Ion (Ioannis) Androutsopoulos

Curriculum Vitae

Last update: June 6, 2026

1. CONTACT INFORMATION

Postal address:	Department of Informatics, Athens University of Economics and Business, Patission 76, GR-104 34 Athens, Greece
Office phone:	+30 210 8203 122
E-mail:	ion@aueb.gr
Personal Web page:	http://www.aueb.gr/users/ion/

2. CURRENT RESEARCH INTERESTS

Artificial Intelligence, especially **Natural Language Processing** (NLP), and particularly:

- **machine learning** in NLP, especially **deep learning** and **large language models**,
- **question answering**, **retrieval augmented generation**, **multi-step reasoning** for document collections,
- **dialog systems**, especially porting them to **new languages and knowledge domains**,
- **sentiment analysis** and **emotion recognition** from text and speech,
- **improving online discussions**, including detecting and handling **toxic posts and disinformation**, and using large language models as **facilitators**,
- **image to text generation**, mostly generating diagnostic tags and diagnostic reports from **medical images**,
- NLP for **biomedical, legal, financial** data and NLP in the **digital humanities**,
- NLP **tools for Greek**.

3. STUDIES

1992–96: Department of Artificial Intelligence, University of Edinburgh, U.K. **PhD in Artificial Intelligence**. Thesis title: *A Principled Framework for Constructing Natural Language Interfaces to Temporal Databases.*

1991–92: Department of Artificial Intelligence, University of Edinburgh, U.K. **MSc in Information Technology/Knowledge-Based Systems**, Natural Language Processing theme. (Official graduation: 1994). Grade: distinction (excellent). Dissertation title: *Interfacing a Natural Language Front-End to a Relational Database.* Dissertation grade: A (excellent).

1986–91: Department of Electrical Engineering, National Technical University of Athens, Greece. "**Diploma**" **in Electrical Engineering** (5-year MSc-level studies), Computer Engineering theme. Grade: 9.06/10 (excellent). Dissertation title: *Construction of a Lexical and Syntax Analyzer for Modern Greek.* Dissertation grade: 10/10 (excellent).

4. WORK EXPERIENCE

2020–present: Professor, Department of Informatics, Athens University of Economics and Business (AUEB).

2013–2019: Associate Professor, Department of Informatics, Athens University of Economics and Business (AUEB).

2007–2013: Assistant Professor (tenured since 2011), Department of Informatics, Athens University of Economics and Business (AUEB).

2002–2007: Lecturer, Department of Informatics, AUEB.

2000–02: Post-doctoral Fellow, Software and Knowledge Engineering Laboratory, Institute of Informatics and Telecommunications, NCSR "Demokritos".

1998–2000: Research Associate, Software and Knowledge Engineering Laboratory, Institute of Informatics and Telecommunications, NCSR "Demokritos".

1997–98: Military service (18 months), Hellenic Army. Specialty: Computer Programmer.

1996–97: Research Scientist, Language Technology Group, Microsoft Research Institute, Macquarie University, Sydney, Australia.

5. TEACHING EXPERIENCE

5.1 Undergraduate courses

- ***Artificial Intelligence*** (2002–03 to 2013–14, 2015–16 to present). Dept. of Informatics, AUEB.
- ***Human-Computer Interaction*** (2002–03, 2013–14 to 2024–25). Dept. of Informatics, AUEB.
- ***Computer Programming in C++*** (2001–02 to 2008–09, 2011–12, 2012–13). Dept. of Informatics, AUEB.

5.2 Graduate courses

- ***Natural Language Processing*** (2010–11 to present, formerly "Language Technology"). MSc "Information Systems" and MSc "Computer Science", Dept. of Informatics, AUEB.
- ***Text Analytics*** (2015–16 to present). MSc "Data Science", Dept. of Informatics, AUEB.
- ***Social Network Analysis and Social Media Analytics*** (2015–16). MSc "Business Analytics", Dept. of Management Science and Technology, AUEB. Jointly with G. Giaglis and H. Papageorgiou.
- ***Logic and Artificial Intelligence*** (2003–04 to 2012–13). MSc "Computer Science", Dept. of Informatics, AUEB.

5.3 Student supervision: See https://www.aueb.gr/users/ion/students.html

6. AWARDS, DISTINCTIONS, FELLOWSHIPS

2025: The August 2025 "Updated science-wide author databases of standardized citation indicators" of J.P.A. Ioannidis (Stanford University) included me in the **top 2% scientists worldwide**, for **2024 publications** and **career-long publications up to 2024** inclusive.
https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/8

2025: AUEB's NLP Group ranked **1st in concept detection, 5th in caption prediction, 1st in explainability** in **ImageCLEFmed Caption 2025**.
https://www.imageclef.org/2025/medical/caption

2025: Excellent teaching award (for undergraduate courses of the academic year 2023–24), Department of Informatics, **AUEB**, for teaching the "**Artificial Intelligence**" course.
https://www.aueb.gr/el/content/vraveia-exairetikis-epidosis-sti-didaskalia-apo-oikonomiko-panepistimio-athinon-2

2024: The August 2024 "Updated science-wide author databases of standardized citation indicators" of J.P.A. Ioannidis (Stanford University) included me in the **top 2% scientists worldwide**, for **2023 publications** and **career-long publications up to 2023** inclusive.
https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/7
https://www.aueb.gr/el/content/pagkosmia-anagnorisi-gia-diakekrimenoys-epistimones-toy-oikonomikoy-panepistimioy-athinon

2024: Excellent teaching award for teaching in the MSc program "Computer Science" (for the academic year 2021–22), Department of Informatics, **AUEB**, for teaching the "**Natural Language Processing**" course.

2024: AUEB's NLP Group ranked **2nd in concept detection** and **4th in caption prediction** in **ImageCLEFmed Caption 2024**. https://www.imageclef.org/2024/medical/caption

2023: **Excellent teaching award** (for undergraduate courses of the academic year 2021–22), Department of Informatics, **AUEB**, for teaching the "**Artificial Intelligence**" course. https://www.aueb.gr/el/content/vraveia-exairetikis-epidosis-sti-didaskalia-apo-oikonomiko-panepistimio-athinon

2023: Excellent teaching award for teaching in the MSc program "Computer Science" (for the academic year 2021–22), Department of Informatics, **AUEB**, for teaching the "**Natural Language Processing**" course.

2023: AUEB's NLP Group ranked **1st in concept detection** and **3nd in caption prediction** in **ImageCLEFmed Caption 2023**. https://www.imageclef.org/2023/medical/caption

2022: AUEB's NLP Group ranked **1st in concept detection** and **2nd in caption prediction** in **ImageCLEFmed Caption 2022**. https://www.imageclef.org/2022/medical/caption

2021: AUEB's NLP Group ranked **1st in concept detection** and **2nd in caption prediction** in **ImageCLEFmed Caption 2021**. https://www.imageclef.org/2021/medical/caption

2020: AUEB's NLP Group received a **BioASQ award** for ranking in the **top 2 positions in 4 out of 5 document retrieval test batches**, and for ranking **1st in 4 out of 5 snippet retrieval test batches**. http://bioasq.org/participate/eighth-challenge-winners

2020: AUEB's NLP Group ranked at **positions 1, 2, 6** among 49 systems in the **ImageCLEFmed Caption 2020** competition. https://www.aueb.gr/en/content/team-auebs-department-informatics-was-ranked-1st-position-imageclefmed-caption-2020
https://www.imageclef.org/2020/medical/caption/

2019: AUEB's NLP Group ranked at **positions 1, 2, 3, 5** among approx. 60 systems in **ImageCLEFmed Caption 2019**. https://www.imageclef.org/2019/medical/caption/

2019: AUEB's NLP Group received a **BioASQ award** for ranking **1st in 4 out of 5 document and snippet retrieval test batches**. https://www.aueb.gr/en/content/team-auebs-department-informatics-wins-bioasq-prize-0

2019: Excellent teaching award (for undergraduate courses of 2018–19), Department of Informatics, **AUEB**, for teaching the "**Artificial Intelligence**" course.
https://www.aueb.gr/sites/default/files/opanews-feb-mar_19.pdf

2018: AUEB's NLP Group received a **BioASQ award** for ranking **1st in 3 out of 5 document retrieval test batches and all 5 snippet retrieval test batches**.
https://www.aueb.gr/en/content/team-auebs-department-informatics-wins-bioasq-prize

2011: Mentor of the students G. Karakatsiotis and V. Pterneas, who won the **1st prize** (inter-operability category) in **Microsoft's International Imagine Cup 2011**.

