

**Europass  
Curriculum Vitae**



**Personal information**

Surname(s) / First name(s) **Dr. NAKOV / Preslav, Ph.D.**

Address(es) 42a Hristo Belchev St, Sofia 1000, Bulgaria

Telephone(s) +359 (2) 981-6075 Mobile: +359 (885) 449-810

E-mail nakov@cs.berkeley.edu; nakov@lml.bas.bg; preslav.nakov@gmail.com

Nationality Bulgarian

Date of birth January 26, 1977

Gender Male

**Work experience**

<p>Dates</p> <p>Occupation or position held</p> <p>Main activities and responsibilities</p> <p>Name and address of employer</p> <p>Type of business or sector</p>	<p>January'2008 – now</p> <p>Linguistic Modeling Department, (<a href="http://lml.bas.bg">http://lml.bas.bg</a>) Institute for Parallel Processing, Bulgarian Academy of Sciences 25A Acad. G. Bonchev St, 1113 Sofia, Bulgaria Research Institution</p>
<p>Dates</p> <p>Occupation or position held</p> <p>Main activities and responsibilities</p> <p>Name and address of employer</p> <p>Type of business or sector</p>	<p>January'2008 – now</p> <p>Consultant</p> <p>Consulting on NLP-related issues</p> <p>OntoText Lab, 135 Tsarigradsko Chaussee, Sofia 1784, Bulgaria (<a href="http://www.ontotext.com/">http://www.ontotext.com/</a>) Information and Communication Technologies</p>
<p>Dates</p> <p>Occupation or position held</p> <p>Main activities and responsibilities</p> <p>Name and address of employer</p> <p>Type of business or sector</p>	<p>January'2008 – now</p> <p>Part-time lecturer</p> <p>Founded a course on <i>Practical Computational Linguistics: Statistical Approach</i> (<a href="http://statnlp.devbg.org/">http://statnlp.devbg.org/</a>) Sofia University, Faculty of Mathematics and Informatics 5 James Bourchier St, 1113 Sofia, Bulgaria University</p>
<p>Dates</p> <p>Occupation or position held</p> <p>Main activities and responsibilities</p> <p>Name and address of employer</p> <p>Type of business or sector</p>	<p>May'2003 - December'2007</p> <p>Research Assistant</p> <p><u>BioText project</u>: Design and implementation of statistical approaches to natural language processing, supporting advanced and particular search needs of bioscience researchers. <a href="http://biotext.berkeley.edu">http://biotext.berkeley.edu</a> School of Information, University of California Berkeley, 102 South Hall, Berkeley, CA 94720-4600, USA University, Research Institution</p>

Dates	June, 2005 – August, 2005
Occupation or position held	Internship
Main activities and responsibilities	Research on trying to improve word alignments using additional languages
Name and address of employer	Information Sciences Institute, University of Southern California, 4676 Admiralty Way, Marina del Rey, CA 90292-6601, USA ( <a href="http://www.isi.edu">http://www.isi.edu</a> )
Type of business or sector	University, Research Institution
Dates	February, 2000 – June, 2002
Occupation or position held	Programmer Analyst
Main activities and responsibilities	Design and implementation of a Natural Language Processing (NLP) engine: Information Extraction (IE) and Information Retrieval (IR).
Name and address of employer	Rila Solutions; 27 Acad. G. Bonchev Street, 1113 Sofia, Bulgaria ( <a href="http://www.rila.com/">http://www.rila.com/</a> )
Type of business or sector	Information and Communication Technologies
Dates	November, 1999 – January, 2000
Occupation or position held	Consultant
Main activities and responsibilities	Consulting on questions related to information retrieval and question answering.
Name and address of employer	Rila Software; 27 Acad. G. Bonchev Street, 1113 Sofia, Bulgaria ( <a href="http://www.rila.com/">http://www.rila.com/</a> )
Type of business or sector	Information and Communication Technologies
Dates	December, 1996 – January, 2000
Occupation or position held	RDBMS developer
Main activities and responsibilities	Design and development of RDBMS systems with Oracle.
Name and address of employer	ComSoft Ltd., 5 Pirotska Street, 1004, Sofia, Bulgaria ( <a href="http://www.comsoft.bg">http://www.comsoft.bg</a> )
Type of business or sector	Information and Communication Technologies

