REVIEW

by Prof. Dr. Kiril Ivanov Simov, IICT-BAS

of the materials submitted for participation in a competition

for the academic position of 'Professor'

at IMI-BAS, "SofTIS" Department

In the scientific field: 4. Natural Sciences, Mathematics and Informatics, professional field: 4.6. Informatics and Computer Sciences (Open Science and Innovation in Digital Cultural Heritage)

1. Overview of the candidate's material

By order No. 87/08.09.2026, issued by the Director of the Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences (BAS), I have been appointed as a member of the jury for a competition to fill the academic position of 'Professor'.

Assoc. Prof. Dr. Milena Petrova Dobreva is participating as a candidate in the competition for the position of 'Professor', announced in the State Gazette, Issue 63 of July 10, 2026.

The set of materials submitted by the candidate includes all documents required by the procedure. Associate Professor Dr. Milena Dobreva has met the minimum national requirements under Article 2b, paragraphs 2 and 3, of the Act on the Development of the Academic Staff in the Republic of Bulgaria (ZRASRB).

The assessment of the candidate against the minimum national requirements indicates that she fully meets them for the position of "Professor." I would like to note that the indicators in the sections concerning publications (group G) and citations (group D) significantly exceed the required levels.

The candidate, Assoc. Prof. Dr. Milena Dobreva, has submitted 18 scholarly works, comprising 1 book chapter, 8 articles in peer-reviewed scientific journals, and 9 articles in international conference proceedings. Additionally, a list of textbooks and teaching aids is provided. The first category includes 5 textbooks for the 6th and 9th grades, while the second category consists of teaching aids and a teacher's manual—totaling 4 items. Nine of these publications have been cited in 47 other publications.

For the competition, Assoc. Prof. Dr. Milena Dobreva has included her participation in and leadership of 9 projects. These are categorized into four groups: participation in a national scientific

or educational project (2); leadership of a national scientific or educational project (1); participation in an international scientific or educational project (3); and leadership of a Bulgarian team within an international scientific or educational project (3).

2. Brief biographical details of the candidate

The candidate's professional background is quite impressive. Assoc. Prof. Dr. M. Dobreva earned a Master's degree in Informatics from the Faculty of Mathematics and Informatics at Sofia University "St. Kliment Ohridski" in 1991, and in 1999, she defended her doctoral dissertation in the field of Informatics and Applied Mathematics before the Higher Attestation Commission (HAC). The title of her dissertation was "Modeling Variability in Old Bulgarian Texts." From 1992 to 2001, she held various research positions at the Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences (IMI-BAS), serving as an Associate Professor there from 2001 to 2008. In October 2004, she established the "Computational Humanities with a Center for Digitization" section, which she led until May 2008. From May 2008 to February 2012, she worked as a researcher at the Faculty of Computer and Information Sciences at the University of Strathclyde, Scotland. She subsequently served as an Associate Professor at the University of Malta (2012–2017) and at University College London's branch in Doha, Qatar (2017–2020). In 2020, he returned to Bulgaria and subsequently worked as an R3-rank researcher at the GATE Institute at Sofia University "St. Kliment Ohridski" (2020–2023). Since 2023, he has been working as a Senior Lecturer (Associate Professor) in Information Behaviour at the University of Strathclyde, Scotland. Since December 2023, he has held the position of R4-rank researcher at the IMI-BAS.

The complete list of publications by Assoc. Prof. Dr. Milena Dobreva comprises 161 items. Of these, 151 are in English, 9 in Bulgarian, and one in Russian. One book has been translated from English into Arabic, Bulgarian, Greek, Spanish, Russian, and Serbian. Four publications have an impact factor, and 24 have an SJR; the list also includes one edited volume, two book presentations, one policy analysis, and four summaries of project results. The candidate has received seven distinctions and awards and is a member of four international organizations.

3. Overview of the candidate's activity

Assoc. Prof. Dobreva has presented the scientific contributions of the works submitted for the competition across five thematic areas: A. Computational methods for processing, searching, and accessing digital cultural heritage; B. User models, requirements, and evaluation of digital information systems; C. Collections as data (datafication), semantic connectivity, and open infrastructures for digital cultural heritage; D. Open science, civic engagement, and co-creation of knowledge; E. Education and competencies for data analysis in the field of digital cultural heritage.