2000–2002: Post-doctoral Fellowship, Institute of Informatics and Telecommunications, **NCSR "Demokritos"**, Greece.

1992–1994: Scholarship for PhD studies from the **Greek State Scholarships Foundation**.

1991: Arthur Andersen award for the **highest grade** among the **1991 graduates** of the **Computer Engineering** theme of **NTUA**'s Department of Electrical Engineering.

7. PUBLICATIONS

7.1 Books

B1. I. Androutsopoulos, *Exploring Time, Tense and Aspect in Natural Language Database Interfaces*, John Benjamins, 2002.

7.2 Journal articles

J21. N.C. Gkoumas, G.N. Leledakis, E.G. Pyrgiotakis and I. Androutsopoulos, "Bank Competition, Loan Portfolio Concentration and Stock Price Crash Risk: the Role of Tone Ambiguity". *British Journal of Management*, 36: 323-341, 2025.

J20. A.G. Katsafados, G.N. Leledakis, E.G. Pyrgiotakis, I. Androutsopoulos and Manos Fergadiotis, "Machine Learning in Bank Merger Prediction: A Text-based Approach". *European Journal of Operational Research*, 312(2): 783–797, 2024.

J19. T. Sommerschield, Y. Assael, J. Pavlopoulos, V. Stefanak, A. Senior, C. Dyer, J. Bodel, J. Prag, I. Androutsopoulos, N. de Freitas, "Machine Learning for Ancient Languages: A Survey". *Computational Linguistics*, 49(3):703–747, 2023.

J18. A.G. Katsafados, G.N. Leledakis, E.G. Pyrgiotakis, I. Androutsopoulos, I. Chalkidis and M.Fergadiotis, "Textual Information and IPO Underpricing: A Machine Learning Approach". *Journal of Financial Data Science*, 5(2):100–135, 2023.

J17. Y. Assael, T. Sommerschield, B. Shillingford, M. Bordbar, J. Pavlopoulos, M. Chatzipanagiotou, I. Androutsopoulos, J. Prag and N. de Freitas, "Restoring and attributing ancient texts using deep neural networks". *Nature*, 603, pp. 280–283, 2022.

J16. J. Pavlopoulos, V. Kougia, I. Androutsopoulos and D. Papamichail, "Diagnostic Captioning: A Survey". *Knowledge and Information Systems*, 64:1691–1722, 2022.

J15. A. Xenos, J. Pavlopoulos, I. Androutsopoulos, L. Dixon, J. Sorensen, L. Laugier, "Toxicity Detection can be Sensitive to the Conversational Context". *First Monday*, 27(9), 2022.

J14. K. Papantoniou, P. Papadakos, T. Patkos, G. Flouris, I. Androutsopoulos, D. Plexousakis, "Deception Detection in Text and its Relation to the Cultural Dimension of Individualism/Collectivism". *Natural Language Engineering*, 28(5):545–606, 2022.

J13. A. Katsafados, I. Androutsopoulos, I. Chalkidis, E. Fergadiotis, G. Leledakis, E. Pyrgiotakis, "Using Textual Analysis to Identify Merger Participants: Evidence from the U.S. Banking Industry". *Finance Research Letters*, 42:101949, 2021.

J12. G. Tsatsaronis1, G. Balikas, P. Malakasiotis, I. Partalas, M. Zschunke, M.R. Alvers, D. Weissenborn, A. Krithara, S. Petridis, D. Polychronopoulos, Y. Almirantis, J. Pavlopoulos, N. Baskiotis, P. Gallinari, T. Artiéres, A. Ngonga, N. Heino, E. Gaussier, L. Barrio-Alvers, M. Schroeder, I. Androutsopoulos, G. Paliouras, "An Overview of the BioASQ Large-Scale Biomedical Semantic Indexing and Question Answering Competition". *BMC Bioinformatics*, 16:138, 2015.

J11. A. Kosmopoulos, I. Partalas, E. Gaussier, G. Paliouras, I. Androutsopoulos, "Evaluation Measures for Hierarchical Classification: a Unified View and Novel Approaches". *Data Mining and Knowledge Discovery*, 29:820–865, Springer, 2015.

J10. I. Androutsopoulos, G. Lampouras, D. Galanis, "Generating Natural Language Descriptions from OWL Ontologies: the NaturalOWL System". *Journal of Artificial Intelligence Research*, 48:671–715, 2013.

J9. I. Androutsopoulos and P. Malakasiotis, "A Survey of Paraphrasing and Textual Entailment Methods". *Journal of Artificial Intelligence Research*, 38:135–187, 2010.

J8. P. Constantopoulos, C. Dallas, I. Androutsopoulos, S. Angelis, A. Deligiannakis, D. Gavrilis, Y. Kotidis, C. Papatheodorou, "DCC&U: An Extended Digital Curation Lifecycle Model". *International Journal of Digital Curation*, 1(4):34–45, 2009.

J7. E.F. Mageirou, D.K. Vassilakis, I. Androutsopoulos, "Teaching Nonzero Sum Games Using a Diagrammatic Determination of Equilibria". *INFORMS Transactions on Education*, 8(3):115–124, 2008.

J6. G. Lucarelli, X. Vasilakos, I. Androutsopoulos, "Named Entity Recognition in Greek Texts with an Ensemble of Support Vector Machines and Active Learning". *International Journal on Artificial Intelligence Tools*, 16(6):1015–1045, World Scientific, 2007.

J5. I. Androutsopoulos, J. Oberlander, V. Karkaletsis, "Source Authoring for Multilingual Generation of Personalised Object Descriptions". *Natural Language Engineering*, 13(3):191–233, Cambridge University Press, 2007.

J4. A. Isard, J. Oberlander, I. Androutsopoulos, C. Matheson, "Speaking the Users' Languages". *IEEE Intelligent Systems*, special issue on Advances in Natural Language Processing, 18(1):40–45, 2003.

J3. G. Sakkis, I. Androutsopoulos, G. Paliouras, V. Karkaletsis, C.D. Spyropoulos, P. Stamatopoulos, "A Memory-Based Approach to Anti-Spam Filtering for Mailing Lists". *Information Retrieval*, 6(1):49–73, Kluwer, 2003.

J2. I. Androutsopoulos, G.D. Ritchie, P. Thanisch, "Time, Tense and Aspect in Natural Language Database Interfaces". *Natural Language Engineering*, 4(3): 229–276, Cambridge University Press, 1998.

J1. I. Androutsopoulos, G.D. Ritchie, P. Thanisch, "Natural Language Interfaces to Databases – An Introduction". *Natural Language Engineering*, 1(1):29–81, Cambridge University Press, 1995.

7.3 Book chapters

K3. J. Calder, A.C. Melengoglou, C. Callaway, E. Not, F. Pianesi, I. Androutsopoulos, C.D. Spyropoulos, G. Xydas, G. Kouroupetroglou, M. Roussou, "Multilingual Personalized

Information Objects". In O. Stock and M. Zancanaro (Eds.), *Multimodal Intelligent Information Presentation*, pp. 177–201, Springer, 2005.

K2. I. Androutsopoulos and M. Aretoulaki, "Natural Language Interaction". Chapter 35 in R. Mitkov (Ed.), *Handbook of Computational Linguistics*, pp. 629–649, Oxford University Press, 2003.

K1. I. Androutsopoulos and G.D. Ritchie, "Database Interfaces". Chapter 9 in R. Dale, H. Moisl and H. Somers (Eds.), *Handbook of Natural Language Processing*, pp. 209–240, Marcel Dekker, 2000.

7.4 Conference articles (peer reviewed)

C112. S. Eleftheriou, T. Giannakopoulos, T. Stafylakis and I. Androutsopoulos, "From Pretraining to Robustness: Benchmarking SSL Models for Noise-Robust Speech Emotion Recognition". Proceedings of the *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP 2026), Barcelona, Spain, pp. 19497-19501, 2026.

C111. O.S. Chlapanis, D. Galanis, N. Aletras and I. Androutsopoulos, **"**GreekBarBench: A Challenging Benchmark for Free-Text Legal Reasoning and Citations". Findings of the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2025), Suzhou, China, 2025.