### Education and training

Dates	August'2002 – December'2007
Title of qualification awarded	Ph.D.
Principal subjects/occupational skills covered	<p><u>Advisor:</u> Marti Hearst</p> <p><u>Course Work:</u> Lexical Semantics (Lynn Nichols), Semantics (Line Mikkelsen), Syntax and Semantics I &amp; II (Line Mikkelsen), Applied Natural Language Processing (Marti Hearst), Statistical Natural Language Processing (Dan Klein), An AI Approach to NLP (Robert Wilensky), Statistical Learning Theory I &amp; II (Michael Jordan), Neural Computation and Language (Jerome Feldman), Combinatorial Algorithms and Data Structures (Christos Papadimitriou), Reinforcement Learning (Stuart Russell), Computational Biology for Computer Scientists (Gene Myers), Advanced Topics in Computer Systems (Eric Brewer), Network Flows and Graphs (Dorit Hochbaum)</p> <p><u>Ph.D. thesis:</u> <i>Using the Web as an Implicit Training Set: Application to Noun Compound Syntax and Semantics</i> (<a href="http://www.eecs.berkeley.edu/Pubs/TechRpts/2007/EECS-2007-173.html">http://www.eecs.berkeley.edu/Pubs/TechRpts/2007/EECS-2007-173.html</a>).</p>
Name and type of organisation providing education and training	University of California at Berkeley, USA Department of Electrical Engineering and Computer Science, Computer Science division
Dates	1996 - 2001
Title of qualification awarded	M.Sc. degree
Principal subjects/occupational skills covered	<ul style="list-style-type: none"> <li>◆ First specialty: Informatics (Computer Science), <ul style="list-style-type: none"> <li>● Major 1: Artificial Intelligence</li> <li>● Major 2: Information and Communication Technologies</li> </ul> </li> <li>◆ Second specialty: Mathematics and Informatics (pedagogical).</li> </ul> <p>§ GPA 6.0 (Bulgarian grading system: 2-6 scale, 6 is the best grade)</p> <p>§ Diploma work: <i>Recognition and Morphological Classification of Unknown Words for German</i></p>
Name and type of organization providing education and training	Sofia University "St. Kliment Ohridski", Sofia, Bulgaria Dept. of Mathematics and Informatics (FMI)
Dates	1991-1996
Title of qualification awarded	Secondary School Diploma
Principal subjects/occupational skills covered	Specialty: Mathematics
Name and type of organization providing education and training	Secondary School of Mathematics and Natural Sciences "Vassil Drumev", Veliko Turnovo, Bulgaria

Dates 1991-1996  
 Title of qualification awarded Secondary School Diploma (equivalency exams)  
 Principal subjects/occupational skills covered Specialty: French  
 Name and type of organization providing education and training Secondary School for Foreign Languages "Prof. Dr. Assen Zlatarov", Veliko Turnovo, Bulgaria

Dates 1991-1996  
 Title of qualification awarded Secondary School Diploma (equivalency exams)  
 Principal subjects/occupational skills covered Specialty: Russian  
 Name and type of organization providing education and training Secondary School for Foreign Languages "Prof. Dr. Assen Zlatarov", Veliko Turnovo, Bulgaria

**Personal skills and competences**

Mother tongue(s) **Bulgarian**

Other language(s)  
 Self-assessment European level (\*)

**English**  
**Russian**  
**French**  
**Spanish**  
**Italian**  
**Portuguese**  
**German**  
**Turkish**

Understanding				Speaking				Writing	
Listening		Reading		Spoken interaction		Spoken production			
C2	Proficient	C2	Proficient	C2	Proficient	C2	Proficient	C2	Proficient
C2	Proficient	C2	Proficient	C2	Proficient	C2	Proficient	C2	Proficient
C2	Proficient	C2	Proficient	C1	Proficient	C2	Proficient	C1	Proficient
C2	Proficient	C1	Proficient	B2	Independent	C1	Proficient	B2	Independent
C1	Proficient	C1	Proficient	B2	Independent	C1	Proficient	B2	Independent
B2	Independent	C1	Proficient	B2	Independent	B2	Independent	B2	Independent
B1	Independent	B2	Independent	B1	Independent	B1	Independent	B2	Independent
A1	Basic	A1	Basic	A1	Basic	A1	Basic	A1	Basic

Social skills and competences - I am a good team worker.  
 - After five years at the University of California at Berkeley, I adapt easily to international environments.  
 - I speak eight languages, which allows me to communicate with many colleagues in their mother tongue.

Organisational skills and competences Mentor: Graduate student mentor for a SUPERB-IT student, UC Berkeley, EECS dept., CS division, Summer 2006.

Coach: Coach of the Sofia University team at the ACM, International Collegiate Programming Contest, South-Eastern Europe. October 20-22, 2001. Bucharest, Romania.

Computer skills and competences Programming languages: Java, C/C++, Perl, Lisp, Prolog, Pascal, Basic, SQL, etc.  
 Operating Systems: Linux/Unix, Windows

Driver's license Category B

## Additional information

### Awards

#### Special awards:

- Bulgarian President's John Atanasoff Award for achievements in the development of the information society, December 2003
- Sofia University Rector's award Best student for the academic year 1997/1998
- National Diploma for achievements in Computer Science and Philosophy by the Ministry of Education, Science and Technologies, 1996 (there were total 54: in sciences and arts)
- Award by the Foundation for helping the Bulgarian high education Sofia - Frankfurt-am-Main for high education results and distinguished qualifications and knowledge for the academic year 2000/2001
- Award by the Eureka foundation for the academic 1999/2000 and 2000/2001 years

