Online Database of 14th-century South Slavonic Manuscripts. Research Results and Perspectives

Marco Scarpa[1, 2][0000-0002-9320-5990], Marta Riparante[1], Desislava Paneva-Marinova[3][0000-0001-5998-687X]

[1] Cyrillo-Methodian Research Center at the Bulgarian Academy of Sciences, 13, Moskovska Str., Sofia, 1000, Bulgaria
[2] University of Messina, Department of Ancient and Modern Civilizations, Santissima Annunziata District, A. Giuffre Str., Messina, 98168, Italy
[3] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 8, G. Bonchev Str., Sofia, 1113, Bulgaria
marco.kmnc@gmail.com, martariparante96@gmail.com, dessi@cc.bas.bg

Abstract. This paper shares the experience of the Kopisti14 project team in achieving a common interdisciplinary scientific result through active collaboration between philologists and computer scientists, each contributing according to their expertise. Some early results of this collaboration are presented. They show how the data collected and processed together offer some interesting insights and raise questions for future research.

Keywords: Digital Humanities, South Slavic Manuscripts, Typologies of Manuscripts.

1 Introduction

In recent decades the use of digital technologies for the traditional humanistic disciplines has seen considerable development and evolution, to the extent that this new area of scholarly activity can be circumscribed under the name of digital humanities. In the case of manuscript studies, their inherently complex and interdisciplinary nature is already a challenge. Palaeography and codicology explore the script and the physical construction of the book, while philology and linguistic analysis address the textual data encoded in its folia. The manuscript is thus a complex object, the elements of which can be analysed in different ways according to different points of view. Therefore, technological tools can not only help but also combine all the different features in an entirely different way, changing the way mediaeval manuscripts are consulted, researched, and disseminated to reveal hitherto unknown and unknowable information. In the field of Slavic codicology and, to a much lesser extent, Slavic paleography, several new projects have been developed in this regard (Scarpa et al., 2022). Our research

project "Fourteenth Century South Slavonic scribes and scriptoria (palaeographical attribution and online repertorium)", also called Kopisti14 project (https://kopisti14.kmnc.bg) also aims to make a relevant contribution resulting from the active collaboration within the scientific team of philological and computer specialists, an important point presented in this paper.

2 IT for Humanities Research. Toward Sustainable Interdisciplinary Collaboration for Common Scientific Result

The recent years' experience offers us the opportunity for some theoretical considerations on the scientific methodology and the collaboration between the disciplines in discussion. From the philologist's point of view, in fact, the contribution of digital sciences is very often perceived in two opposing ways. A first approach sees information technologies as a promising new working tool, which facilitates better and faster performance of certain tasks. Philologists therefore seek to acquire as many computer skills as possible in order to directly apply them in their research. The limitation of this approach is that the humanities scholar faces the risk of always remaining an amateur in computer science, who does not have the necessary means to study in depth, as well as that of staying constantly updated on the evolution of the discipline. A second approach is typical of those who, while intuiting the importance of the digital tool, do not fully understand its mechanisms and try to reproduce traditional ways of working and methodology as much as possible. This very often results in outsourcing to external specialists the creation of tools that he will later learn to deal with, but without understanding their real potential and the new needs that come from using them. Frequently, then, it will be difficult to improve the digital tool, either because the computer scientist does his work and then steps away from the project, or because the philologist is unwilling to further innovations, with the consequence here, too, that digital tools soon age out of use. These two approaches often have the additional consequence that promising projects are created, but they sometimes run aground already in the short term and often remain independent and one-off experiments, without a concrete possibility of development and updating over time.

Conversely, the Kopisti14 research project we are working on brings philologists and computer scientists together in one team, following a methodological approach that has collaboration among specialists as its cornerstone. From the very beginning of the process of creating the research itself, each specialist puts his or her professional and scientific expertise to work, but also practises constant dialogue and collaboration between disciplines, in which there is neither a main nor an auxiliary one, but both are concurrent to the common scientific result. While the participation of computer scientists in meetings on philological topics enables them to become better oriented with respect to the needs presented by philologists (Trifonova, 2022), philologists can better relate to the possibilities offered by computer technology during technical meetings. The work we have accomplished together in recent years shows us how equally positive this collaboration is in the development of solid and prospectively enduring tools as

well as in the opening of new research perspectives. This is also demonstrated by the numerous joint publications that have appeared in both computer science and philology journals (Scarpa et al., 2021; Scarpa et al., 2022; Paneva-Marinova et al., 2022a, Paneva-Marinova et al., 2022b). Moreover, we consider that the continuity perspective of the research we are pursuing can be all the more concrete considering the collaboration between two advanced research centres of the Bulgarian Academy of Sciences, such as the Cyrillo-Methodian Research Center and the Institute of Mathematics and Informatics, together then also with other researchers who are part of other institutions in Bulgaria but also abroad. The intentional inclusion of doctoral students and young scholars, then, makes them key players in the necessary training in this type of research, as well as later in the introduction to targeted professional activities. We are aware that we are not the only ones pursuing this path, as well as that we are still taking our first steps, but we believe it is really important to bring our experience about the methodological approach to scientific research.

In the next section, some early results of this collaboration are presented. They show how the data collected and processed together offer some interesting insights and raise questions for future research. These are some initial processing of the data that were included in the Southern Slavonic manuscripts database of the 14th century (Scarpa et al., 2021).

3 Results

There are currently 1051 South Slavonic manuscripts catalogued in the database, and it is estimated that another 200 codices will be included when the online repertorium is completed. Considering all the manuscripts that were produced during the fourteenth century, it was essential for us to also consider the nearly 80 codices that have been dated from the thirteenth century to the fourteenth century of our interest.

