# Shae McFadden

Ph.D. Candidate · Artificial Intelligence & Cybersecurity

shae.mcfadden@kcl.ac.uk | mcfadden-s.github.io | McFadden-S | shae-mcfadden-27b840234 | Shae McFadden

## Summary

Currently a Ph.D. candidate, supervised by Dr. Fabio Pierazzi and Dr. Nicola Paoletti, at King's College London, investigating the applications of deep reinforcement learning to cybersecurity. A passion for the cross-section between artificial intelligence and cybersecurity resulted in collaborations with the Systems Security Research Lab (S2Lab) at UCL and the AI for Cyber Defence Research Centre (AICD) at The Alan Turing Institute. Graduated from King's College London with a first-class B.Sc. (Hons) in Computer Science specialising in Artificial Intelligence. During undergraduate studies, was awarded the Layton Research Award, the Alan Fairbourn Memorial Prize, and the Associateship of King's College.

## Experience

### Ph.D. Candidate in Computer Science

*London, UK*

King's College London

*October 2023 - Present*

- Thesis on "The Applications of Deep Reinforcement Learning for Cybersecurity", supervised by Dr. Fabio Pierazzi.
- Received funding from a NMES Faculty Studentship.

### Visiting Researcher @ S2Lab

*London, UK*

University College London

*January 2024 - Present*

- Aided in the extension of the Tesseract framework and ACSAC Cybersecurity Artifacts Competition.

### Visiting Collaborator @ AICD

*London, UK*

The Alan Turing Institute

*October 2023 - Present*

- A member of the DARPA AI Cyber Challenge (AIxCC) team, which developed a LLM pipeline for automated vulnerability detection and repair.
- Developed a deep reinforcement learning (DRL)-based autonomous security testing tool for denial-of-service in GraphQL applications.

### King's Undergraduate Research Fellow

*London, UK*

King's College London

*July 2022 - September 2022*

- Performed evaluations on the impact of poisoning on machine learning classifiers for malware detection under the effects of concept drift.

## Education

### B.Sc. (Hons) in Computer Science (Artificial Intelligence)

*London, United Kingdom*

King's College London

*2020 - 2023*

- First Class Honours with a specialization in Artificial Intelligence.
- Dissertation on "Adversarial Machine Learning Evaluation of the MaMaDroid Feature Space", which evaluated the impact of data poisoning alongside different mitigation strategies on a malware classifier over time.

## Technical Skills

### Python Programming

Projects using Python:

- Published several papers which used Python to perform machine learning evaluations.
- Current maintainer of three open-source Python Github repositories (Wendigo [link], RPAL [link], and Tesseract [link]).

### Machine Learning & Deep Learning

Projects using machine learning:

- Several publications applying machine learning (either supervised or deep reinforcement learning) to cybersecurity problems.
- Undergraduate Dissertation focusing on the safe application of supervised learning to malware classification.

### Tech Stack

Experience from prior projects

- Python, PyTorch, Scikit-Learn, CleanRL, Stable-Baselines3, Gymnasium, Hugging Face Transformers, NumPy, Git, Docker

### Leadership & Teamwork

Experiences that have developed leadership and teamwork skills:

- Collaborating with researchers from various institutions on research projects.
- Leading 4 to 10 person teams on software development projects in undergraduate studies.
- Captain of a 12 person fencing team.

# Publications

## Published

2024 **The Impact of Active Learning on Availability Data Poisoning for Android Malware Classifiers**, **Shae McFadden**, Zeliang Kan, Lorenzo Cavallaro, Fabio Pierazzi. [paper], [code]. *ARTMAN co-located with ACSAC*

2024 **Wendigo: Deep Reinforcement Learning for Denial-of-Service Query Discovery in GraphQL**, **Shae McFadden**, Marcello Maugeri, Chris Hicks, Vasilios Mavroudis, Fabio Pierazzi. [paper], [code]. *DLSP co-located with IEEE S&P*

2023 **Poster: RPAL–Recovering Malware Classifiers from Data Poisoning using Active Learning**, **Shae McFadden**, Zeliang Kan, Lorenzo Cavallaro, Fabio Pierazzi. [paper], [code]. *ACM CCS Poster*

## Preprints

2025 **One Pic is All it Takes: Poisoning Visual Document Retrieval Augmented Generation with a Single Image**, Ezzeldin Shereen, Dan Ristea, **Shae McFadden**, Burak Hasircioglu, Vasilios Mavroudis, Chris Hicks. *arXiv*

2024 **Impact of "TESSERACT: Eliminating Experimental Bias in Malware Classification across Space and Time"**, **Shae McFadden**, Zeliang Kan, Daniel Arp, Feargus Pendlebury, Roberto Jordaney, Johannes Kinder, Fabio Pierazzi, Lorenzo Cavallaro. [paper], [code]. *ACSAC Artifacts Competition*