#### Other awards:

- ACM, International Collegiate Programming Contest, South-Eastern Europe, Bucharest, Romania: 2nd place (2000) and 3rd place (1999)
- III Balkan Olympiad in Informatics: bronze medal (1995)
- II International Olympiad in Philosophy: 10th place (1995)
- Spring Conference of the Bulgarian Mathematicians' Union (student section), 1st award (1995)
- Contest in Informatics of PC Magazine/Bulgaria: 1st award plus PC (1997), 2nd award (1999)
- Contest in Informatics of ComputerNews newspaper: 2nd (1996) and 2nd (1999) awards
- Contest in Informatics of Computer Magazine: 1st (1999, 1997), 2nd (1998, 1996, 1995), 3rd (1993) and 5th (1994) awards
- Winter Mathematical Holidays (Informatics Section): 3rd (1995) and 7th (1996) awards
- Interuniversity Programming Competition: 5th (1999) and 5th (2000) awards
- French Language Kangourou Mathematical Competition: 7th place in Bulgaria (1996)

### Scholarships and Fellowships

- Fulbright grant — 10 months, Ph.D. in USA, starting Fall 2002
- UC Berkeley Graduate Fellowship, academic year 2002/2003

### Professional Membership

- Association of Computational Linguistics (ACL), since 2003
- Association for Computing Machinery (ACM), since 2000
- Bulgarian Artificial Intelligence Association (BAIA), since 2008
- Bulgarian Union of Automation and Informatics (SAI), since 2004
- Internet Society Bulgaria (ISOC), since 2003
- Bulgarian Association of Software Developers (BSAD), since 2004 - honor member
- Association for Information Security (ISECA), since 2004

### Research Projects

- **@ UC Berkeley**
  - ✓ BioText: The project combines a flexible, efficient, platform-independent database system infrastructure with statistical approaches to natural language processing, in order to support the advanced and particular search needs of bioscience researchers. (<http://biotext.berkeley.edu>)
- **past**
  - ✓ Joint Research Project with the Hamburg University: Recognition and Morphological Classification of Unknown Words for German. Partly developed during my visit to the Natural Language Systems Division of the Computer Science Department, Hamburg University (Prof. Walther von Hahn).
  - ✓ INCO Copernicus'98 Joint Research Project #977074. LARFLAST "Learning Foreign Language Scientific Terminology". Funded by the EC, November 1998 – October 2001
  - ✓ § Latent Semantic Analysis module design and implementation in C++, 2001.

## Teaching Experience

- Lecturer
  - ✓ *Practical Computational Linguistics: Statistical Approach*, Sofia University, dept. FMI (Fall 2008)  
<http://statnlp.devbg.org/>
  - ✓ *Design and Analysis of Computer Algorithms*, Sofia University, dept. FMI (Fall 2001)
  - ✓ *Data Modeling and Database Design (Oracle course)*, Sofia University, dept. FMI (Spring 2001, Fall 2000, Spring 1999)
  - ✓ *Data Warehousing. Fundamentals (Oracle course)*, Sofia University, dept. FMI (Fall 2000)
  - ✓ *Computer Programming and Algorithms* (specialized course for secondary school students with special interests), Bulgarian Academy of Sciences (Fall 2000)
- Teaching Assistant
  - ✓ *Applied Natural Language Processing (SIMS 256)*, UC Berkeley, SIMS dept., Fall 2006, Prof. Marti Hearst (I gave 1 lecture)
  - ✓ *Applied Natural Language Processing (SIMS 290-2)*, UC Berkeley, SIMS dept., Fall 2004, Prof. Marti Hearst (I gave 2 lectures)
  - ✓ *Introduction to Artificial Intelligence (CS 188)*, UC Berkeley, EECS dept., CS division, Spring 2004, Prof. Stuart Russell (I gave 1 lecture)
  - ✓ *Languages for System Programming*, Sofia University, dept. FMI (Fall 2001), Prof. St. Bachvarov
  - ✓ *Object-oriented Programming*, Sofia University, dept. FMI (Spring 2001), Prof. St. Bachvarov
  - ✓ *Data Structures and Programming*, Sofia University, dept. FMI (Spring 2000, Spring 1999), Prof. St. Bachvarov
  - ✓ *Introduction to Programming*, Sofia University, dept. FMI (Fall 1999), Prof. St. Bachvarov
  - ✓ *Algebra*, University College London (Spring 1999), Dr. R. Hill (one month, during my visit to UCL)
  - ✓ *Data Structures*, Sofia University, dept. FMI (Fall 1998), Prof. D. Shishkov
- Teacher (High School)
  - ✓ *Informatics*, High School "Frédéric Joliot-Curie", Sofia, Bulgaria, pedagogical practice required by the Sofia University, dept. FMI for my second specialty "Mathematics and Informatics", Fall 2001: 2 months.
  - ✓ *Mathematics*, High School "Frédéric Joliot-Curie", Sofia, Bulgaria, pedagogical practice required by the Sofia University, dept. FMI for my second specialty "Mathematics and Informatics", Fall 2001: 2 months.