These categories are linked to "The 2012 ACM Computing Classification System," which situates them within the broader context of informatics and computer science. For each category, the specific computational problem addressed by the candidate's work is identified, along with the primary type of contribution and the specific papers that substantiate these contributions. Furthermore, these contributions are linked to various projects and initiatives, demonstrating the breadth of their impact.

A. Computer-based methods for processing, searching, and accessing digital cultural heritage

Three publications are related to this area: [1] Beshirov et al. (2025) – Post-OCR text correction for Bulgarian historical documents; [9] Darányi, Wittek & Dobreva (2012) – Using wavelet analysis for text categorization in digital libraries; [16] Beshirov et al. (2022) – DuoSearch: A Novel Search Engine for Bulgarian Historical Documents.

In [1], the challenges associated with standard OCR systems applied to documents from the period 1945–1978 are discussed. To address these issues, a dataset was created to facilitate the post-editing of digitized content. Large language models, retrained for this purpose, are employed to identify problematic instances and perform corrections. The proposed method improves output quality by 16% to 25% compared to previous approaches. In [9], a method for the automatic categorization of text using Support Vector Machines (SVM) is presented. This solution was applied to approximately 6,000 items from Strathprints—indexed using Library of Congress Subject Headings—and its applicability was evaluated. [16] describes a search system for a collection of historical newspapers (pre-1945) that enhances search capabilities by applying content recognition methods to post-OCR data.

B. User models, requirements, and evaluation of digital information systems

The following publications are related to this area: [5] Sykes et al. (2010) – A new focus on end users: Eye-tracking analysis for digital libraries; [6] Dobreva et al. (2010) – Digital natives and specialized digital libraries: A study of Europeana users; [7] Birrell et al. (2010) – Constituencies of use: Representative usage scenarios in international digital library user studies; [8] Birrell et al. (2011) – The DiSCmap project: Digitization of special collections; [8] Birrell et al. (2011) – The DiSCmap project: Digitization of special collections; [12] Hughes et al. (2015) – Assessing and measuring the impact of a digital collection in the humanities: the SPHERE project; [4] Candela, Do-breva et al. (2026) – A Use Case Lens on Digital Cultural Heritage; and [15] Dobreva, Stefanov & Ivanova (2022) – Data Spaces for Cultural Heritage: Insights from GLAM Innovation Labs.

The publications present research on user needs, user behavior, the assessment of content importance, usage scenarios, and related topics. This work is associated with the DiSCmap project, led by Assoc. Prof. Dobreva, and the SPHERE project, in which she participated. She has continued this line of research within the Europeana Network Association. In [5], eye-tracking is employed to evaluate the user interface—assessing both its functionality and user experience—by analyzing how users direct their attention toward various interface elements. In [6], a combined research design is developed and applied to compare the behavior of different user groups. In [7], user scenarios are formalized as a structuring tool for international studies involving users of digital libraries. Building on this experience, [4] presents a systematic methodology for developing usage scenarios within the context of digital cultural heritage. In [8], an approach for prioritizing the digitization of special collections is developed, based on the needs of end-users. In [12], a framework is established for evaluating specialized digital resources from parliamentary archives; the focus shifts from usability to the value and significance of these resources.

C. Collections as data (datafication, collection as data), semantic connectivity, and open infrastructures for digital cultural heritage

The following publications are related to this area: [2] Candela et al. (2025) – A checklist to publish collections as data in GLAM institutions; [4] Candela, Dobreva et al. (2026) – A Use Case Lens on Digital Cultural Heritage; [15] Dobreva, Stefanov & Ivanova (2022) – Data Spaces for Cultural Heritage: Insights from GLAM Innovation Labs; and [17] Candela et al. (2024) – A Systematic Review of Wikidata in GLAM Institutions.