C110. K. Korre, D. Tsirmpas, N. Gkoumas, E. Cabalé, D. Myrtzani, T. Evgeniou, I. Androutsopoulos and J. Pavlopoulos, "Evaluation and Facilitation of Online Discussions in the LLM Era: A Survey". Proceedings of the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2025), Suzhou, China, 2025.

C109. A. Chatzipapadopoulou, I. Pantelidis, F. Charalampakos, M. Samprovalaki, G. Moschovis, P. Kaliosis, K.V. Dalakleidi, J. Pavlopoulos and I. Androutsopoulos, "AUEB NLP Group at ImageCLEF medical 2025". Proceedings of the *Conference and Labs of the Evaluation Forum* (CLEF 2025), Madrid, Spain, 2025.

C108. C. Vlachos, N. Stylianou, A. Fiotaki, S. Methenitis, E. Palogiannidi, T. Stafylakis and I. Androutsopoulos, "Building Open-Retrieval Conversational Question Answering Systems by Generating Synthetic Data and Decontextualizing User Questions". Proceedings of the *26th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (SIGDIAL 2025), Avignon, France, 2025.

C107. N. Nikolaidis, N. Stefanovitch, P. Silvano, D. Dimitrov, R. Yangarber, N. Guimaraes, E. Sartori, I. Androutsopoulos, P. Nakov, G. Da San Martino, J. Piskorski, "PolyNarrative: A Multilingual, Multilabel, Multi-domain Dataset for Narrative Extraction from News Articles". Proceedings of the *63rd Annual Meeting of the Association for Computational Linguistics* (ACL 2025), Vienna, Austria, 2025.

C106. L. Loukas, F. Billert, M. Fergadiotis, P. Malakasiotis and I. Androutsopoulos, "EDGAR-CRAWLER: From Raw Web Documents to Structured Financial NLP Datasets". Companion Proceedings of the *ACM Web Conference* (WWW 2025), pp. 761–764, Sydney, Australia, 2025.

C105. I. Chasandras, O.S. Chlapanis and I. Androutsopoulos, "AUEB-Archimedes at RIRAG-2025: Is Obligation Concatenation Really all you Need?". Proceedings of the *Workshop on Regulatory Natural Language Processing* (RegNLP 2025) of the *International Conference on Computational Linguistics* (COLING 2025), Abu Dhabi, UAE, 2025.

C104. O.S. Chlapanis, D. Galanis and I. Androutsopoulos, "LAR-ECHR: A New Legal Argument Reasoning Task and Dataset for Cases of the European Court of Human Rights". Proceedings of the *Workshop on Natural Legal Language Processing* (NLLP 2024) of the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2024), Miami, Florida, USA, 2024.

C103. M. Samprovalaki, A. Chatzipapadopoulou, G. Moschovis, F. Charalampakos, P. Kaliosis, J. Pavlopoulos and I. Androutsopoulos, "AUEB NLP Group in ImageCLEF medical 2024". Proceedings of the *Conference and Labs of the Evaluation Forum* (CLEF 2024), Grenoble, France, 2024. (**Highlighted talk**.)

C102. P. Kaliosis, J. Pavlopoulos, F. Charalampakos, G. Moschovis and Ion Androutsopoulos, "A Data-Driven Guided Decoding Mechanism for Diagnostic Captioning". *Findings of the 62nd Annual Meeting of the Association for Computational Linguistics* (ACL 2024), Bangkok, Thailand, 2024.

C101. C. Vlachos, T. Stafylakis and Ion Androutsopoulos, "Comparing Data Augmentation Methods for End-to-End Task-Oriented Dialog Systems". *Findings of the 62nd Annual Meeting of the Association for Computational Linguistics* (ACL 2024), Bangkok, Thailand, 2024.

C100. A. Toumazatos, J. Pavlopoulos, I. Androutsopoulos and S. Vassos, "Still all Greeklish to me: Greeklish to Greek Transliteration". Proceedings of the *Joint International Conference on Computational Linguistics, Language Resources and Evaluation* (LREC-COLING 2024), Torino, Italy, pp. 15309–15319, 2024.

C99. O. Chlapanis, I. Androutsopoulos and D. Galanis, "Archimedes-AUEB at SemEval-2024 Task 5: LLM explains Civil Procedure". Proceedings of the *18th International Workshop on Semantic Evaluation* (SemEval 2024) of the *Annual Conference of the North American Chapter of the Association for Computational Linguistics* (NAACL 2024), Mexico City, Mexico, pp. 1607–1622, 2024.

C98. N. Gkouti, P. Malakasiotis, S. Toumpis and I. Androutsopoulos, "Should I Try Multiple Optimizers when Fine-tuning Pre-trained Transformers for NLP tasks? Should I Tune their Hyperparameters?". Proceedings of the *18th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2024), Malta, 2024.

C97. I. Stogiannidis, S. Vassos, P. Malakasiotis and I. Androutsopoulos, "Cache me if you Can: an Online Cost-aware Teacher-Student Framework to Reduce the Calls to Large Language Models". *Findings of the Conference on Empirical Methods in Natural Language Processing* (EMNLP 2023), Singapore, 2023.

C96. K. Korre, J. Pavlopoulos, J. Sorensen, L. Laugier, I. Androutsopoulos, L. Dixon and A. Barron-cedeno, "Harmful Language Datasets: An Assessment of Robustness". Proceedings of the *7th Workshop on Online Abuse and Harms* (WOAH 2023) of the *61st Annual Meeting of the Association for Computational Linguistics* (ACL 2023), Toronto, Canada, 2023.

C95. P. Kaliosis, G. Moschovis, F. Charalampakos, J. Pavlopoulos and I. Androutsopoulos, "AUEB NLP Group at ImageCLEFmedical Caption 2023". *Working Notes of the Conference and Labs of the Evaluation Forum* (CLEF 2023, https://ceur-ws.org/Vol-3497/), Thessaloniki, Greece, 2023.

C94. D. Mamakas, P. Tsotsi, I. Androutsopoulos and I. Chalkidis, "Processing Long Legal Documents with Pre-trained Transformers: Modding LegalBERT and Longformer". Proceedings of the *4th Workshop on Natural Legal Language Processing* (NLLP 2022) of the

Conference on Empirical Methods in Natural Language Processing (EMNLP 2022), Abu Dhabi, United Arab Emirates, 2022.

C93. F. Charalampakos, G. Zachariadis, J. Pavlopoulos, V. Karatzas, C. Trakas and I. Androutsopoulos, "AUEB NLP Group at ImageCLEFmedical Caption 2022". Proceedings of *CLEF 2022 – Conference and Labs of the Evaluation Forum*, Bologna, Italy, 2022.

C92. S. Xenouleas, A. Tsoukara, G. Panagiotakis, I. Chalkidis and I. Androutsopoulos, "Realistic Zero-Shot Cross-Lingual Transfer in Legal Topic Classification". Proceedings of the *12th Hellenic Conference on Artificial Intelligence* (SETN 2022), Corfu, Greece, 2022.

C91. I. Chalkidis, A. Jana, D. Hartung, M. Bommarito, I. Androutsopoulos, D.M. Katz and N. Aletras, "LexGLUE: A Benchmark Dataset for Legal Language Understanding in English". Proceedings of the *60th Annual Meeting of the Association for Computational Linguistics* (ACL 2022), Dublin, Ireland, 2022.

C90. J. Pavlopoulos, L. Laugier, A. Xenos, J. Sorensen and I. Androutsopoulos, "From the Detection of Toxic Spans in Online Discussions to the Analysis of Toxic-to-Civil Transfer". Proceedings of the *60th Annual Meeting of the Association for Computational Linguistics* (ACL 2022), Dublin, Ireland, 2022.

C89. L. Loukas, M. Fergadiotis, I. Chalkidis, E. Spyropoulou, P. Malakasiotis, I. Androutsopoulos and G. Paliouras, "FiNER: Financial Numeric Entity Recognition for XBRL Tagging". Proceedings of the *60th Annual Meeting of the Association for Computational Linguistics* (ACL 2022), Dublin, Ireland, 2022.

C88. D. Pappas, P. Malakasiotis and I. Androutsopoulos, "Data Augmentation for Biomedical Factoid Question Answering". Proceedings of the *21st Workshop on Biomedical Natural Language Processing* (BioNLP 2022) at the *60th Annual Meeting of the Association for Computational Linguistics* (ACL 2022), Dublin, Ireland, 2022.

