

---

CV

---

JOHN PAVLOPOULOS

## TABLE OF CONTENTS

---

---

<b>ID .....</b>	<b>3</b>
<b>RESEARCH INTERESTS.....</b>	<b>4</b>
<b>EDUCATION .....</b>	<b>5</b>
<b>PROFESSIONAL EXPERIENCE .....</b>	<b>6</b>
<b>TEACHING EXPERIENCE .....</b>	<b>7</b>
<b>SCHOLARSHIPS AND AWARDS .....</b>	<b>8</b>
<b>PUBLICATIONS.....</b>	<b>9</b>
<b>RESEARCH PROJECTS.....</b>	<b>13</b>
<b>MISCELLANEOUS SCIENTIFIC ACHIEVEMENTS.....</b>	<b>14</b>

## ID

---

Name/Surname: Ioannis (John) Pavlopoulos  
Date of Birth: 6 August 1983  
Email: [ipavlopoulos@aueb.gr](mailto:ipavlopoulos@aueb.gr)  
Website: <https://ipavlopoulos.github.io>

## RESEARCH INTERESTS

---

My research is focused on Natural Language Processing, specifically on:

- **Machine** and **deep learning** for text and images
- **Data mining** (emotion, named entities) from large collections
- **Classification** for early diagnosis or filtering
- **Generation** of natural language.

## EDUCATION

---

2010 – 2014: Department of Informatics, Athens University of Economics and Business, Greece. PhD in Natural Language Processing titled as “Aspect Based Sentiment Analysis” ([link](#)).

2008 – 2009: Department of Informatics, University of Edinburgh, U.K. MSc in Artificial Intelligence.

2001 – 2007: School of Applied Mathematics and Physical Sciences, National Technical University of Athens, Greece.

## PROFESSIONAL EXPERIENCE

---

- [Department of Informatics, Athens University of Economics and Business](#)
  - Assistant Professor
  - Machine learning in information retrieval and data mining for large document collections
- [Ca'Foscari University of Venice, Department of Humanities](#)
  - Visiting Scholar in NLP (1/12/2021 – 31/8/2022).
  - Co-organisation of the [HTREC](#) academic challenge
  - Member of the editorial board of the [magazén international journal](#).
- [Stockholm University, Department of Computer and System Sciences](#)
  - Senior Lecturer (fixed-term) in NLP (1/6/2020 – 30/11/2021).
  - Postdoctoral Researcher (1/10/2019 – 31/5/2020).
- [Google Jigsaw](#)
  - Research associate (1/10/2018 -- today).
- [Cognitiv+](#)
  - Data Scientist (1/9/2018 – 1/4/2019)
    - Layout and skew detection in images of legal documents.
    - Data annotation manager.
- [Liquid Media](#) / Straintek
  - Senior Researcher (1/7/2016 – 1/7/2018)
    - Leading the team developing APIs for abusive language detection.
    - Multilingual abusive language detection.
    - Deep learning.
  - Postdoctoral Researcher (1/7/2015 – 1/7/2016)
    - Machine learning for abusive language detection.
    - Lead author of funding proposals.
- [Hellenic Army](#) (Co11, Co12)
  - Software Development (1/9/2014 – 1/6/2015).
- [BioASQ](#) (1/9/2012 – 1/9/2014)
  - Study of evaluation measures for biomedical QA systems.
  - Development of word embeddings from biomedical corpora ([available](#)).
- [IMC Technologies](#)
  - R&D Engineer (1/9/2010 – 1/9/2012)
    - Text search and web crawling,
    - Data mining for e-Government.
- DBNet, AUEB.
  - Scientific Associate (1/9/2009 – 1/9/2010)
    - Application of Markov models on web time-series data.
- [NCSR Demokritos](#) (1/7/2007 – 1/8/2008)
  - Scientific Associate.
    - Development of Neural Networks.
    - Co-authoring of EU proposals.
- [Oracle Hellas](#)
  - Internship in software development (1/10/2005 – 1/12/2005)

## TEACHING EXPERIENCE

---

### Teaching:

- Natural Language Processing, MSc in Digital Methods for the Humanities, Athens University of Economics and Business, Greece.
- Practical Data Science, MS in Data Science, Athens University of Economics and Business, Greece.
- Natural Language Processing, PhD course, Stockholm University, Sweden.
- Visiting Lecturer in Data Science, Stockholm University, Sweden.