The publications in this group discuss the transition toward data spaces for digital cultural heritage, including the use of intelligent approaches (publication [15]), the linking and integration of data with knowledge networks such as Wikidata [17], and the application of FAIR principles. They also cover the development and evaluation of digital collections, as well as usage scenarios. This area is linked to several initiatives in which Assoc. Prof. Dobreva participates, including the International GLAM Labs Community, the Europeana Network Association, and DARIAH-EU. The DISTILL project (funded as a "Petar Beron" project) and the FOCUS project are also highlighted.

Overall, the work discussed in these publications contributes to various aspects of the infrastructural concept of data spaces—such as intelligent services, the semantic linking of data and collections, the application of FAIR data principles, and the formalization of usage patterns.

D. Open science, civic participation, and co-creation of knowledge

The following publications relate to this area: [13] Dobreva (2016) – Collective knowledge and creativity: The future of citizen science in the humanities; [14] Jennings, Dobreva & Devreni-Koutsouki (2017) – Towards user engagement models for citizen science; [1] Beshirov et al. (2025) – Post-OCR text correction for Bulgarian historical documents; [2] Candela et al. (2025) – A checklist to publish collections as data in GLAM institutions; [4] Candela, Dobreva et al. (2026) – A Use Case Lens on Digital Cultural Heritage; [15] Dobreva, Stefanov & Ivanova (2022) – Data Spaces for Cultural Heritage: Insights from GLAM Innovation Labs; and [17] Candela et al. (2024) – A Systematic Review of Wikidata in GLAM Institutions.

The publications in this group introduce the concept of citizen science within the humanities, its implementation via digital infrastructures and tools, its relationship with cultural institutions, and the facilitation of communication between large groups of volunteers and researchers. It is demonstrated how open access to data, software, and other assets leads to the reuse of created resources and software, which directly links these activities to open science and the FAIR principles. A conceptual framework for citizen science in the humanities is formulated in [13]. In [14], the factors and mechanisms for engaging citizens with digital cultural heritage are systematized, and a model-oriented approach is developed for activities based on the participation of diverse user groups. The remaining publications in this group present examples and aspects related to the FAIR data principles.

This line of work is supported by the practical experience of Assoc. Prof. Dobreva as the leader of the ECHOING project, which focuses on the restoration of cultural heritage through higher-education-driven open innovation.

E. Education and competencies for data analysis in digital cultural heritage

The following publications are related to this area: [3] Dobreva & Candela (2025) – Introducing Collections as Data in Postgraduate and Professional LIS Education; [10] Dobreva & Oliver (2015) – A tale of two countries: Digital curation education in Malta and New Zealand; [11] Dobreva, Angelova & Agre (2015) – Bridging the gap between digital libraries and e-learning; [18] Dobreva (2026) – Digitised cultural heritage as data: Emerging trends in its use in school-level computer science education.

The publications in this group cover research related to education in the field of digital collection creation, specifically within professional training [10]. A conceptual framework has been developed for the closer integration of digital libraries and e-learning. It is demonstrated that the mere existence of large digital collections does not guarantee their use for educational purposes;

rather, resources must be designed and integrated as an active component of the learning process. New competencies required for utilizing "Collections as Data" have been identified—highlighting a need for library and information science professionals skilled in areas such as data cleaning, visualization, and more. Additionally, a model has been developed for the use of digitized cultural heritage within school-level computer science education. The achievements and experience gained through the development of the five thematic areas demonstrate a broad range of knowledge and skills, providing an excellent foundation for research, the training of new specialists, and the attainment of outstanding results consistent with the expectations for a professor. All the submitted publications reflect the personal contribution of Assoc. Prof. Dobreva.

5. Critical remarks and recommendations

I have no critical remarks regarding the candidate.

7. Conclusion

Having reviewed the materials and scholarly works submitted by the candidate for this competition, and having analyzed their significance and contributions, I confidently offer a positive assessment. I recommend that the esteemed scientific jury select Assoc. Prof. Dr. Milena Dobreva for the academic position of "Professor" in Professional Field 4.6: Informatics and Computer Sciences—specializing in "Informatics" (Open Science and Innovation in Digital Cultural Heritage)—to serve the needs of the "SoftTIS" Department at the Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences (IMI-BAS).

02.11.2026 г. Reviewer: Prof. Dr. Kiril Simov