C87. L. Loukas, M. Fergadiotis, I. Androutsopoulos and P. Malakasiotis, "EDGAR-CORPUS: Billions of Tokens Make The World Go Round". Proceedings of the *Economics and Natural Language Processing Workshop* (EcoNLP 2021) of the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2021), on-line and Punta Cana, Dominican Republic, 2021.

C86. F. Charalampakos, V. Karatzas, V. Kougia, J. Pavlopoulos and I. Androutsopoulos, "AUEB NLP Group at ImageCLEFmed Caption Tasks 2021". Proceedings of the *Conference and Labs of the Evaluation Forum* (CLEF 2021), Bucharest, Romania, 2021.

C85. I. Chalkidis, M. Fergadiotis and I. Androutsopoulos, "MultiEURLEX – A Multi-lingual and Multi-label Legal Document Classification Dataset for Zero-shot Cross-lingual Transfer". Proceedings of the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2021), on-line and Punta Cana, Dominican Republic, 2021.

C84. D. Pappas and I. Androutsopoulos, "A Neural Model for Joint Document and Snippet Ranking in Question Answering for Large Document Collections". Proceedings of the *Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (ACL-IJCNLP 2021), held on-line, 2021.

C83. I. Chalkidis, M. Fergadiotis, D. Tsarapatsanis, N. Aletras, I. Androutsopoulos and P. Malakasiotis, "Paragraph-level Rationale Extraction through Regularization: A case study on European Court of Human Rights Cases". Proceedings of the *Annual Conference of the North*

American Chapter of the Association for Computational Linguistics (NAACL 2021), held on-line, 2021.

C82. G. Vernikos, K. Margatina, A. Chronopoulou and I. Androutsopoulos, "Domain Adversarial Fine-Tuning as an Effective Regularizer". *Findings of the Conference on Empirical Methods in Natural Language Processing* (EMNLP 2020), held on-line, 2020.

C81. I. Chalkidis, M. Fergadiotis, S. Kotitsas, P. Malakasiotis, N. Aletras and I. Androutsopoulos, "An Empirical Study on Large-Scale Multi-Label Text Classification Including Few and Zero-Shot Labels". Proceedings of the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2020), held on-line, 2020.

C80. I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras and I. Androutsopoulos, "LEGAL-BERT: The Muppets straight out of Law School". *Findings of the Conference on Empirical Methods in Natural Language Processing* (EMNLP 2020), held on-line, 2020.

C79. D. Pappas, P. Stavropoulos and I. Androutsopoulos, "AUEB-NLP at BioASQ 8: Biomedical Document and Snippet Retrieval". Proceedings of the *8th BioASQ workshop* at the *Conference and Labs of the Evaluation Forum* (CLEF 2020), Thessaloniki, Greece, 2020.

C78. V. Kougia, J. Pavlopoulos and I. Androutsopoulos, "Medical Image Tagging by Deep Learning and Retrieval". Proceedings of the *International Conference of the Cross-Language Evaluation Forum for European Languages* (CLEF 2020), Best of CLEF 2019 labs, held on-line, pp. 154–166, 2020.

C77. B. Karatzas, J. Pavlopoulos, V. Kougia and I. Androutsopoulos, "AUEB NLP Group at ImageCLEFmed Caption 2020". *Working Notes of the Conference and Labs of the Evaluation Forum* (CLEF 2020), Thessaloniki, Greece, 2020.

C76. J. Koutsikakis, I. Chalkidis, P. Malakasiotis and I. Androutsopoulos, "Greek-BERT: The Greeks Visiting Sesame Street". Proceedings of the 11th Hellenic Conference on Artificial Intelligence (SETN 2020), held on-line, 2020.

C75. J. Pavlopoulos, J. Sorensen, L. Dixon and I. Androutsopoulos, "Toxicity Detection: Does Context Really Matter?". Proceedings of the *58th Annual Meeting of the Association for Computational Linguistics* (ACL 2020), held on-line, 2020.

C74. P. Stavropoulos, D. Pappas, I. Androutsopoulos and R. McDonald, "BIOMRC: A Dataset for Biomedical Machine Reading Comprehension". Proceedings of the *19th Workshop on Biomedical Natural Language Processing* (BioNLP 2020) of the *58th Annual Meeting of the Association for Computational Linguistics* (ACL 2020), held on-line, 2020.

C73. I. Chalkidis, M. Fergadiotis, P. Malakasiotis and I. Androutsopoulos, "Neural Contract Element Extraction Revisited". Proceedings of the *Document Intelligence Workshop* of the *33rd Conference on Neural Information Processing Systems* (NeurIPS 2019), Vancouver, Canada, 2019.

C72. D. Pappas, R. McDonald, G.I. Brokos and I. Androutsopoulos, "AUEB at BioASQ 7: Document and Snippet Retrieval". Proceedings of the *7th BioASQ workshop* at the *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases* (ECML-PKDD), Würzburg, Germany, 2019.

C71. S. Xenouleas, P. Malakasiotis, M. Apidianaki and I. Androutsopoulos, "SUM-QE: a BERT-based Summary Quality Estimation Model". Proceedings of the *2019 Conference on*

Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP 2019), Hong Kong, China, 2019.

C70. I. Chalkidis, I. Androutsopoulos and N. Aletras, "Neural Legal Judgment Prediction in English". Proceedings of the *57th Annual Meeting of the Association for Computational Linguistics* (ACL 2019), Florence, Italy, pp. 4317–4323, 2019.

C69. I. Chalkidis, M. Fergadiotis, P. Malakasiotis and I. Androutsopoulos, "Large-Scale Multi-Label Text Classification on EU Legislation". Proceedings of the *57th Annual Conference of the Association for Computational Linguistics* (ACL 2019), Florence, Italy, pp. 6314–6322, 2019.

C68. S. Kotitsas, D. Pappas, I. Androutsopoulos, R. McDonald and M. Apidianaki, "Embedding Biomedical Ontologies by Jointly Encoding Network Structure and Textual Node Descriptors". Proceedings of the *18th Workshop on Biomedical Natural Language Processing* (BioNLP 2019) of the *57th Annual Meeting of the Association for Computational Linguistics* (ACL 2019), Florence, Italy, pp. 298–308, 2019.

C67. M. Kyriakakis, I. Androutsopoulos, J. Gines i Ametlle and A. Saudabayev, "Transfer Learning for Causal Sentence Detection". Proceedings of the *18th Workshop on Biomedical Natural Language Processing* (BioNLP 2019) of the *57th Annual Meeting of the Association for Computational Linguistics* (ACL 2019), Florence, Italy, pp. 292–297, 2019.

C66. C. Baziotis, I. Androutsopoulos, I. Konstas and Alexandros Potamianos, "SEQ^3: Differentiable Sequence-to-Sequence-to-Sequence Autoencoder for Unsupervised Abstractive Sentence Compression". Proceedings of the *Annual Conference of the North American Chapter of the Association for Computational Linguistics* (NAACL-HLT 2019), Minneapolis, USA, pp. 673–681, 2019.

C65. V. Kougia, J. Pavlopoulos and I. Androutsopoulos, "A Survey on Biomedical Image Captioning". Proceedings of the Workshop on Shortcomings in Vision and Language (SiVL 2019) of the *Annual Conference of the North American Chapter of the Association for Computational Linguistics* (NAACL-HLT 2019), Minneapolis, USA, pp. 26–36, 2019.

C64. I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras and I. Androutsopoulos, "Extreme Multi-Label Legal Text Classification: A Case Study in EU Legislation". Proceedings of the *Workshop on Natural Legal Language Processing* (NLLP 2019) of the *Annual Conference of the North American Chapter of the Association for Computational Linguistics* (NAACL-HLT 2019), Minneapolis, USA, pp. 78–87, 2019.

C63. J. Pavlopoulos, N. Thain, I. Androutsopoulos and L. Dixon. "ConvAI at SemEval-2019 Task 6: Offensive Language Identification and Categorization with Perspective and BERT". Proceedings of the *13th International Workshop on Semantic Evaluation* (SemEval) of the *Annual Conference of the North American Chapter of the Association for Computational Linguistics* (NAACL-HLT 2019), Minneapolis, USA, pp. 571–576, 2019.