### Co-teaching:

- Natural Language Processing, MSc in Digital Methods for the Humanities, Athens University of Economics and Business, Greece.

## SCHOLARSHIPS AND AWARDS

---

- Secured a 'Google Research Award' for postdoctoral studies.
- 'PriceWaterhouseCoopers' scholarship for postgraduate studies.
- 'Thomaidio Award' for the best BSc thesis in the University.
- 'Ganioti-Papageorgi' scholarship for undergraduate studies.
- 'Aglaias Koufodimou' scholarship for undergraduate studies.

## PUBLICATIONS

---

### Peer-reviewed scientific articles published in conference proceedings:

- Σ1. Vrettaros J., **Pavlopoulos J.**, Vouros G. & Drigas A. S. (2008). The development of a self-assessment system for the learners answers with the use of GPNN. *World Summit on Knowledge Society (WSKS)*, Athens, Greece, pp. 332—340.
- Σ2. Voudigari E., **Pavlopoulos J.** & Vazirgiannis M. (2011). A Framework for Web Page Rank Prediction. *Engineering Applications of Neural Networks / Artificial Intelligence Applications and Innovations*, Corfu, Greece, pp. 240—249.
- Σ3. Alexopoulos P., **Pavlopoulos J.**, Wallace, M. & Καφετζής Κ. (2011). Exploiting ontological relations for automatic semantic tag recommendation. *International Conference on Semantic Systems*, Graz, Austria, pp. 105—110.
- Σ4. Alexopoulos P., **Pavlopoulos J.** & Mylonas P. (2012). Learning vague knowledge from socially generated content in an enterprise framework. *IFIP Advances in Information and Communication Technology*, Artificial Intelligence Applications and Innovations, Halkidiki, Greece, pp. 510—519.
- Σ5. Anadiotis G., Kafentzis K., **Pavlopoulos J.** & Westerski A. (2012). Building consensus via a semantic web collaborative space. WWW, Lyon, France, pp. 1097—1106.
- Σ6. **Pavlopoulos J.** & Androutsopoulos I. (2014). Multi-granular aspect aggregation in aspect-based sentiment analysis. *EACL*, Gothenburg, Sweden, pp. 78—87.
- Σ7. Karampatsis P. M., Pavlopoulos J. & Malakasiotis P. (2014). AUEB: Two stage sentiment analysis of social network messages. *SemEval in ACL*, Dublin, Ireland, pp. 114—118.
- Σ8. **Pavlopoulos J.** & Androutsopoulos I. (2014). Aspect term extraction for sentiment analysis: New datasets, new evaluation measures and an improved unsupervised method. *Workshop on Language Analysis for Social Media in EACL*, Gothenburg, Sweden, pp. 44—52.
- Σ9. Pontiki M., Galanis D., **Pavlopoulos J.**, Papageorgiou C., Androutsopoulos I. & Manandhar, S. (2014). Semeval-2014 task 4: Aspect based sentiment analysis. *SemEval in COLING*, Dublin, Ireland, pp. 27—35.
- Σ10. Xenos D., Theodorakakos P., **Pavlopoulos J.**, Malakasiotis P. & Androutsopoulos I. (2016). AUEB-ABSA at SemEval-2016 Task 5: Supervised Machine Learning for Aspect Based Sentiment Analysis. *SemEval in NAACL:HLT*, San Diego, California, USA, pp. 312—317
- Σ11. Giorgis S., Rousas A., **Pavlopoulos J.**, Malakasiotis P. & Androutsopoulos I. (2016). Aueb.twitter.sentiment at SemEval-2016 Task 4: A weighted ensemble of SVMs for Twitter sentiment analysis. *SemEval in NAACL:HLT*, San Diego, California, USA, pp. 96—99.
- Σ12. **Pavlopoulos J.**, Malakasiotis P. & Androutsopoulos I. (2017). Deeper Attention to abusive user comment moderation. *EMNLP*, Copenhagen, Denmark, pp. 1125—1135.
- Σ13. **Pavlopoulos J.**, Malakasiotis P. & Androutsopoulos I. (2017). Deep Learning for User Comment Moderation. *Abusive Language Workshop in ACL*, Vancouver, BC, Canada, pp. 25—35.

