

UNITED INSTITUTE OF INFORMATICS PROBLEMS  
OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS

**International Scientific Conference  
on the Automatic Processing of Natural-Language  
Electronic Texts “NooJ’2015”**

**NOOJ 2015**

Abstracts

June 11–13, 2015, Minsk, Belarus

Minsk  
UIIP NASB  
2015

УДК 004.91

**International Scientific Conference on the Automatic Processing of Natural-Language Electronic Texts “NooJ’2015”** : Abstracts (11–13 June, 2015, Minsk, Belarus). – Minsk : UIIP NASB, 2015. – 80 p.  
ISBN 978-985-6744-89-4.

This volume contains the abstracts of the International conference “NooJ 2015”. The research presented covers different aspects of natural language processing using NooJ, including formalizing such levels of linguistic phenomena as syllabification, phonemic and prosodic transcription, multiword units and discontinuous expressions, local and structural syntax; transformational syntax and paraphrase generation, semantic analysis and machine translation, etc.

Abstracts are published in the form presented by authors.

У дадзеным зборніку прадстаўлены тэзісы дакладаў Міжнароднай канферэнцыі “NooJ 2015”. Разглядаюцца розныя аспекты апрацоўкі натуральнай мовы з выкарыстаннем лінгвістычнага асяроддзя распрацоўкі NooJ, улічваючы фармалізаваўне такіх напрамкаў лінгвістычнага аналізу як склададзяленне, фанетычная і прасадычная транскрыпцыі, устойлівыя выразы і дыскрэтныя слоўныя канструкцыі, лакальны і структурны сінтаксісы, трансфармацыйны сінтаксіс і перафразаванне, семантычны аналіз і машынны пераклад і г. д.

Тэзісы друкуюцца ў выглядзе, пададзеным аўтарамі.

**Scientific Editors:**

DSc in Engineering B.M. Lobanov,  
PhD in Engineering Yu.S. Hetsevich

**ISBN 978-985-6744-89-4**

© United Institute of Informatics  
Problems of the National Academy  
of Sciences of Belarus, 2015

# PROCESSING OF PUBLICATION REFERENCES IN BELARUSIAN AND RUSSIAN ELECTRONIC TEXTS

D. Dzenisiuk, Yu. Hetsevich

United Institute of Informatics Problems of the NAS of Belarus, Minsk

*e-mail:* d.denissyuk@gmail.com

Entity recognition is a subtask of information extraction that serves to locate and classification of elements in text into pre-defined categories such as the names of persons, organizations, expressions of time, quantities, monetary values, percentages, references etc. [1]. Information which one can get from the references is useful for libraries, publishing houses or research institutions, since it has information about authors or editors, publishers or universities, name and location of the publishing house etc.

The aim of the research is to develop NooJ grammars [2] for processing of references in Belarusian and Russian texts. First step is to develop a set of queries which allow retrieving information about author and co-authors of the publication, name of the publication, location and year, overall number of pages or certain page of the book or the journal. Second step is to combine those queries into syntactic grammars, which can be used on large corpora.

By the beginning of the conference is planned to improve grammars for the Belarusian and Russian languages, and also to make grammar for processing of references designed according to the standard of the Chicago Manual of Style.

## References

1. Get a publication reference! [Electronic resource]. – 2014. – Mode of access : <http://corpus.by/publicationReference>. – Date of access : 10.12.2014.
2. Nooj Manual // NooJ: A Linguistic Development Environment [Electronic resource]. – 2014. – Mode of access : <http://www.nooj4nlp.net>. – Date of access : 10.12.2014.

## CONTENTS

<b>PREFACE</b> .....	5
<b>Ben Ali H., Rhazi A., Aouini M.</b> Translating Arabic Active Sentences into English Passive Sentences using NooJ Platform.....	7
<b>Benet V.</b> Semantic Tags for NooJ Russian Dictionary .....	9
<b>Blanco X.</b> A Hierarchy of Semantic Labels for Spanish Dictionaries .....	10
<b>Chernyshevich M., Stankevitch V.</b> A Hybrid Approach to Extracting and Encoding Disorder Mentions from Clinical Notes.....	12
<b>Collec Clerc V.</b> Mixed Prolog and NooJ Approach in Japanese Benefactive Constructions.....	14
<b>Buono di M.P.</b> Semi-Automatic Indexing and Parsing Information on the Web with NooJ.....	16
<b>Duran M.</b> The Annotation of Compound Suffixation Structure of Quechua Verbs.....	18
<b>Dzenisiuk D., Hetsevich Yu.</b> Processing of Publication References in Belarusian and Russian Electronic Texts.....	20
<b>Ghezaiel N., Haddar K.</b> Study and Resolution of Arabic Lexical Ambiguity through the Transduction on Text Automaton .....	21
<b>Hetsevich Yu., Borodina J.</b> Using NooJ for the Processing of Satellite Data .....	23
<b>Hetsevich Yu., Okrut T., Lobanov B.</b> Grammars for the Sentence into Phrase Segmentation: Punctuation Level.....	25
<b>Hiuntar A., Zahariev V.</b> Grammars for Making Written Orthographic Words from Transcribed Spoken Language .....	26
<b>Kaigorodova L., Hetsevich Yu., Nikalaenka K., Prakapovich R., Gerasuto S., Sychou U.</b> Language Modelling for Robots-Human Interaction .....	28
<b>Kirova M.</b> Translating Spacial and Temporal Deixis in Near Languages: A Comparative Classification Approach with NooJ.....	30