C62. V. Kougia, J. Pavlopoulos and I. Androutsopoulos, "AUEB NLP Group at ImageCLEFmed Caption 2019". *Working Notes of the Conference and Labs of the Evaluation Forum* (CLEF 2019, http://ceur-ws.org/Vol-2380/), Lugano, Switzerland, 2019.

C61. R. McDonald, G. Brokos and I. Androutsopoulos, "Deep Relevance Ranking Using Enhanced Document-Query Interactions". Proceedings of the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2018), Brussels, Belgium, pp. 1849–1860, 2018.

C60. G. Brokos, P. Liosis, R. McDonald, D. Pappas and I. Androutsopoulos, "AUEB at BioASQ 6: Document and Snippet Retrieval". Proceedings of the workshop *BioASQ: Large-scale Biomedical Semantic Indexing and Question Answering*, at the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2018), Brussels, Belgium, pp. 30–39, 2018.

C59. V. Pertsas, P. Constantopoulos and I. Androutsopoulos, "Ontology Driven Extraction of Research Processes". Proceedings of the *17th International Semantic Web Conference* (ISWC 2018), Monterey, CA, USA, pp. 162–178, 2018. **Nominated for best student paper award.**

C58. I. Chalkidis, I. Androutsopoulos and A. Michos, "Obligation and Prohibition Extraction Using Hierarchical RNNs". Proceedings of the *56th Annual Meeting of the Association for Computational Linguistics* (ACL 2018), Melbourne, Australia, pp. 254–259 (short papers), 2018.

C57. D. Pappas, I. Androutsopoulos and H. Papageorgiou, "BioRead: A New Dataset for Biomedical Reading Comprehension". Proceedings of the *11th Language Resources and Evaluation Conference* (LREC 2018), Miyazaki, Japan, pp. 2771–2776, 2018.

C56. M. Vougioukas, I. Androutsopoulos and G. Paliouras, "Identifying Retweetable Tweets with a Personalized Global Classifier". Proceedings of the *10th Hellenic Conference on Artificial Intelligence* (SETN 2018), Patras, 2018.

C55. I. Chalkidis and I. Androutsopoulos, "A Deep Learning Approach to Contract Element Extraction". Proceedings of the *30th International Conference on Legal Knowledge and Information Systems* (JURIX 2017), pp. 155–164, Luxembourg, 2017.

C54. J. Pavlopoulos, P. Malakasiotis and I. Androutsopoulos, "Deeper Attention to Abusive User Content Moderation". Proceedings of the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2017), Copenhagen, Denmark, pp. 1136–1146, 2017.

C53. J. Pavlopoulos, P. Malakasiotis, J. Bakagianni and I. Androutsopoulos, "Improved Abusive Comment Moderation with User Embeddings". Proceedings of the workshop *Natural Language Processing Meets Journalism*, at the *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2017), Copenhagen, Denmark, pp. 51–55, 2017.

C52. J. Pavlopoulos, P. Malakasiotis and I. Androutsopoulos, "Deep Learning for User Comment Moderation". Proceedings of the *1st Workshop on Abusive Language Online* (ALW1), at the 55th *Annual Meeting of the Association for Computational Linguistics* (ACL 2017), Vancouver, Canada, pp. 25–35, 2017.

C51. M. Vougioukas, I. Androutsopoulos and G. Paliouras, "A Personalized Global Filter To Predict Retweets". Proceedings of the *25th Conference on User Modeling, Adaptation and Personalization* (UMAP 2017), Bratislava, Slovakia, pp. 393–394, 2017.

C50. I. Chalkidis, I. Androutsopoulos and A. Michos, "Extracting Contract Elements". Proceedings of the *16th International Conference on Artificial Intelligence and Law* (ICAIL 2017), London, UK, pp. 19–28, 2017.

C49. G. Brokos, P. Malakasiotis and I. Androutsopoulos, "Using Centroids of Word Embeddings and Word Mover's Distance for Biomedical Document Retrieval in Question Answering". Proceedings of the *15th Workshop on Biomedical Natural Language Processing* (BioNLP 2016), at the *54th Annual Meeting of the Association for Computational Linguistics* (ACL 2016), Berlin, Germany, pp. 114–118, 2016.

C48. M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. AL-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, V. Hoste, M. Apidianaki, X. Tannier, N. Loukachevitch, E. Kotelnikov, N. Bel, S.M. Jiménez-Zafra and G. Eryigit, "SemEval-2016 Task 5: Aspect Based Sentiment Analysis". Proceedings of the *10th International Workshop on Semantic Evaluation* (SemEval 2016), San Diego, CA, USA, pp. 19-30, 2016.

C47. S. Giorgis, A. Rousas, J. Pavlopoulos, P. Malakasiotis and I. Androutsopoulos, "aueb.twitter.sentiment at SemEval-2016 Task 4: A Weighted Ensemble of SVMs for Twitter Sentiment Analysis". Proceedings of the *10th International Workshop on Semantic Evaluation* (SemEval 2016), San Diego, CA, USA, pp. 96–99, 2016.

C46. D. Xenos, P. Theodorakakos, J. Pavlopoulos, P. Malakasiotis and I. Androutsopoulos, "AUEB-ABSA at SemEval-2016 Task 5: Supervised Machine Learning for Aspect Based Sentiment Analysis". Proceedings of the *10th International Workshop on Semantic Evaluation* (SemEval 2016), San Diego, CA, USA, pp. 324–329, 2016.

C45. D. Pappas, I. Androutsopoulos and H. Papageorgiou, "Anger Detection in Call Center Dialogues". Proceedings of the *6th IEEE Conference on Cognitive Infocommunications* (CogInfoCom 2015), Gyor, Hungary, 2015.

C44. P. Malakasiotis, E. Archontakis, I. Androutsopoulos, D. Galanis and H. Papageorgiou, "Biomedical Question-Focused Multi-document Summarization: ILSP and AUEB at BioASQ3". Proceedings of the *QA Lab of the 6th Conference and Labs of the Evaluation Forum* (CLEF 2015), Toulouse, France, 2015.

C43. M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar and I. Androutsopoulos, "SemEval-2015 Task 12: Aspect Based Sentiment Analysis". Proceedings of the *9th International Workshop on Semantic Evaluation* (SemEval 2015), at the *Conference of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies* (NAACL HLT 2015), Denver, Colorado, USA, pp. 486–495, 2015.

C42. A. Kosmopoulos, G. Paliouras and I. Androutsopoulos, "The Effect of Dimensionality Reduction on Large Scale Hierarchical Classification". Proceedings of the *5th Conference and Labs of the Evaluation Forum* (CLEF 2014), Sheffield, UK, pp. 160–171, 2014.

C41. M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos and S. Manandhar, "SemEval-2014 Task 4: Aspect Based Sentiment Analysis". Proceedings of the *8th International Workshop on Semantic Evaluation* (SemEval 2014), at the *25th International Conference on Computational Linguistics* (COLING 2014), Dublin, Ireland, pp. 27–35, 2014.

C40. J. Pavlopoulos and I. Androutsopoulos, "Multi-Granular Aspect Aggregation in Aspect-Based Sentiment Analysis". Proceedings of the *14th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2014), Gothenburg, Sweden, pp.78–87, 2014.

C39. J. Pavlopoulos and I. Androutsopoulos, "Aspect Term Extraction for Sentiment Analysis: New Datasets, New Evaluation Measures and an Improved Unsupervised Method". Proceedings of the *5th Workshop on Language Analysis for Social Media* (LASM 2014), *14th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2014), Gothenburg, Sweden, pp. 44–52, 2014.

C38. G. Lampouras and I. Androutsopoulos, "Using Integer Linear Programming in Concept-to-Text Generation to Produce More Compact Texts". Proceedings of the *51st Annual Meeting of the Association for Computational Linguistics* (ACL 2013), Sofia, Bulgaria, pp. 561–566 (short papers), 2013.

C37. G. Lampouras and I. Androutsopoulos, "Using Integer Linear Programming for Content Selection, Lexicalization, and Aggregation to Produce Compact Texts from OWL Ontologies". Proceedings of the *14th European Workshop on Natural Language Generation* (ENLG 2013), at the *51st Annual Meeting of the Association for Computational Linguistics* (ACL 2013), Sofia, Bulgaria, pp. 51–60, 2013.