The countries whose archives contain the codices of our interest, which have been catalogued within our database, are nineteen, namely Bosnia and Herzegovina, Bulgaria, Croatia, Germany, Greece, Hungary, Ireland, Vatican, Moldova, Montenegro, Northern Macedonia, Romania, Russia, Serbia, Slovakia, Slovenia, United Kingdom, Ukraine and United States of America. In particular, Mount Athos is the territory that not only holds, along with Serbia, almost 55% of the total number of manuscripts, but where the most fervent manuscript production occurred. In fact, as many as 29 manuscripts have been assigned by scholars to the scriptorium of Hilandar Monastery.

Regarding the main materials used for the production of manuscripts, namely parchment and paper, retrieved data show that paper has a higher percentage (54%) than parchment codices (40%), while the remaining percentage includes mixed texts, i.e., those codices that contain folia of both materials, as a result of later readjustments and manipulations of the original manuscript. Observation of the data shows that while for the first half of the century, manuscripts were written predominantly on parchment (87%), with an almost marginal presence of paper texts, from the second half of the century onward, codices are written mainly on paper, reaching 80% of texts in paper as early as the third quarter of the fourteenth century.

One feature that has been taken into account in our database concerns the orthography of texts. The manuscripts that are currently in the database have the Middle Bulgarian orthography, the Serbian orthography of Raska and the one of Resava, and the Bosnian orthography. Texts with an average Bulgarian orthography account for 25%, while texts with a generic Serbian orthography constitute 70% of the South Slavic manuscript collection, with only 5 manuscripts with a Bosnian orthography. It should be pointed out that about 200 manuscripts have not yet been assigned to a particular orthography because the catalogues consulted did not report this information. One of the future tasks of our research project will be to analyse these codices and categorise them relative to the grammatical norm used. Despite the large gap, the peak of production for both Serbian and Bulgarian orthography manuscripts is to be inscribed in the decade from 1350 to 1360. Considering the context, these data confirm the prosperity of the historical period, on the one hand with the reign of Ivan Aleksander as tsar of Bulgaria from 1331 to 1371 and on the other hand with the reign of Stephen Uroš IV Dušan. Despite his death in 1355, the emperor of the Serbs and Romans was very active in the renovation of churches and monasteries, as well as in manuscript production.

As for those who were responsible for copying the manuscripts and then left traces of their names among the papers, we can claim to know as many as 112, namely 88 Serbian copyists, 23 Bulgarian copyists and one Bosnian copyist. The most famous are Lavrentiy, Gruban, Teodosij or even Danilac Lefteyed, Georgi Radoslav or Ioan. To these and others, more than two manuscripts have been attributed.

As for the typologies that have been previously standardised (Scarpa, et al., 2022), the two macro categories we consider here are those of biblical texts and those of liturgical texts. If for liturgical texts, the total percentages of typologies belonging comes to 38%, with Liturgicon (6.34%), Euchologion (4.48%), Synaxarion (2.24%), Menaion (15.11%), Octoechos (6.90%), Menologion (1.49%), Horologion (1.68%), for biblical texts we have a percentage of almost 48%, with Apostols (11.94%), Four Gospels (33.58%), Full Gospel Lectionary (1.87%) and Prophetologion (0.56%).

4 Conclusions

The aim of the presented project is to reach, effectively use and store in digital format the knowledge about the South Slavonic manuscripts, scriptoria and copyists of the 14th century. The developed Repertorium (https://kopisti.kmnc.bg/bg) successfully manages objects and knowledge by greatly supporting the research work of humanities professionals. In this paper we have shown how through the fruitful collaboration of humanities and computing professionals much higher and contributory scientific results are achieved.

Future project activities will focus on the development of the system and its functionality to support the authors' handwriting recognition, school recognition, *etc.* Solutions based on AI, neural networks, *etc.*, are currently being explored to most effectively implement the requested functionality.

Acknowledgements.

This research is supported by the Bulgarian Scientific Fund under the research project № KP-06-N50/4 30.11.2020 “Fourteenth Century South Slavonic Scribes and Scriptoria (Palaeographical Attribution and Online Repertorium)”.

References

Paneva-Marinova, D., Scarpa, M., Riparante, M., Goynov, M., Luchev, D., Pavlova, L., & Zlatkov, L. (2022). Machine-readable Descriptive Structures to Study Medieval South Slavic Scriptoria and Scribes. *TEM journal, 11*(2), 675-682. https://doi.org/10.18421/TEM112-22.

Scarpa, M., Riparante, M., Paneva-Marinova, D., Goynov, M., & Luchev, D. (2022). Online Database for South Slavonic Manuscripts from the 14th c. *Palaeobulgarica, XLVI*(1), 181-202.

Scarpa, M., Burlacu, C., & Paneva-Marinova, D. (2021). New Digital Life for Fourteenth Century South Slavonic Scriptoria. *Digital Presentation and Preservation of Cultural and Scientific Heritage, 11*, 57-64. https://doi.org/10.55630/dipp.2021.11.5.

Trifonova, I. (2022). South Slavonic Scribes and Scriptoria in the Light of Contemporary Palaeographic Research, *Palaeobulgarica, XLVI*(2), 173-183.

Paneva-Marinova, D., Goynov, M., Stoykov, J., & Pavlov, R. (2022).Studying the Fourteenth Century South Slavic Scribes and Scriptoria through Modern IT Tools. In *Proceedings of the 16th annual International Technology, Education and Development Conference, IATED* (pp. 2170-2176). https://doi.org/10.21125/inted.2022.

Received: April 10, 2023
Reviewed: May 20, 2023
Finally Accepted: June 15, 2023