## Programme committees

- EMNLP'2008: Conference on Empirical Methods on Natural Language Processing, 2008
- AAAI'2008: Twenty-Third AAAI Conference on Artificial Intelligence, 2008
- LREC'2008: Sixth International Conference on Language Resources and Evaluation, 2008
- MWE'2008: Towards a Shared Task for Multiword Expressions
- LDALSK'2008: Language Diversity and the Acquisition of Linguistic-Semantic Knowledge, 2008
- Open PhD Workshop on Technology-Enhanced Learning and Semantics, Software and Services, in conjunction with AIMSAS'2008
- Machine Learning and Intelligent Systems, in conjunction with the Second International Symposium Recent Developments in Cryptography and Information Security, 2008
- RANLP'2007: Recent Advances in Natural Language Processing, 2007
- SemEval'2007: 4th International Workshop on Semantic Evaluations, June 23-24, Prague, Czech Republic, 2007 (task organizer)
- ACL workshop: A Broader Perspective on Multiword Expressions, June 28, Prague, Czech Republic, 2007
- Recent Advances in Natural Language Processing (RANLP'2005), Borovets, Bulgaria, 2005
- HLT/NAACL'2004, Student session, Boston, MA, USA, 2004

## List of publications

### Books

- Nakov P., P. Dobrikov. Programming = ++Algorithms; Third edition. TopTeam Co Publishing. Sofia, 2005. (706 pages) \*
- Nakov P. Fundamentals of the Computer Algorithms. Third revised edition. TopTeam Co Publications. Sofia. 2001. (400 pages)
- Nakov S. & al. Programming for .NET Framework. Vol. 1. Faber. 2005. Sofia. (801 pages).

\* Some university courses using this textbook:

- ✓ *Design and Analysis of Computer Algorithms*, Sofia University "St. Kliment Ohridski"
- ✓ *Advanced programming*, New Bulgarian University, Sofia
- ✓ *Algorithm Synthesis and Analysis*, Technical University of Varna
- ✓ *Algorithms and Programs*, University of Plovdiv "Paisij Hilendarski"

### Papers

2008

1. Preslav Nakov and Marti Hearst. Solving Relational Similarity Problems Using the Web as a Corpus. In Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL'08). Columbus, OH, USA. 2008.
2. Preslav Nakov. Improved Statistical Machine Translation Using Monolingual Paraphrases. In Proceedings of the European Conference on Artificial Intelligence (ECAI'08), Patras, Greece, 2008.
3. Preslav Nakov. Noun Compound Interpretation Using Paraphrasing Verbs: Feasibility Study. In Proceedings of the 13<sup>th</sup> International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA'08), Varna, Bulgaria, 2008.
4. Preslav Nakov. Paraphrasing Verbs for Noun Compound Interpretation. In Proceedings of the Workshop on Multiword Expressions (MWE'08), in conjunction with the Language Resources and Evaluation conference, Marrakech, Morocco, 2008.
5. Preslav Nakov. Improving English-Spanish Statistical Machine Translation: Experiments with Domain Adaptation, Sentence-Level Paraphrasing, Tokenization, and Recasing. In Proceedings of the Third Workshop on Statistical Machine Translation (WMT'08), in conjunction with ACL'2008.
6. Larry Smith, Lorraine K. Tanabe, Rie Johnson nee Ando, Cheng-Ju Juo, I-Fang Chung, Chun-Nan Hsu, Yu-Shi Lin, Roman Klinger, Christoph M. Friedrich, Kuzman Ganchev, Manabu Torii, Hongfang Liu, Barry Haddow, Craig A. Struble, Richard J. Pavinelli, Andreas Vlachos, William A. Baumgartner Jr., Lawrence Hunter, Bob Carpenter, Richard Tzong-Han Tsai, Hong jie Dai, Feng Liu, Yifei Chen, Chengjie Sun, Sophia Katrenko, Pieter Adriaans, Christian Blaschke, Rafel Torres Perez, Mariana Neves, Preslav Nakov, Anna Divoli, Manuel Mana, Jacinto Mata-Vazquez, and W. John Wilbur. Overview of BioCreative II Gene Mention Recognition. *Journal of Genome Biology*, 2008.
7. Roxana Girju, Preslav Nakov, Vivi Nastase, Stan Szpakowicz, Peter Turney, Deniz Yuret. Classification of Semantic Relations between Nominals. *Journal of Language Resources and Evaluation*, Special Issue on "Computational Semantic Analysis of Language: SemEval-2007 and Beyond", 2008.

