

# Jakub Grzywaczewski

✉ jgrzywaczewski.com   ✉ jakubzgrzywaczewski@gmail.com   🌐 ZetextJG

📄 jakub-grzywaczewski-5b6801236   📞 790 200 186   📍 Warsaw, Poland

## Education

- |                                                                                                                 |                     |
|-----------------------------------------------------------------------------------------------------------------|---------------------|
| <b>Warsaw University of Technology</b> , MSc in Data Science – Warsaw, Poland                                   | Feb 2025 – present  |
| <b>Warsaw University of Technology</b> , BSc(Eng) in Data Science – Warsaw, Poland<br>Graduated Summa cum laude | Oct 2021 – Feb 2025 |

## Experience

- |                                                                                                                                                                                                                                        |                      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|
| <b>Research Software Engineer</b> , Centre for Credible AI (CCAI) – Warsaw, Poland                                                                                                                                                     | Oct 2025 – present   |
| Working on projects pushing the frontier of explainable AI, including model validation with Generative AI.                                                                                                                             |                      |
| <ul style="list-style-type: none"> <li>• Co-authored a scientific paper on explaining the decision-making process of deep lung cancer models using generative/diffusion models. Results were accepted for ICML 2026</li> </ul>         |                      |
| <b>Research Software Engineer</b> , MI2 – Warsaw, Poland                                                                                                                                                                               | June 2024 – Oct 2025 |
| Worked on explainability of deep learning models using generative and diffusion models.                                                                                                                                                |                      |
| <ul style="list-style-type: none"> <li>• Co-authored and presented a scientific paper at ICLR 2025</li> </ul>                                                                                                                          |                      |
| <b>Python Developer</b> , Samsung Research – Warsaw, Poland                                                                                                                                                                            | Aug 2022 – June 2024 |
| Worked on NLP research projects and developed internal CLI and web tools for data processing.                                                                                                                                          |                      |
| <ul style="list-style-type: none"> <li>• Developed and maintained NLP pipelines (automatic, CLI, Web) for both research and production</li> <li>• Completed time management training and applied PRINCE2 methodology basics</li> </ul> |                      |

## Publications

- |                                                                                                                                                                                                                                                                                                                                                            |      |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <b>Your CLIP has 164 dimensions of noise: Exploring the embeddings covariance eigenspectrum of contrastively pretrained vision-language transformers</b><br><b>Jakub Grzywaczewski</b> , Dawid Płudowski, Przemysław Biecek<br>(arXiv preprint arXiv:2605.14893)                                                                                           | 2026 |
| <b>Auditing Sybil: Explaining Deep Lung Cancer Risk Prediction Through Generative Interventional Attributions</b><br>Bartłomiej Sobieski, <b>Jakub Grzywaczewski</b> , Karol Dobiczek, Mateusz Wajcik, Tomasz Bartczak, Patryk Szatkowski, Matthew Tivnan, Przemysław Biecek<br>(The Forty-Third International Conference on Machine Learning (ICML 2026)) | 2026 |
| <b>Rethinking visual counterfactual explanations through region constraint</b><br>Bartłomiej Sobieski, <b>Jakub Grzywaczewski</b> , Bartłomiej Sadlej, Matthew Tivnan, Przemysław Biecek<br>(The Thirteenth International Conference on Learning Representations (ICLR 2025))                                                                              | 2024 |

## Projects

- |                                                                                                                                                                                                                             |             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>Serolyzer (PvSTATEM)</b> – <a href="https://mini-pw.github.io/SerolyzeR/">https://mini-pw.github.io/SerolyzeR/</a>                                                                                                       | 2024 – 2026 |
| an R package for automated analysis of serological data, currently used in the PvSTATEM global initiative                                                                                                                   |             |
| <ul style="list-style-type: none"> <li>• Currently maintaining the package, implemented parsing methods for many MBA data format</li> <li>• Designed and implemented a Shiny Application for non-technical users</li> </ul> |             |

High Performance Open-Source Python+C package for Covariance Matrix shrinkage and LDA

- Implemented novel covariance shrinkage methods in C, with bindings in Python

## Skills

---

**Languages:** Python, Rust, C, Java, JavaScript, TypeScript, TeX

**Data Science:** NumPy, Pandas, Scikit-learn, PyTorch (Fabric), Transformers, SHAP, vLLM, ONNX, R

**Deep Learning & AI:** Generative Models, Diffusion Models, Explainable AI, Computer Vision, NLP

**Databases:** PostgreSQL, MySQL, MS SQL Server, MongoDB, Redis, Hadoop

**Compute & Cloud:** SLURM, Kubeflow, GCP, Kubernetes

**Web Development:** Vue, Svelte, Flask, FastAPI

**DevOps & Tooling:** Git, Linux, Vim, Jenkins, GitHub Actions

## About Me

---

My journey with programming began in 2017 when I taught myself basic C++. Over the years, my dedication and hard work have significantly expanded my skill set. In 2022, I joined Samsung Research as a Python Developer. My interest in machine learning was quickly recognized, and I have contributed to numerous Natural Language Processing (NLP) projects, both in research and implementation. Additionally, I developed internal command-line interfaces (CLI) and web tools for data processing. For nearly two years, I worked as a Research Software Engineer at MI<sup>2</sup>, where I focused on explaining the decision-making processes of deep learning models using generative models, including diffusion models. This role also allowed me to present a scientific paper at the prestigious International Conference on Learning Representations (ICLR) in 2025. Currently, I am working with the Centre for Credible AI in Warsaw. Here, I am extending methods for model validation into medical and space domains, while also deepening my knowledge in contrastive learning models. Beyond my professional and academic pursuits, I served as the Vice President of the Mathematics Science Club.

## Languages

---

**Polish:** Native speaker

**English:** C1 Proficient

## Interests

---

**Generative Models for Explainable AI:** Generative AI, Diffusion Models

**Explainable AI:** Model Validation, Interpretability, Depoisoning

**Mathematics:** Probability, Statistics, Linear Algebra

**Hobbies:** Ping-pong, Music, Cooking