- Σ14. **Pavlopoulos J.**, Malakasiotis P., Bakagianni G. & Androutsopoulos I. (2017). Improved Abusive Comment Moderation with User Embeddings. *Natural Language Processing meets Journalism Workshop in EMNLP*, Copenhagen, Denmark, pp. 51–55.
- Σ15. Kougia V., **Pavlopoulos J.** & Androutsopoulos I. (2019). A Survey of Biomedical Image Captioning. *Shortcomings in Vision and Language NAACL Workshop*, Minneapolis, MN, USA, pp. 26-36.
- Σ16. Kougia V., **Pavlopoulos J.** & Androutsopoulos I. (2019). AUEB NLP Group at ImageCLEFmed Caption 2019. *International Conference and Labs of the Evaluation Forum for European Languages*.
- Σ17. Kougia V., **Pavlopoulos J.** & Androutsopoulos I. (2020). Medical Image Tagging with Neural Retrieval and Deep Learning. *LNCS* (online), pp. 154-166.
- Σ18. **Pavlopoulos J.** & Papapetrou P. (2020). Clinical Predictive Keyboard using Statistical and Neural Language Modeling. *IEEE Computer Based Medical Systems* (online), pp. 293-296.
- Σ19. **Pavlopoulos J.** & Sorensen, J. & Dixon, L. & Thain, N. & Androutsopoulos I. (2020). Toxicity Detection: Does Context Really Matter? *ACL* (online), pp. 4296-4305.
- Σ20. Karatzas V., Kougia V., **Pavlopoulos J.** & Androutsopoulos I. (2020). AUEB NLP Group at ImageCLEFmed Caption 2020. *International Conference and Labs of the Evaluation Forum for European Languages* (online).
- Σ21. Korre K. & **Pavlopoulos J.** (2020). ERRANT: Assessing and Improving Grammatical Error Type Classification. *Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (online), pp. 85-89.
- Σ22. Laughier L., **Pavlopoulos J.**, Sorensen J. & Dixon L. (2021). Civil Rephrases Of Toxic Texts With Self-Supervised Transformers. *EACL* (online), pp. 1442-1461.
- Σ23. **Pavlopoulos J.**, Laughier L, Sorensen J. & Androutsopoulos I. (2021). Semeval-2021 task 5: Toxic Spans Detection. *SemEval in ACL-IJCNLP* (online), pp. 59-69.
- Σ24. **Pavlopoulos J.** & Papapetrou P. (2021). Customized Neural Predictive Medical Text:A Use-Case on Caregivers. *International Conference on Artificial Intelligence in Medicine* (online), pp. 438-443.
- Σ25. Xenos A., **Pavlopoulos J.** (corresponding author) & Androutsopoulos I. (2021). Context Sensitivity Estimation in Toxicity Detection. *Workshop on Online Abuse and Harms in ACL* (online), pp. 140-145.
- Σ26. Korre K., Chatzipanagiotou M. & **Pavlopoulos J.** (2021) ELERRANT: Automatic Grammatical Error Type Classification for Greek. *RANLP* (online), pp. 713-722.
- Σ27. Kougia V. & **Pavlopoulos J.** (2021) Multimodal or Text? Retrieval or BERT? Benchmarking Classifiers for the Shared Task on Hateful Memes. *Workshop on Online Abuse and Harms (WOAH 2021)*, (online), pp. 220-225.
- Σ28. Miliou I., **Pavlopoulos J.** & Papapetrou P. (2021). Sentiment Nowcasting During the COVID-19 Pandemic. *International Conference on Discovery Science*. Springer, Cham. (online), pp. 218-228.