- 2007
1. **Ph.D. thesis:** Using the Web as an Implicit Training Set: Application to Noun Compound Syntax and Semantics., Nakov, Preslav Ivanov. EECS Department, University of California, Berkeley. Technical Report No. UCB/EECS-2007-173. December 20, 2007.
  2. Marti A. Hearst, Anna Divoli, Harendra Guturu, Alex Ksikes, Preslav Nakov, Michael A. Wooldridge, and Jerry Ye, BioText Search Engine: beyond abstract search. Bioinformatics Advance Access published on June 1, 2007, DOI 10.1093/bioinformatics/btm301.
  3. Girju R., Nakov P., Nastase V., Szpakowicz S., Turney P., Yuret D. SemEval-2007 Task 04: Classification of Semantic Relations between Nominals. In Proceedings of SemEval-2007 Workshop co-located with ACL-2007, Prague, June 23-24, 2007.
  4. Nakov P. and Hearst M. UCB: System Description for SemEval Task #4. In Proceedings of SemEval-2007 Workshop co-located with ACL-2007, Prague, June 23-24, 2007.
  5. Nakov P. and Hearst M. UCB System Description for the WMT 2007 Shared Task. In Proceedings of Second Workshop on Statistical Machine Translation (WMT'2007) co-located with ACL-2007, pp.212-215, Prague, June 23, 2007.
  6. Nakov, P., and Divoli A. BioText Report for the Second BioCreative Challenge., In Proceedings of the Second BioCreative Challenge Evaluation Workshop: 23-25 April 2007; Madrid; pp. 297-306.
  7. Nakov, P., Nakov S., and Paskaleva E. Improved Word Alignments Using the Web as a Corpus. In Proceedings of RANLP'2007, pp.400-405, Borovetz, Bulgaria, September 27-29, 2007.
  8. Nakov, S., Nakov P., and Paskaleva E. Cognate or False Friend? Ask the Web! In Proceedings of the Workshop on Acquisition and Management of Multilingual Lexicons, held in conjunction with RANLP'2007, pp.55-62, Borovetz, Bulgaria, September 30, 2007.
  9. Nakov P., Pacovski V., and Paskaleva E. Extracting Translation Lexicons from Bilingual Corpora: Application to South-Slavonic Languages. In Proceedings of the Workshop on Common Natural Language Processing Paradigm For Balkan Languages, held in conjunction with RANLP'2007, pp. 23-31, Borovetz, Bulgaria, September 26, 2007.
- 2006
1. Nakov, P., and Hearst, M. Using Verbs to Characterize Noun-Noun Relations. In Proceedings of the 12<sup>th</sup> International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA'2006), Bulgaria, September 2006.
  2. Divoli A., Hearst A., Nakov P., Schwartz A., Ksikes A. BioText Team Report for the TREC 2006 Genomics Track. In Proceedings of Text Retrieval Conference (TREC'2006), Gaithersburg, MD, 2006.
  3. Preslav Nakov, Veno Pacovski. Acquiring "False Friends" from Parallel Corpora: Application to South Slavonic Languages. In Slavcheva, M., Angelova, G., Simov, K. (Eds.) Readings in Multilinguality, Selected Papers from Young Researchers in BIS-21++. INCOMA Ltd., Shoumen, Bulgaria. 2006.
- 2005
1. Nakov, P., and Hearst, M. Search Engine Statistics Beyond the n-gram: Application to Noun Compound Bracketing. In Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005), Ann Arbor, MI, June 2005.
  2. Nakov, P., and Hearst, M. Using the Web as an Implicit Training Set: Application to Structural Ambiguity Resolution. In Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT-NAACL'05), Vancouver, Canada, 2005.
  3. Nakov, P., and Hearst, M. A Study of Using Search Engine Page Hits as a Proxy for n-gram Frequencies. In Proceedings of the Recent Advances in Natural Language Processing (RANLP'05), Borovets, Bulgaria, 2005.
  4. Nakov, P., Schwartz, A., Wolf, B, and Hearst, M. Scaling Up BioNLP: Application of a Text Annotation Architecture to Noun Compound Bracketing. In ACL/ISMB BioLINK SIG: Linking Literature, Information and Knowledge for Biology (BioLINK'2005), Detroit, MI, June 2005.
  5. Nakov, P., Schwartz, A., Wolf, B, and Hearst, M. Supporting Annotation Layers for Natural Language Processing. In the ACL 2005 Poster/Demo Track Ann Arbor, MI, June 2005.
  6. Nakov, P. and Paskaleva E. Dictionary, statistical and Web knowledge in shallow parsing procedures for inflectional languages, In Proceedings of International Workshop on Language and Speech Infrastructure for Information Access in the Balkan Countries, RANLP'97, Borovets, Bulgaria, 2005, pp. 39-46.