2024 **SoK: On Closing the Applicability Gap in Automated Vulnerability Detection**, Ezzeldin Shereen, Dan Ristea, Sanyam Vyas, **Shae McFadden**, Madeleine Dwyer, Chris Hicks, Vasilios Mavroudis. [paper]. *arXiv*

2024 **TESSERACT: Eliminating Experimental Bias in Malware Classification across Space and Time (Extended Version)**, Zeliang Kan, **Shae McFadden**, Daniel Arp, Feargus Pendlebury, Roberto Jordaney, Johannes Kinder, Fabio Pierazzi, Lorenzo Cavallaro. [paper], [code]. *arXiv*

# Talks & Presentations

**Recent Advances in Resilient and Trustworthy Machine Learning Workshop (ARTMAN) co-located with ACSAC** — *Hawaii, USA*

Presenter for <The Impact of Active Learning on Availability Data Poisoning for Android Malware Classifiers>. *2024*

- Presented my paper published at the workshop on the use of active learning to recover malware classifiers from poisoning.

**Deep Learning Security & Privacy Workshop (DLSP) co-located with IEEE S&P** — *San Francisco, USA*

Presenter for <Wendigo: Deep Reinforcement Learning for Denial-of-Service Query Discovery in GraphQL>. *2024*

- Presented my paper published at the workshop on the use of DRL to discover low-rate DoS queries in GraphQL applications.

**Machine Learning and Cyber Security Symposium @ Imperial** — *London, UK*

Presenter for <Deep Reinforcement Learning for Denial-of-Service Query Discovery in GraphQL>. *2024*

- Presented my work on using deep reinforcement learning to discover low-rate denial-of-service queries in GraphQL applications.

**WorldCUR*BCUR** — *Warwick, UK*

Presenter for <Adversarial Machine Learning Evaluation of The MaMaDroid Feature Space>. *2023*

- Presented my work on the robustness of the MaMaDroid feature space and its interplay with concept drift mitigation techniques.

# Honours & Awards

## Competitions

2025 **LLMail-Inject: Adaptive Prompt Injection Challenge**, Mindrake Team Member, *IEEE SaTML*. *Copenhagen, DEN*

2024 **Cybersecurity Artifacts Competition**, Finalist, *ACSAC*. *Hawaii, USA*

2024 **AIxCC Semifinal Competition**, Mindrake Team Member, *DEFCON 32* *Las Vegas, USA*

## Awards

2024 **Top Reviewer**, awarded to recognize a reviewer's professionalism and diligence, AISec. *Salt Lake City, USA*

2023 **Layton Research Award**, given not merely for success in passing examinations but for the best promise of aptitude and genius for original scientific work, King's College London. *London, UK*

2023 **Alan Fairbourn Memorial Prize**, awarded to the student who produces the most meritorious final year project in the Department of Informatics, King's College London. *London, UK*

2023 **Associateship of King's College**, elected by the Academic Board of King's College London as an 'Associate of King's College' (AKC), King's College London. *London, UK*

2020 **Governor General Academic Medal**, awarded to the student graduating with the highest grade point average from a Canadian high school, University of Winnipeg Collegiate. *Winnipeg, CAN*

## Program Committees

| | | |
|---|---|---|
| 2025 | **18th ACM Workshop on Artificial Intelligence and Security (AISec)**, PC Member | *Taipei, Taiwan* |
| 2025 | **4th Workshop on Rethinking Malware Analysis (WoRMA)**, PC Member | *Venice, Italy* |
| 2025 | **8th Deep Learning Security and Privacy Workshop (DLSP)**, PC Member | *San Francisco, USA* |
| 2024 | **17th ACM Workshop on Artificial Intelligence and Security (AISec)**, PC Member | *Salt Lake City, USA* |
| 2024 | **21st Conference on Detection of Intrusions and Malware & Vulnerability Assessment (DIMVA)**, Sub-reviewer | *Lausanne, Switzerland* |
| 2024 | **33rd USENIX Security Symposium**, Sub-reviewer | *Philadelphia, USA* |

## Extracurricular Activity

### Men's First Fencing Team (King's College London)

*London, UK*

Captain *2022 - 2024*

- Led the team to winning the premier league and the national championships in 2023.

### Canadian National Fencing Team

*Canada*

Member *2017 - 2020*

- World Fencing Grand Prix in Montreal, Canada, 2020. U18 World Championships in Torun, Poland, 2019. U18 Fencing Pan American Championships in Bogata, Columbia, 2019. Commonwealth Fencing Championships in Canberra, Australia, 2018.