C36. D. Galanis, G. Lampouras and I. Androutsopoulos, "Extractive Multi-Document Summarization with Integer Linear Programming and Support Vector Regression". Proceedings of the *24th International Conference on Computational Linguistics* (COLING 2012), Mumbai, India, 2012.

C35. G.Tsatsaronis, M. Schroeder, G. Paliouras, Y. Almirantis, I. Androutsopoulos, E. Gaussier, P. Gallinari, T. Artieres, M. Alvers, M. Zschunke and A. Ngonga, "BioASQ: A Challenge on Large-Scale Biomedical Semantic Indexing and Question Answering". Proceedings of the *AAAI Fall Symposium on Information Retrieval and Knowledge Discovery in Biomedical Text*, Arlington, VA, USA, pp. 92–98, 2012.

C34. V. Panagiotopoulou, I. Varlamis, I. Androutsopoulos and G. Tsatsaronis, "Word Sense Disambiguation as an Integer Linear Programming Problem". Proceedings of the *7th Hellenic Conference on Artificial Intelligence* (SETN 2012), Lamia, Greece, pp. 33–40, 2012.

C33. P. Malakasiotis and I. Androutsopoulos, "A Generate and Rank Approach to Sentence Paraphrasing". Proceedings of the *Conference on Empirical Methods on Natural Language Processing* (EMNLP 2011), Edinburgh, UK, pp. 96–106, 2011.

C32. D. Galanis and I. Androutsopoulos, "A New Sentence Compression Dataset and Its Use in an Abstractive Generate-and-Rank Sentence Compressor". Proceedings of the *Language Generation and Evaluation Workshop* (UCNLG+Eval), at the *Conference on Empirical Methods on Natural Language Processing* (EMNLP 2011), Edinburgh, UK, pp. 1–11, 2011.

C31. D. Galanis and I. Androutsopoulos, "An Extractive Supervised Two-Stage Method for Sentence Compression". Proceedings of *Human Language Technologies: The 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics* (HLT-NAACL 2010), Los Angeles, CA, USA, pp. 885–893, 2010.

C30. G. Lampouras and I. Androutsopoulos, "Finding Short Definitions of Terms on Web Pages". Proceedings of the *2009 Conference on Empirical Methods in Natural Language Processing* (EMNLP 2009), Suntec, Singapore, pp. 1270–1279, 2009.

C29. D. Vogiatzis, D. Galanis, V. Karkaletsis, I. Androutsopoulos and C.D. Spyropoulos, "A Conversant Robotic Guide to Art Collections". Proceedings of the *2nd Workshop on Language Technology for Cultural Heritage Data*, *6th International Conference on Language Resources and Evaluation* (LREC 2008), Marrakech, Morocco, 2008.

C28. J. Oberlander, G. Karakatsiotis, A. Isard and I. Androutsopoulos, "Building an Adaptive Museum Gallery in Second Life". Proceedings of *Museums and the Web*, Montreal, Quebec, Canada, 2008.

C27. P. Malakasiotis and I. Androutsopoulos, "Learning Textual Entailment using SVMs and String Similarity Measures". Proceedings of the *ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*, *45th Annual Meeting of the Association for Computational Linguistics* (ACL 2007), Prague, Czech Republic, pp. 42–47, 2007.

C26. D.K. Vassilakis, I. Androutsopoulos and E.F. Mageirou, "A Game-Theoretic Investigation of the Effect of Human Interactive Proofs on Spam E-mail". Proceedings of the *4th Conference on Email and Anti-Spam* (CEAS 2007), Microsoft Research Silicon Valley, Mountain View, CA, USA, 2007.

C25. D. Galanis and I. Androutsopoulos, "Generating Multilingual Descriptions from Linguistically Annotated OWL Ontologies: the NaturalOWL System". Proceedings of the *11th European Workshop on Natural Language Generation* (ENLG 2007), Schloss Dagstuhl, Germany, pp. 143–146, 2007.

C24. G. Tsatsaronis, M. Vazirgiannis and I. Androutsopoulos, "Word Sense Disambiguation with Spreading Activation Networks Generated from Thesauri". Proceedings of the *20th International Joint Conference on Artificial Intelligence* (IJCAI 2007), Hyderabad, India, pp. 1725-1730, 2007.

C23. V. Metsis I. Androutsopoulos and G. Paliouras, "Spam Filtering with Naive Bayes – Which Naive Bayes?". Proceedings of the *3rd Conference on Email and Anti-Spam* (CEAS 2006), Microsoft Research Silicon Valley, Mountain View, CA, USA, 2006.

C22. G. Lucarelli and I. Androutsopoulos, "A Greek Named-Entity Recognizer that Uses Support Vector Machines and Active Learning". Proceedings of the *4th Hellenic Conference on Artificial Intelligence* (SETN 2006), Herakleion, Crete, Greece, 2006.

C21. I. Androutsopoulos and D. Galanis, "A Practically Unsupervised Learning Method to Identify Single-Snippet Answers to Definition Questions on the Web". Proceedings of the *Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing* (HLT/EMNLP 2005), Vancouver, Canada, pp. 323–330, 2005.

C20. I. Androutsopoulos, E.F. Mageirou and D.K. Vassilakis, "A Game Theoretic Model of Spam E-Mailing". Proceedings of the *2nd Conference on Email and Anti-Spam* (CEAS 2005), Stanford University, CA, USA, 2005.

C19. I. Androutsopoulos, S. Kallonis and V. Karkaletsis, "Exploiting OWL Ontologies in the Multilingual Generation of Object Descriptions". Proceedings of the *10th European Workshop on Natural Language Generation* (ENLG 2005), Aberdeen, UK, pp. 150–155, 2005.

C18. S. Miliaraki and I. Androutsopoulos, "Learning to Identify Single-Snippet Answers to Definition Questions". Proceedings of the *20th International Conference on Computational Linguistics* (COLING 2004), Geneva, Switzerland, pp. 1360– 1366, 2004.

C17. E. Michelakis, I. Androutsopoulos, G. Paliouras, G. Sakkis and P. Stamatopoulos, "Filtron: A Learning-Based Anti-Spam Filter". Proceedings of the *1st Conference on Email and Anti-Spam* (CEAS 2004), Mountain View, CA, USA, 2004.

C16. A. Dimitromanolaki and I. Androutsopoulos, "Learning to Order Facts for Discourse Planning in Natural Language Generation". Proceedings of the *9th European Workshop on Natural Language Generation*, *10th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2003), Budapest, Hungary, pp. 23–30, 2003.

C15. I. Androutsopoulos, D. Spiliotopoulos, K. Stamatakis, A. Dimitromanolaki, V. Karkaletsis and C.D. Spyropoulos, "Symbolic Authoring for Multilingual Natural Language Generation". In I.P. Vlahavas and C.D. Spyropoulos (Eds.), *Methods and Applications of Artificial Intelligence* (proceedings of the *2nd Hellenic Conference on Artificial Intelligence –*

SETN 2002), Thessaloniki, Greece, pp. 131–142. Lecture Notes in Artificial Intelligence 2308, Springer, 2002.

C14. G. Petasis, V. Karkaletsis, G. Paliouras, I. Androutsopoulos and C.D. Spyropoulos, "Ellogon: A New Text Engineering Platform". Proceedings of the *3rd International Conference on Language Resources and Evaluation* (LREC 2002), Las Palmas, Canary Islands, Spain, pp. 72–78, 2002.

C13. G. Petasis, V. Karkaletsis, D. Farmakiotou, I. Androutsopoulos and C.D. Spyropoulos, "A Greek Morphological Lexicon and its Exploitation by a Greek Controlled Language Checker". Proceedings of the *8th Panhellenic Conference in Informatics* (EPY 2001), Nicosia, Cyprus, vol. 1, pp. 80–89, 2001.

C12. G. Sakkis, I. Androutsopoulos, G. Paliouras, V. Karkaletsis, C.D. Spyropoulos and P. Stamatopoulos, "Stacking Classifiers for Anti-Spam Filtering of E-Mail". Proceedings of the *6th Conference on Empirical Methods in Natural Language Processing* (EMNLP 2001), pp. 44–50, Carnegie Mellon University, Pittsburgh, USA, 2001.

C11. I. Androutsopoulos, J. Koutsias, K.V. Chandrinos and C.D. Spyropoulos, "An Experimental Comparison of Naive Bayesian and Keyword-Based Anti-Spam Filtering with Personal E-Mail Messages". Proceedings of the *23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (SIGIR 2000), Athens, Greece, pp. 160–167, 2000.