- 2004
1. Preslav Nakov, Ariel Schwartz, Marti Hearst. Citances: Citation Sentences for Semantic Analysis of Bioscience Text. Workshop on Search and Discovery in Bioinformatics at SIGIR'04, Sheffield, UK, July 2004.
  2. Preslav Nakov, Ariel Schwartz, Emilia Stoica, Marti Hearst. BioText Team Report for the TREC 2004 Genomics Track. In Proceedings of the Text REtrieval Conference (TREC'04), Gaithersburg, MD, USA, 2004.
  3. Archana Ganapathi, Preslav Nakov, Ariel Schwartz, Marti Hearst. Supporting Annotation Layers for Natural Language Processing, manuscript, February 2004.
  4. Preslav Nakov, Elena Paskaleva. Robust Ending Guessing Rules with Application to Slavonic Languages. In Proceedings of the 3rd workshop on ROBust Methods in Analysis of Natural Language Data (ROMAND), an International Workshop in Association with COLING'04, pp. 76-85, Geneva, August 29, 2004.
  5. Galia Angelova, Albeno Strupchanska, Ognyan Kalaydijev, Milena Yankova, Svetla Boytcheva, Irena Vitanova, Preslav Nakov. Towards Deeper Understanding and Personalisation in CALL. In Proceedings of "eLearning for Computational Linguistics and Computational Linguistics for eLearning", an International Workshop in Association with COLING'04. pp. 45-52, Geneva, August 28, 2004.
  6. Preslav Nakov, Panayot Dobrikov. Non-Parametric SPAM Filtering based on kNN and LSA. In Proceedings of the 33th National Spring Conference of the Bulgarian Mathematicians Union, Borovets, Bulgaria, April 1-4, 2004.
  7. Preslav Nakov, Elena Valchanova, Galia Angelova. Towards deeper understanding of the latent semantic analysis performance. In Recent Advances in Natural Language Processing III (RANLP'03), Nicolov, Nicolas, Kalina Bontcheva, Galia Angelova and Ruslan Mitkov (eds.), pp. 297-306, Johns Benjamins, 2004.
  8. Preslav Nakov, Yury Bonev, Galia Angelova, Evelyn Gius, Walther von Hahn. Guessing morphological classes of unknown German nouns. In Recent Advances in Natural Language Processing III (RANLP'03), Nicolov, Nicolas, Kalina Bontcheva, Galia Angelova and Ruslan Mitkov (eds.), pp. 347-356, Johns Benjamins, 2004.
  9. Ivanka Atanassova, Preslav Nakov, Svetlin Nakov. FineArtsDict – a program for creation and management of bilingual dictionaries with glosses. In Proceedings of the VIth International Symposium "Projects on Comparative Studies of Russian and Other Languages". pp. 14-20, Belgrade, Serbia and Montenegro, June 1-4, 2004. (in Russian).



- 2003
1. Nakov P., M. Hearst. Category-based Pseudowords. In Proceedings of HLT-NAACL'03 (Companion Volume), pp. 67-69, Edmonton, Canada, May 2003.
  2. Bhalotia G., P. Nakov, A. Schwartz, M. Hearst. BioText Team Report for the TREC 2003 Genomics Track. In Proceedings of the Text REtrieval Conference (TREC'03), pp. 612-621, Gaithersburg, MD, USA, 2003.
  3. Nakov P., E. Valchanova, G. Angelova. Towards Deeper Understanding of LSA Performance. In Proceedings of Recent Advances in Natural Language Processing (RANLP'03). pp. 311-318. Borovetz, Bulgaria, September 10-12, 2003.
  4. Nakov P., Bonev Y., G. Angelova, E. Gius, W. von Hahn. Guessing Morphological Classes of Unknown German Nouns. In Proceedings of Recent Advances in Natural Language Processing (RANLP'03). pp. 319-326. Borovetz, Bulgaria, September 10-12, 2003.
  5. Nakov P. BulStem: Design and Evaluation of Inflectional Stemmer for Bulgarian. In Proc. Workshop on Balkan Language Resources and Tools (1st Balkan Conference in Informatics), Thessaloniki, Greece, November 21, 2003.
  6. Atanassova I, Nakov S., Nakov P. ArtsSemNet: from Bilingual Dictionary to Bilingual Semantic Network. In Proceedings of Workshop on Balkan Language Resources and Tools (1st Balkan Conference in Informatics), Thessaloniki, Greece, November 21, 2003.
  7. Nakov P. Building an Inflectional Stemmer for Bulgarian. In Proc. 4th International Conference on Computer Systems and Technologies (ICCST'03), pp. 419-424, Sofia, June 19-20, 2003.
  8. Atanassova I., Nakov P., Nakov S., ArtsSemNet: A Bilingual Semantic Network for Bulgarian and Russian Fine Arts Terminology, In Proceedings of Naval Scientific Forum, vol. 3 (Humanities. Applied Linguistics and Foreign Language Teaching), pp. 222-229. Varna, Bulgaria. 2003. (in Russian)
  9. Dobrikov P, P. Nakov. The Architecture of a Corporate Information and News Engine. In Proceedings of the 4th International Conference on Computer Systems and Technologies (ICCST'03), pp. 425-430, Sofia, June 19-20, 2003.
- 2002
1. Angelova G., S. Boytcheva, O. Kalaydjiev, S. Trausan-Matu, P. Nakov and A. Strupchanska. Adaptivity in Web-Based CALL. In Proceedings of the 15th European Conference on Artificial Intelligence (ECAI'02), pp. 445-449. Lyon, France. July 21-26 2002.
  2. Nakov, P., G. Angelova, W. von Hahn. Automatic Recognition and Morphological Classification of Unknown German Nouns. FBI-HH-B-243/02, Bericht 243, Fachbereich Informatik, Universitaet Hamburg, September 2002.
  3. Nakov P., S. Terzieva. Latent Semantic Analysis for Notional Structures Investigation. In Proceedings of the Annual Congress of the European Society for Philosophy and Psychology (ESPP'02). Lyon, France, July 10-13, 2002.
  4. Nakov P. MorphoClass - Recognition and Morphological Classification of Unknown Words for German. In Proceedings of Semantic Authoring, Annotation & Knowledge Markup (SAAKM'02), Workshop at ECAI 2002. Lyon, France. July 22-26, 2002.
  5. Nakov P. Ending-Guessing Rules for Morphological Classification of German Nouns. In Proceedings of CompSysTech'02. Sofia, Bulgaria. pp. II.10-1—II.10-6. June 20-21, 2002.
  6. Atanassova I., Nakov P, Nakov S. Semantic Technique for Automatic Hyponym Chains Extraction from Terminological Dictionaries. In Proceedings of the VIIIth International Symposium (MAPRIAL'02). Theoretical and methodological problems of Russian as a foreign language at the beginning of the XXI century. Veliko Turnovo, Bulgaria. pp. 309-314. April 4-5, 2002. (in Russian)
  7. Atanassova I., Nakov P, Nakov S. Information Technologies Helping the Linguist-Explorer. In Proceedings of the VIIIth International Symposium (MAPRIAL'02). Theoretical and Methodological problems of Russian as a foreign language at the beginning of the XXI century. Veliko Turnovo, Bulgaria. pp. 304-309. Veliko Turnovo, Bulgaria. April 4-5, 2002. (in Russian)
  8. Jeliakov J., P. Nakov. Extended Boolean Operations in Latent Semantic Indexing Search. In Proceedings of the Spring Conference of Bulgarian Mathematicians Union. Bulgaria. April, 2002.