- Σ29. Ljungman, J., Lislevand, V., **Pavlopoulos J.**, Farazouli A., Lee, Z., Papapetrou P. & Fors U. (2021). Automated Grading of Exam Responses: An Extensive Classification Benchmark. International Conference on Discovery Science. Springer, Cham. (online), pp. 3-18.
- Σ30. Fasoi M., **Pavlopoulos J.** & Konstantinidou M. (2021). Computational Authorship Analysis of Homeric Language. Digital Humanities Workshop. (online), pp. 78-88, <https://dl.acm.org/doi/10.1145/3526242.3526256>
- Σ31. Chatzipanagiotou M., Ewa Machotka & **Pavlopoulos J.** (2021). Automated recognition of geographical named entities in titles of Ukiyo-e prints. Digital Humanities Workshop. (online), pp. 70-77, <https://doi.org/10.1145/3526242.3526254>.
- Σ32. **Pavlopoulos J.**, Laugier L., Xenos A., Sorensen J. & Androutsopoulos I. (2022). From the Detection of Toxic Spans in Online Discussions to the Analysis of Toxic-to-Civil Transfer. Association for Computational Linguistics. (online), pp. 3721-3734
- Σ33. **Pavlopoulos J.**, Machotka E. & Liagkou K. (2022). A Study of Distant Viewing of ukiyo-e prints. Language Resources and Evaluation Conference, Μασσαλία, Γαλλία, pp. 5879-5888.
- Σ34. Korre K. & **Pavlopoulos J.** (2022). Enriching Grammatical Error Correction Resources for Modern Greek. Language Resources and Evaluation Conference, Μασσαλία, Γαλλία, pp. 4984—4991.
- Σ35. Platanou P., **Pavlopoulos J.** & Papaioannou G. (2022). Handwritten Paleographic Greek Text Recognition: A Century-Based Approach. Language Resources and Evaluation Conference, Μασσαλία, Γαλλία, pp. 6585-6589.
- Σ36. **Pavlopoulos J.**, Χένοζ Α. & Picca D. (2022). Sentiment Analysis of Homeric Text: The 1st Book of Iliad. Language Resources and Evaluation Conference, Μασσαλία, Γαλλία, pp. 7071-7077.
- Σ37. Zonios C., **Pavlopoulos J.** & Likas A. (2022). Transformer-Based Music Language Modelling and Transcription. Proceedings of the 12th Hellenic Conference on Artificial Intelligence. Corfu, Greece, pp. 1-8.
- Σ38. Dritsa K., Thoma A., **Pavlopoulos J.** & Louridas P. (2022). A Greek Parliament Proceedings Dataset for Computational Linguistics and Political Analysis. Advances in Neural Information Processing Systems, 35, pp. 28874-28888.
- Σ39. **Pavlopoulos J.**, & Likas A. (2024). Polarized Opinion Detection Improves the Detection of Toxic Language. In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 1946-1958.

#### Peer-reviewed journals:

- Π1. Vrettaros J., **Pavlopoulos J.**, Drigas A. Σ. & Hrisagis K., “GPNN techniques in learning assessment systems”. the International Journal of Technology Enhanced Learning. Inderscience Publishers, pp. 415-429, 3(4), 2011.

