it is correct and useful." This chapter outlines the author's self-acknowledged limitations across several categories: text corpus specificity constraints (primarily Bulgarian media and OCR noise), content model constraints (reliance on pre-trained models not domain-optimized, and subjective parameter choices), behavioral data limitations, as well as technical constraints related to computational load during real-time processing.

The fifth chapter delivers the conclusion, summarizes the main results, and outlines future directions for developing the proposed personalization software architecture in content management environments.

The dissertation's contributions are formulated as scientific and applied, forming a logical and natural conclusion to the research. I accept the classification and formulation of contributions proposed by the candidate.

3. Critical Remarks

In the final version of the dissertation, the candidate fully addressed the critical remarks and recommendations I made during the preliminary discussion.

I appreciate the choice of the experimental corpus comprising digitized issues of multi-topic periodicals, which introduces specific challenges to text processing. The proposed software architecture was verified on digitized Bulgarian periodical press issues. These multi-topic

periodicals (newspapers) possess a specific structure and dynamics, adding extra practical value to the research.

The valuable results achieved do not exclude aspects that point toward future research directions rather than requiring critical remarks. Heavy reliance on synthetically generated data to demonstrate the hybrid algorithm's stability in complex scenarios reveals a certain limitation. For behavioral testing purposes, the candidate used an artificially constructed dataset of 60 text documents organized into three clusters. This warrants further investigation into the algorithm's behavior under real operational conditions. An investigation into the added computational complexity of the semantic layer in a live operational environment with large volumes and high load is also necessary.

The application of modern natural language processing software tools to these specific archival resources also requires data quality reviews on multiple levels.

These remarks do not diminish the value of the research; rather, they point out possible directions for enhancing and expanding future work.

I declare that no plagiarism has been proven according to statutory procedures in the presented dissertation and scientific papers involved in this procedure.

4. Significance of Contributions to Science and Practice

The publications appearing in prestigious international journals indexed in global scientometric databases demonstrate the candidate's capability to conduct scientific research in modern areas of information technology development. A Scopus check for Emanuela Mitreva shows an h-index = 2 with 16 documents, while Web of Science indexes 11 publications with an h-index = 2. The four observed independent citations confirm the initial scientific impact of the published results. I highlight the 2026 publication in the prestigious journal *Applied Sciences* (MDPI) dedicated to a hybrid approach for personalized, intelligent content recommendation in digital libraries, as well as a publication in a Springer series describing a multi-component similarity metric.

According to the Regulations for the Implementation of the Development of Academic Staff in the Republic of Bulgaria Act, these publications earn the candidate a total of 174 points, significantly exceeding the minimum requirement of 30 points under Indicator Γ for a PhD dissertation.

The registered presence and citation count of the research in global scientometric databases (Scopus and Web of Science) serve as an objective measure of international recognition for PhD candidate Emanuela Mitreva in her chosen scientific field.

Emanuela also has documented participation in national projects: "Young Scientists and Postdoctoral Students – 2" and CLADA-BG – National Interdisciplinary Research E-Infrastructure for Resources and Technologies for Bulgarian Language and Cultural Heritage, integrated within the European infrastructures CLARIN and DARIAH (CLaDA-BG).

Quality of the Abstract (Autorating)

The abstract spans 42/39 pages (Bulgarian/English) and satisfies all preparation requirements, accurately reflecting the results, content, and scientific contributions of the dissertation.

5. Conclusion

Emanuela Dimitrova Mitreva's dissertation represents a timely and technology-intensive piece of research. It delivers effective solutions by implementing natural language processing models and named-entity recognition algorithms while building a practically applicable and technically robust software architecture for personalized access in digital libraries. The work fully complies with the requirements for obtaining the educational and scientific degree "Doctor" (PhD) and addresses critical needs of modern society.

The scientific contributions, supported by 5 publications (including 1 single-authored), satisfy all conditions for awarding the degree. The achieved results exceed the minimum requirements set by the Development of Academic Staff in the Republic of Bulgaria Act (ZRASRB), its implementing regulations, and the relevant regulations of BAS and IMI-BAS.

The results achieved provide me with full reason to give a **positive evaluation** and recommend that the esteemed Scientific Jury award the educational and scientific degree "**Doctor**" (PhD) to Mag. **Emanuela Dimitrova Mitreva** in professional field 4.6. "Informatics and Computer Science", doctoral program "Informatics".

August 11, 2026

Signature:

(Prof. Tatiana Atanasova, PhD)