- 2001
1. **M.Sc.thesis:** Nakov P. Recognition and Morphological Classification of Unknown Words for German. Sofia University, Faculty of Mathematics and Informatics, Department of Information Technologies. Sofia. July, 2001.
  2. Nakov P. Latent Semantic Analysis for German literature investigation. In Proceedings of the 7th Fuzzy Days'01, International Conference on Computational Intelligence. B. Reusch (Ed.): LNCS 2206. pp. 834-641. Dortmund, Germany. October 1-3, 2001.
  3. Nakov P., Popova A., Mateev P. Weight functions impact on LSA performance. In Proceedings of the EuroConference Recent Advances in Natural Language Processing (RANLP'01). pp. 187-193. Tzigov Chark, Bulgaria, September 5-7, 2001.
  4. Nakov P. Latent Semantic Analysis for Bulgarian Literature. In Proceedings of the Spring Conference of Bulgarian Mathematicians Union. pp. 279-284. Borovetz, Bulgaria. 2001.
  5. Nakov P., Terzieva S. Study of notional structures based on latent semantic analysis. In Proceedings of the Annual Journal of the University of Chemical Technology and Metallurgy, vol 37/2001. Sofia, Bulgaria. 2001.
  6. Terzieva S., Nakov P., Handjieva S. Investigating the Degree of Adequacy of the Relations in the Concept Structure of Students using the Method of Latent Semantic Analysis. In Proceedings of the Bulgarian Computer Science Conference on Computer Systems and Technologies (CompSysTech'01). Sofia, Bulgaria. 2001.
  7. Atanassova I., Nakov P. The Impact of the Segmentation on the Automatic Hyponyms Extraction from Terminological Dictionaries. In Proceedings of the Conference on Contemporary Achievements in the Philological Sciences and the Foreign Language University Education. pp. 140-155, Veliko Turnovo, Bulgaria. May 15-17, 2001. (in Bulgarian)
  8. Atanassova I., Nakov P. Term and Document from the Point of View of the Latent Semantic Analysis. In Proceedings of the International Conference "Technologies, Safety and Ecology", pp.(69)193-205. Veliko Turnovo, Bulgaria. June 21-22, 2001. (in Bulgarian)
  9. Nakov P. Intelligent Agents for Personalised Information Delivery. In Proceedings of the Naval Scientific Forum, vol. 4 (Mechanical Engineering and Mathematics. Information Technology), pp. 284-291. Varna, Bulgaria. 2001. (in Bulgarian)
  10. Nakov P. Latent Semantic Analysis for Russian literature investigation. In Proceedings of the Naval Scientific Forum, vol. 4 (Mechanical Engineering and Mathematics. Information Technology), pp. 292-299. Varna, Bulgaria. 2001.
  11. Nakov P., Atanassova I. Automatic hyponyms extraction from Bulgarian and Russian terminological dictionaries. In Proceedings of the Naval Scientific Forum, vol. 3 (Applied Linguistics and Foreign Language Teaching), pp.327-335. Varna, Bulgaria. 2001. (in Bulgarian)
  12. Atanassova I., Nakov P. Factors Influencing the Automatic Hyponyms Extraction from Terminological Dictionaries using Latent Semantic Analysis. Jubilee scientific conference, Shumen University "Ep. Konstantin Preslavski". Shumen, Bulgaria. November, 2001. (in Russian)
  13. Terzieva S., Nakov P. Research in the Notional Structures of the Declarative Memory of Students using Latent Semantic Analysis. Bulgarian Journal of Psychology. vol. 1-2/2001. Sofia, Bulgaria. 2001. (in Bulgarian)
- 2000
1. Nakov P. Getting Better Results with Latent Semantic Indexing. In Proceedings of the Students Presentations at the European Summer School in Logic Language and Information (ESSLLI'00). pp. 156-166. Birmingham, UK. August 2000.
  2. Nakov P. Web Personalization Using Extended Boolean Operations with Latent Semantic Indexing. In Lecture Notes in Artificial Intelligence — 1904 (Springer). Artificial Intelligence: Methodology, Systems and Applications. 9th International Conference (AIMSA'00), pp. 189-198. Varna, Bulgaria, September 2000.
  3. Paskaleva E., Nakov P., Angelova G., Racheva P., Mateev P.: Matching Text Meaning for Information Retrieval Tasks in Inflectional Languages, Fifth TELRI European Seminar in Corpus Linguistics: How to Extract Meaning from Corpora, Ljubljana, Slovenia, September 22-24, 2000.
  4. Nakov P. Latent Semantic Analysis of Textual Data. In Proceedings of the International Conference on Computer Systems and Technologies (CompSysTech'00). pp. V.3-1–V.3-5. Sofia, Bulgaria. June 2000.
- 1996
1. Nakov P. Object-Oriented Technology in the Contemporary Computer Programming. In Summaries of Student Topics. 38th London International Youth Science Forum. London. 1996.