C10. I. Androutsopoulos and R. Dale, "Selectional Restrictions in HPSG". Proceedings of the *18th International Conference on Computational Linguistics* (COLING 2000), Saarbrucken, Germany, pp. 15–20, 2000.

C9. I. Androutsopoulos, J. Koutsias, K.V. Chandrinos, G. Paliouras and C.D. Spyropoulos, "An Evaluation of Naive Bayesian Anti-Spam Filtering". Proceedings of the *Workshop on Machine Learning in the New Information Age*, *11th European Conference on Machine Learning* (ECML 2000), Barcelona, Spain, pp. 9–17, 2000.

C8. I. Androutsopoulos, G. Paliouras, V. Karkaletsis, G. Sakkis, C.D. Spyropoulos and P. Stamatopoulos, "Learning to Filter Spam E-mail: A Comparison of a Naive Bayesian and a Memory Based-Approach". Proceedings of the *Workshop on Machine Learning and Textual Information Access*, *4th European Conference on Principles and Practice of Knowledge Discovery in Databases* (PKDD 2000), Lyon, France, pp. 1–13, 2000.

C7. K.V. Chandrinos, I. Androutsopoulos, G. Paliouras and C.D. Spyropoulos, "Automatic Web Rating: Filtering Obscene Content on the Web". Proceedings of the *4th European Conference on Research and Advanced Technology for Digital Libraries* (ECDL 2000), Lisbon, Portugal, pp. 403–406, 2000.

C6. G. Paliouras, V. Karkaletsis, I. Androutsopoulos and C.D. Spyropoulos, "Learning Rules for Large-Vocabulary Word Sense Disambiguation: a Comparison of Various Classifiers". Proceedings of the *2nd International Conference on Natural Language Processing* (NLP 2000), Patra, Greece. Lecture Notes in Artificial Intelligence, 1835, pp. 383–394, Springer, 2000.

C5. I. Androutsopoulos, "Temporal Meaning Representations in a Natural Language Front-End". In M. Gergatsoulis and P. Rondogiannis (Eds.), *Intensional Programming II* (Proceedings of the *12th International Symposium on Languages for Intensional Programming*, Athens, 1999), pp. 197–213, World Scientific, 2000.

C4. G. Petasis, G. Paliouras, V. Karkaletsis, C.D. Spyropoulos and I. Androutsopoulos, "Resolving Part-of-Speech Ambiguity in the Greek Language Using Learning Techniques". *Machine Learning in Human Language Technology* (Proceedings of the ACAI Workshop), N. Fakotakis et al. (Eds.), pp. 29–34, Chania, Greece, 1999.

C3. I. Androutsopoulos, G.D. Ritchie and P. Thanisch, "A Framework for Natural Language Interfaces to Temporal Databases". Proceedings of the *20th Australasian Computer Science Conference*, Sydney, Australia. Australian Computer Science Communications, 19(1):307–315, 1997.

C2. I. Androutsopoulos, G.D. Ritchie and P. Thanisch, "Experience Using TSQL2 in a Natural Language Interface". In J. Clifford and A. Tuzhilin (Eds.), *Recent Advances in Temporal Databases* (Proceedings of the *International Workshop on Temporal Databases*, Zurich, Switzerland), pp. 113–132. Springer-Verlag, Workshops in Computing Series, 1995.

C1. I. Androutsopoulos, G.D. Ritchie and P. Thanisch, "MASQUE/SQL: an Efficient and Portable Natural Language Query Interface for Relational Databases". Proceedings of the *6th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*, pp. 327–330, Edinburgh, UK, 1993.

7.5 System demonstrations at conferences

S6. L. Loukas, N. Smyrnioudis, C. Dikonomaki, S. Barbakos, A. Toumazatos, J. Koutsikakis, M. Kyriakakis, M. Georgiou, S. Vassos, J. Pavlopoulos and I. Androutsopoulos, "GR-NLP-TOOLKIT: An Open-Source NLP Toolkit for Modern Greek". System demonstration, *31st International Conference on Computational Linguistics* (COLING 2025), Abu Dhabi, UAE, 2025.

S5. D. Galanis, G. Karakatsiotis, G. Lampouras and I. Androutsopoulos, "An Open-Source Natural Language Generator for OWL Ontologies and its Use in Protégé and Second Life". System demonstration, *12th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2009), Athens, Greece, 2009.

S4. S. Konstantopoulos, A. Tegos, D. Bilidas, I. Androutsopoulos, G. Lampouras, P. Malakasiotis, C. Matheson and O. Deroo, "Adaptive Natural Language Interaction". System demonstration, *12th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2009), Athens, Greece, 2009.

S3. G. Karakatsiotis, D. Galanis, G. Lampouras and I. Androutsopoulos, "NaturalOWL: Generating Texts from OWL Ontologies in Protege and in Second Life". System demonstration, *18th European Conference on Artificial Intelligence* (ECAI 2008), Patras, Greece, 2008.

S2. S. Konstantopoulos, I. Androutsopoulos, H. Baltzakis, V. Karkaletsis, C. Matheson, A. Tegos and P. Trahanias, "INDIGO: Interaction with Personality and Dialogue Enabled Robots". System demonstration, *18th European Conference on Artificial Intelligence* (ECAI 2008), Patras, Greece, 2008.

S1. V. Karkaletsis, G. Samaritakis, G. Petasis, D. Farmakiotou, I. Androutsopoulos, S. Markantonatou and C.D. Spyropoulos, "A Controlled Language Checker Based on the Ellogon Text Engineering Platform". System demonstration, *2nd Meeting of the North American Chapter of the Association for Computational Linguistics* (NAACL 2001), Carnegie Mellon University, Pittsburgh, USA, pp. 74–75 (companion volume), 2001.

7.6 Recent public talks and articles in the press

See https://www.aueb.gr/users/ion/publications.html - press.

8. PARTICIPATION IN RESEARCH PROJECTS

1. "AIRIS – Mechanism-Informed Multimodal Generative AI for Causal and Dynamical Modelling in Biomedical Research". European Horizon-RIA project (2026–2030). Participating as Researcher of "Archimedes", Research Center "Athena", Athens, Greece.

2. "LLM Mediators – Exploring LLM-based Mediators in Online Discussions". Research project funded by "Archimedes", Research Center "Athena", Athens, Greece (2024–25). Co-Principal Investigator.

3. "Deductive Reasoning in Natural Language". Research project funded by "Archimedes", Research Center "Athena", Athens, Greece (2023–27). Principal Investigator.

4. "Deep Learning for Multilingual Cross-Domain Spoken Dialog Systems in Call Centers". Industrial PhDs program, funded by the Hellenic Federation of Enterprises (SEB) and AUEB (2023–27). Principal Investigator.

5. "Studying toxicity classifiers". Unrestricted research support project funded by Google (2022–23). Principal Investigator.

6. "Creation of AUEB's repository in the CLARIN:EL infrastructure". Subcontract of R.C. "Athena" in the project "National Infrastructure for the Digital Humanities and Arts and Language Research and Innovation" (2021). Principal Investigator of AUEB.

7. "Design, development and validation of the natural language processing modules and training of the models and algorithms". Subcontract in the national Research-Create-Innovate (2nd cycle) project "IntelPact - Novel approach to the management and monitoring of contractual obligations" of the company Cognitiv+ (2020–22). Principal Investigator of AUEB.

8. "ConvAI – Context-aware Abusive Language Detection in Online Conversations". Unrestricted research support project funded by Google (2018–19). Principal Investigator.

9. "ML2P Filter – Machine Learning to Filter Profanity". Google Digitial News Initiative (DNI) project (2016–18). Senior research advisor.

10. "BioQA – Biomedical Question Answering". AUEB Research Support Programme (2014–15). Project coordinator.

11. "BioASQ – A Challenge on Large-Scale Biomedical Semantic Indexing and Question Answering". European FP7 ICT project (2012–14). Principal Investigator of AUEB.

12. "Natural Language Interaction with Semantic Web ontologies" (2010–13). National Heracleitus-II project. Project coordinator.

13. "A Linear Programming Approach to Multi-Document Text Summarization and Natural Language Generation from Ontologies" (2011–12). AUEB Basic Research Funding Program (BRFP) project. Project coordinator.