- Π2. Alexopoulos P., **Pavlopoulos J.** & Mylonas P., "A Semantic Framework for Harvesting Vague Enterprise Knowledge from Microposts". *International Journal on Artificial Intelligence Tools* 23(02), 2014.
- Π3. Tsatsaronis G., Balikas G., Malakasiotis P., Partalas I., Zschunke M., Alvers, M., Weissenborn D., Krithara A., Petridis S., Polychronopoulos D., Almyrantis Y., **Pavlopoulos J.**, Baskiotis N., Gallinari R., Artieres T., Ngomo A.-C., Heino N., Gaussier E.; Barrio-Alvers L., Schroeder M., Androutsopoulos I. & Paliouras G. (2015). An overview of the bioasq large-scale biomedical semantic indexing and question answering competition. *BMC Bioinformatics* 16(1).
- Π4. Kougia V., **Pavlopoulos J.** (corresponding author), Papapetrou P. (2021). RTEx: A novel framework for Ranking, Tagging, and Explanatory diagnostic captioning of radiography exams. *Journal of the American Medical Informatics Association*, volume 28, issue 8, August 2021, pp. 1651-1659, <https://doi.org/10.1093/jamia/ocab046>.
- Π5. Olczak J., **Pavlopoulos J.**, Prijs J., Ijpmma F. F. A., Doornberg J. N., Lundstrom C., Hedlund J., and Gordon M. (2021). Presenting artificial intelligence, deep learning, and machine learning studies to clinicians and healthcare stakeholders: an introductory reference with a guideline and a Clinical AI Research (CAIR) checklist proposal. *Acta Orthopaedica*.
- Π6. **Pavlopoulos J.**, Kougia V., Androutsopoulos I., Papamichail D. (2022). Diagnostic Captioning: A survey. *Knowledge and Information Systems*. 2022 Jun 16:1-32. <https://doi.org/10.1007/s10115-022-01684-7>
- Π7. Assael Y., Sommerschild T., Shillingford B. Bordbar M., **Pavlopoulos J.**, Chatzipanagiotou M., Androutsopoulos I., Prag J., de Freitas N. Restoring and attributing ancient texts using deep neural networks. *Nature* 603, pp. 280-283 (2022). <https://doi.org/10.1038/s41586-022-04448-z>
- Π8. **Pavlopoulos J.**, Konstantinidou M. (2022). Computational Authorship Analysis of the Homeric Poems. *International Journal of Digital Humanities*, pp. 1-20, <https://doi.org/10.1007/s42803-022-00046-7>
- Π9. Xenos A., **Pavlopoulos J.**, Androutsopoulos I., Dixon L., Sorensen J., Laugier L. (2022). Toxicity detection sensitive to conversational context. *First Monday*. <https://doi.org/10.5210/fm.v27i5.12285>.
- Π10. **Pavlopoulos J.**, Likas A. (2023) Distance from Unimodality for the Assessment of Opinion Polarization. *Cognitive Computation*. <https://doi.org/10.1007/s12559-022-10088-2>.
- Π11. Sommerschild T., Assael Y., **Pavlopoulos J.** (equal contribution), et al. (2023). Machine learning for ancient languages: A survey. *Computational Linguistics*, 49(3), pp.703-747.
- Π12. Boulieris P., **Pavlopoulos J.**, Xenos A., Vassalos V. (2023). Fraud detection with natural language processing. *Machine Learning*. Jul 19:1-22.
- Π13. **Pavlopoulos J.**, Romell A., Curman J., Steinert O., Lindgren T., Borg M., Randl K. (2024). Automotive fault nowcasting with machine learning and natural language processing. *Machine Learning*. 113(2):843-61.

## RESEARCH PROJECTS

---

- [EFRA](#): EU Horizon project on food risk analytics (2022-2025).
- [NLP for digital Art History](#): Research project on NLP for Art History (2021-2022).
- [Autograde](#): Research project for the automated grading of electronic exams (2021).
- Research project for EU (European Committee) for the development of a cost model for user-generated content moderation online (2020). Primary Investigator.
- ConvAI - Context-aware abusive language detection in online conversations. Google unrestricted award (Google Research Award, 2019). Primary Investigator.
- [ML2P Filter](#) - Machine learning to filter profanity. Google Digital News Initiative (DNI) project. Primary Investigator (2016-2018).
- [BioASQ](#), EU FP7 ICT. Research associate (2012-2014).

## OTHER SCIENTIFIC ACHIEVEMENTS

---

Reviewer in top-tier conferences, such as NeurIPS, ACL, EMNLP, COLING, NAACL-HLT, and journals such as: *Natural Language Engineering*, *Data Mining and Knowledge Discovery*. Member of the organisation committee of the ML4AL ACL workshop and the program committee of the Abusive Language Workshop. Member of the editorial board of the *magazén* International Journal for Public and Digital Humanities.