## (\* Common European Framework of Reference for Languages

The self-assessment grid is based on the six level scale of the common European framework of reference for languages developed by the Council of Europe. The grid consists of three broad levels as follows:

- Basic user (levels A1 and A2);
- Independent user (levels B1 and B2);
- Proficient user (levels C1 and C2).

### Understanding

#### Listening

**A 1:** I can understand familiar words and very basic phrases concerning myself, my family and immediate surroundings when people speak slowly and clearly.

**A 2:** I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local area, employment). I can catch the main points in short, clear, simple messages and announcements.

**B 1:** I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main points of many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.

**B 2:** I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar. I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect.

**C 1:** I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.

**C 2:** I have no difficulty in understanding any kind of spoken language, whether live or broadcast, even when delivered at fast native speed, provided I have some time to get familiar with the accent.

#### Reading

**A 1:** I can understand familiar names, words and very simple sentences, for example on notices and posters or in catalogues.

**A 2:** I can read very short, simple texts. I can find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables and I can understand short simple personal letters.

**B 1:** I can understand texts that consist mainly of high frequency everyday or job-related language. I can understand the description of events, feelings and wishes in personal letters.

**B 2:** I can read articles and reports concerned with contemporary problems in which the writers adopt particular attitudes or viewpoints. I can understand contemporary literary prose.

**C 1:** I can understand long and complex factual and literary texts, appreciating distinctions of style. I can understand specialised articles and longer technical instructions, even when they do not relate to my field.

**C 2:** I can read with ease virtually all forms of the written language, including abstract, structurally or linguistically complex texts such as manuals, specialised articles and literary works.

### Speaking

#### Spoken interaction

**A 1:** I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.

**A 2:** I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.

**B 1:** I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).

**B 2:** I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.

**C 1:** I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.

**C 2:** I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.

#### Spoken production

**A 1:** I can use simple phrases and sentences to describe where I live and people I know.

**A 2:** I can use a series of phrases and sentences to describe, in simple terms, my family and other people, living conditions, my educational background and my present or most recent job.

**B 1:** I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes and ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my reactions.

**B 2:** I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.

**C 1:** I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.

**C 2:** I can present a clear, smoothly-flowing description or argument in a style appropriate to the context and with an effective logical structure which helps the recipient to notice and remember significant points.

### Writing

**A 1:** I can write a short, simple postcard, for example sending holiday greetings. I can fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form.

**A 2:** I can write short, simple notes and messages. I can write a very simple personal letter, for example thanking someone for something.

**B 1:** I can write simple connected text on topics which are familiar or of personal interest. I can write personal letters describing experiences and impressions.

**B 2:** I can write clear, detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.

**C 1:** I can express myself in clear, well-structured text, expressing points of view at some length. I can write about complex subjects in a letter, an essay or a report, underlining what I consider to be the salient issues. I can select a style appropriate to the reader in mind.

**C 2:** I can write clear, smoothly-flowing text in an appropriate style. I can write complex letters, reports or articles which present a case with an effective logical structure which helps the recipient to notice and remember significant points. I can write summaries and reviews of professional or literary works.