14. "INDIGO – Interaction with Personality and Dialogue Enabled Robots" (2007–10). European IST project. Principal Investigator of AUEB.

15. “XENIOS: User-Robot Interaction Based on Visual Perception, Speech Processing, and Natural Language Generation for Use in Robotic Guidance Systems” (2006–07). Project of the Greek “Information Society” programme. Principal Investigator of AUEB.

16. “Combined Research in the Areas of Information Retrieval, Natural Language Processing, and User Modeling Aiming at the Development of Advanced Search engines for Document Collections” (2006–08). Greek PENED project.

17. “CROSSMARC: Cross-lingual Multi-Agent Retail Comparison” (2001–03). European IST project.

18. “M-PIRO: Multi-lingual Personalised Information Objects” (2000–03). European IST project.

19. “HYGEIOROBOT: Navigation and Control of a Health Service Robot” (2000–01). Greek PENED project.

20. “FILTERIX: Web Page Filtering System” (2000–04). Internal project of NCSR “Demokritos”.

21. “MITOS: Information Retrieval and Mining System – Application to Financial News” (1999–2001). Greek EPET project.

22. “SCHEMATOPOIESIS: Integrated Environment for the Development and Utilization of Sublanguages of Greek” (1999–2001). National EPET project.

23. “ECRAN: Extraction of Content – Research at Near-Market” (1995–1999). European Language Engineering project.

9. PATENTS

- E. Loukas, E. Spyropoulou, P. Malakasiotis, M. Fergadiotis, I. Chalkidis, I. Androutsopoulos, and G. Paliouras. System and method for automatically tagging documents. United States Patent and Trademark Office (USPTO) Patent US17/873,932 (**granted**), 2025. Also, European Office (EPO) Patent Application 21 386 048.9 (pending) and World Intellectual Property Organization (WIPO) Patent Application WO2023006773A1 (pending).

- C.D. Spyropoulos, K. Chandrinos, G. Paliouras, E. Karkaletsis and I. Androutsopoulos, “Method for Probabilistic Information Fusion to Filter Multi-Lingual, Semi-Structured and Multimedia Electronic Content”. Patent no. 1005379, Greek Industrial Property Organization, 2006. Submitted to the European Patent Office.

10. OTHER ACTIVITIES AND INFORMATION

- **Citations: 22,863 according to Google Scholar, not excluding self-citations, h-index: 57 (June 2026).**
- Founder and co-director of AUEB’s Natural Language Processing Group (2002–today).

- Head (2024–today) and Deputy Head (2017–2020) of AUEB’s Department of Informatics.

- President (2025–today) of the Sectoral Scientific Council on Artificial Intelligence (SSC AI) of the Greek National Council for Research, Technology and Innovation (NCRTI).

- Lead Researcher of the "Archimedes" Research Unit, "Athena" Research Centre, Athens, Greece (2023–present).

- Director of the Information Systems and Databases Sector, Department of Informatics, AUEB (2019–2021 and 2023–2024).

- Director of the Information Processing Laboratory, Department of Informatics, AUEB (2018–2021 and 2024).

- Founding head of AUEB's e-Learning Unit (2015–2017).

- Member of the Scientific Committee of the Lifelong Learning Center (KEDIVIM) of AUEB (2017–2024).

- Research Associate, Software and Knowledge Engineering Lab, Institute of Informatics and Telecommunications, NCSR "Demokritos", Athens, Greece (2019–2023).

- Adjunct Researcher of the Digital Curation Unit, Institute for the Management of Information Systems (IMIS), "Athena" Research Centre, Athens, Greece (2007–present).

- Senior Research Advisor of StrainTek (http://www.straintek.com/about), a start-up that received funding from Google's Digital News Initiative (2016–18).

- Co-organizer of the following conferences, workshops, competitions, summer schools:

  1. 5th *Athens Natural Language Processing Summer School* (AthNLP 2026), Athens, Greece, 2026. **Co-organizer.**
  2. *4th Athens Natural Language Processing Summer School* (AthNLP 2025), Athens, Greece, 2025. **Co-organizer.**
  3. *3rd Athens Natural Language Processing Summer School* (AthNLP 2025), Athens, Greece, 2024. **Co-organizer.**
  4. *SemEval 2025 Task 10 "Multilingual Characterization and Extraction of Narratives from Online News".* **Co-organizer.**
  5. *2nd Athens Natural Language Processing Summer School* (AthNLP 2024), Athens, Greece, 2024. **Co-organizer.**
  6. *3rd Workshop on Natural Legal Language Processing* (NLLP 2021) at the *2021 Conference on Empirical Methods in Natural Language Processing* (EMNLP 2021). Online and Punta Cana, Dominican Republic. **Program Committee co-chair.**
  7. *SemEval 2021 Task 5: Toxic Spans Detection.* **Co-organizer.**
  8. *2nd Workshop on Natural Legal Language Processing* (NLLP 2020) of the *2020 Knowledge Discovery in Databases* conference (KDD 2020). Held online. **Program committee co-chair.**
  9. The 11th EETN (Greek) Conference on Artificial Intelligence (SETN 2020). Held online. **Program committee co-chair.**
  10. *17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (NAACL-HLT 2019), Minneapolis, USA, 2019. **Co-chair of the Mixed Topics area**.
  11. *1st Athens Natural Language Processing Summer School* (AthNLP 2019), Athens, Greece, 2019. **Co-organizer.**
  12. *SemEval-2016 Task 5: Aspect Based Sentiment Analysis.* **Co-organizer.**
  13. *SemEval-2015 Task 12: Aspect Based Sentiment Analysis.* **Co-organizer.**
  14. *SemEval-2014 Task 4: Aspect Based Sentiment Analysis.* **Co-organizer.**
  15. *Joint ECML/PKDD – PASCAL Workshop on Large-Scale Hierarchical Classification* (LSHC), *European Conference on Machine Learning and Principles and Practice of*

Knowledge Discovery in Databases (ECML PKDD), Athens, Greece, 2011. **Co-organizer.**

16. *12th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2009), Athens, Greece, 2009. **Co-organizer.**
17. *3rd Conference on Email and Anti-Spam* (CEAS 2006), Microsoft Research Silicon Valley, Mountain View, CA, U.S.A., 2006. **Program Committee co-chair**.
18. *Workshop on End User Semantic Web Interaction, 4th International Semantic Web Conference* (ISWC 2005), Galway, Ireland, 2005. **Co-chair of the Organizing Committee.**
19. *12th International Symposium on Languages for Intensional Programming* (ISLIP 1999), Athens, Greece, 1999. **Co-organizer.**

- Reviewer of the journals: *Journal of Artificial Intelligence Research*, *Transactions of ACL*, *Natural Language Engineering*, *Machine Learning*, *Semantic Web Journal*, *ACM Transactions on Internet Technology*, *ACM Computing Surveys*, *Communications of the ACM*, *IEEE Transactions on Affective Computing, Research on Language and Computation*, *Applied Artificial Intelligence*, *International Journal of Artificial Intelligence Tools, Artificial Intelligence in Medicine*, *Journal of Biomedical Semantics, Computers and Security, Universal Access in the Information Society* etc.

- Frequent reviewer of the top Natural Language Processing Conferences (*ACL, EMNLP, NAACL, EACL, Coling*, LREC etc.) since 1998.

- Reviewer awards:
  - *58th Annual Meeting of the Association for Computational Linguistics* (ACL 2020), held online, 2020. **Outstanding reviewer** (218 out of approx. 2519 reviewers, https://virtual.acl2020.org/plenary_session_opening_remarks.html).
  - *15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (NAACL-HLT 2016), San Diego, CA, USA, 2016. **Best reviewers award** (http://naacl.org/naacl-hlt-2016/best_reviewers.html).
  - *2012 Conference on Empirical Methods in Natural Language Processing and Natural Language Learning* (EMNLP-CoNLL 2012), Jesu Island, Korea, 2012. **Best reviewers award** (http://emnlp-conll2012.unige.ch/best-reviewers.html).
  - *Conference on Empirical Methods in Natural Language Processing* (EMNLP 2010), MIT, MA, USA, 2010. **Best reviewers award** (awarded to 86 out of 460 reviewers).

- Special Secretary of the Hellenic Artificial Intelligence Society (EETN, December 2006 – December 2010).