

El'Manuscript 2018

7th International Conference on Textual Heritage
and Information Technologies

Vienna and Krems, Austria,
14-18 September 2018

Abstracts, Participants, Programme

Vienna: University of Vienna
Krems: Danube University for
Continuing Education,
August 2018

El'Manuscript 2018

7th International Conference on Textual Heritage and
Information Technologies

Vienna and Krems, Austria, 14–18 September 2018

Abstracts, Participants, Programme

Vienna: University of Vienna – Krems: Danube University for
Continuing Education, August 2018

SPONSORS



ADEVA

Jan Kocisek +43 676 5195068
kocisek@utanet.at / @adeva.com



IMPRESSUM

Title

El'Manuscript 2018

7th International Conference on Textual Heritage and Information Technologies. Vienna and Krems, Austria, 14–18 September 2018

Editors

Victor Baranov, Patricia Engel, Jürgen Fuchsbauer, Heinz Miklas

Proof-reading

David J. Birnbaum, Lou Burnard, Heinz Miklas

Photo Credits | Cover

Glagolitic Vienna Folia (Cod. Vind. slav. 136, f. 2r) (front) / Cyrillic Dragotin-palimpsest (Cod. Serd. BNL 880, f. 111r)(back) | Ernst-Georg Hammerschmid

Print

Sofia University Press St. Kliment Ohridski

Institutions

*Department of Slavonic Studies, University of Vienna
Centre of Image and Material Analysis, Vienna, Austria
Danube University for Continuing Education - European Research Centre for Book and Paper
Conservation-Restoration, Krems, Austria
Kalashnikov Izhevsk State Technical University
Society „Textual Heritage“, Russian Federation*

Sponsors

*Faculty of Philological and Cultural Studies, University of Vienna
Danube University for Continuing Education, Krems, Austria
ADEVA, Format werk, CEPA pictures, Verlag Holzhausen
Kremser Humanistische Gesellschaft, ZEISS, Lenzing paper, Österreichische Forschungsgemeinschaft*

© Department of Slavonic Studies, University of Vienna, Austria, 2018

© Kalashnikov Izhevsk State Technical University, Russian Federation, 2018

© Authors, 2018

Verlag Holzhausen - Scientific Publishing

ISBN: 978-3-903207-28-8

Bibliographic information published by the Österreichische Nationalbibliothek and the Deutsche Nationalbibliothek: The ÖNB and the DNB are listing these publications in the Nationalbibliografien; detailed bibliographic data are available on the Internet. For the Österreichische Nationalbibliothek: <http://onb.ac.at>; for the Deutsche Nationalbibliothek: <http://dnb.ddb.de>

HOLZHAUSEN

— Der Verlag —

Maria A. Pilgun

Quantitative text analysis: interpretation of results

Keywords: Text analysis, quantitative analysis, interpretation

In an era of Big Data, artificial intelligence, neural network modeling, and formalized methods of text analysis are helping to transform the concept of textual heritage, and to point toward new ways of archiving digital content. Formalized methods of text analysis provide opportunities to process large text corpora, using encoding, but they do not directly allow the identification of implicit meanings, connotations, and genre specificities. This situation can be compensated by applying informal thematic analysis that focuses on the world of the author, and on his understanding, description, and representation of the situation, which are realized through the construction of meaning. With such an approach, attention focuses on identifying and interpreting the meanings that the author explicitly or implicitly translates, on his interpretations, the reconstruction of positions, types of arguments, etc. This is a mixed method, which combines formalized features (cluster analysis, tonality analysis, thematic modeling, etc.) with informal thematic analysis, which allows the use of triangulation to evaluate the reliability of data and correctly interpret the results.

This report will discuss the strategy and analysis methodology of texts of the 17th–18th centuries (Vesti-Kuranty, Vedomosti) using technology to correctly interpret the results. We also provide a comparative description of the software Jfreq (University of Mannheim, Germany), Yoshikoder (Harvard University, USA), Automap, ORA (CASOS, Carnegie Mellon University, USA, etc.). We conclude that the integration of analysis tools for large collections of structured texts, using sentiment analysis, cluster analysis and psycholinguistic procedures opens promising avenues for interdisciplinary research.

Nadezhda G. Povroznik

Corpus-based historical interpretation: Analyzing the indexes to verbatim records of the early 20th-century State Duma of the Russian Empire

Keywords: Indexes to verbatim records, State Duma, Russian Empire, XML, source structure, activity of Deputies

The significance of indexes to verbatim records as historical sources is that their content is highly structured, with detailed qualitative and quantitative

indicators of the activity of Deputies in the parliament. An analysis of the indexes' structure and their changes allows us to identify opportunities for further study of the source in order to assess its source potential, to provide detailed description of the performance characteristics of Deputies, and to observe changes in the structure of the data submitted.

The Indexes refer to highly structured sources in which the main sections were preserved during all four convocations of the early 20th century. By modeling the Indexes' structure and development using an XML markup scheme, we have been able to explore the source potential of this historical text corpus, and to present detailed descriptions of Deputies' activities. The paper concludes that the indexes have different degrees of completeness of data relating to the personal Deputies' characteristics and their activities, that the structure of markers to verbatim records underwent significant change during the I–IV convenings, but that on the whole it reflects the most significant indicators of the Deputies' activity in the Duma in general.

Acknowledgment

The project was supported by a grant from the Russian Federation Basic Research № 16-06-00569.

Anna V. Ptentsova

Old Russian *pravo* and *krivo*: part-of-speech attribution and grammatical tagging

Keywords: Old Russian, Russian National Corpus, grammatical tagging, part-of-speech ambiguity

In this presentation I discuss the problem of the part-of-speech attribution of two Old Russian words ended in *-o*: *pravo* and *krivo*. This case study is based primarily on the Old Russian subcorpus of the Russian National Corpus (RNC). In addition, I consider tagged texts from the collection of Old Russian manuscripts (<http://www.lrc-lib.ru/>) and historical dictionaries.

Compare:

рекль бо еси вже ꙗ сѣду въ киеву то ꙗ тебе надѣлю <...> нынѣ же ты сѣлъ еси право ли криво ли надѣли же мене

‘You said: if I reign in Kiev, I will vest you <...> Now it is you who is in Kiev; rightly or wrongly, vest me’ (The Kiev Chronicle, 1147).

Here, according to RNC, *pravo* and *krivo* are adverbs, but it's difficult to prove them not to be neuter short adjectives or nouns.



European Research Centre for
Book and Paper
Conservation-Restoration



universität
wien



HOLZHAUSEN

— Der Verlag —

