

# Shae McFadden

PH.D. CANDIDATE · ARTIFICIAL INTELLIGENCE & CYBERSECURITY

[✉ shae.mcfadden@kcl.ac.uk](mailto:shae.mcfadden@kcl.ac.uk) | [🏡 mcfadden-s.github.io](https://mcfadden-s.github.io) | [🔗 McFadden-S](https://McFadden-S) | [🔗 shae-mcfadden-27b840234](https://shae-mcfadden-27b840234.netlify.app) | [🎓 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

KING'S COLLEGE LONDON

London, UK

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

UNIVERSITY COLLEGE LONDON

London, UK

January 2024 - Present

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

### Visiting Collaborator @ AICD

THE ALAN TURING INSTITUTE

London, UK

October 2023 - Present

- A member of the DARPA AI Cyber Challenge (AlxCC) 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.

### Graduate Teaching Assistant

KING'S COLLEGE LONDON

London, UK

January 2024 - May 2024

- MSc Security Testing module labs.
- BSc/MSc Network Security module labs.

### King's Undergraduate Research Fellow

KING'S COLLEGE LONDON

London, UK

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)

KING'S COLLEGE LONDON

London, United Kingdom

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.

## Papers

---

### PUBLICATIONS

#### DRMD: Deep Reinforcement Learning for Malware Detection under Concept Drift,

2026 **Shae McFadden**, Myles Foley, Mario D'Onghia, Chris Hicks, Vasilios Mavroudis, Nicola Paoletti, Fabio Pierazzi. [paper], [code].

AAAI-26

#### The Impact of Active Learning on Availability Data Poisoning for Android Malware Classifiers,

2024 **Shae McFadden**, Zeliang Kan, Lorenzo Cavallaro, Fabio Pierazzi. [paper], [code].

ARTMAN co-located

with ACSAC

#### Wendigo: Deep Reinforcement Learning for Denial-of-Service Query Discovery in GraphQL,

2024 **Shae McFadden**, Marcello Maugeri, Chris Hicks, Vasilios Mavroudis, Fabio Pierazzi. [paper], [code].

DLSP co-located

with IEEE S&P

#### Poster: RPAL-Recovering Malware Classifiers from Data Poisoning using Active Learning,

2023 **Shae McFadden**, Zeliang Kan, Lorenzo Cavallaro, Fabio Pierazzi. [paper], [code].

ACM CCS Poster

### PREPRINTS

#### One Pic is All it Takes: Poisoning Visual Document Retrieval Augmented Generation with a Single

2025 **Image**, Ezzeldin Shereen, Dan Ristea, **Shae McFadden**, Burak Hasircioglu, Vasilios Mavroudis, Chris Hicks.

arXiv

2024	<b>SoK: On Closing the Applicability Gap in Automated Vulnerability Detection</b> , Ezzeldin Shereen, Dan Ristea, Sanyam Vyas, <b>Shae McFadden</b> , Madeleine Dwyer, Chris Hicks, Vasilios Mavroudis. [paper].	arXiv
2024	<b>TESSERACT: Eliminating Experimental Bias in Malware Classification across Space and Time (Extended Version)</b> , Zeliang Kan, <b>Shae McFadden</b> , Daniel Arp, Feargus Pendlebury, Roberto Jordaney, Johannes Kinder, Fabio Pierazzi, Lorenzo Cavallaro. [paper], [code].	arXiv

## NON-ARCHIVALS

2025	<b>One Pic is All it Takes: Poisoning Visual Document Retrieval Augmented Generation with a Single Image (Short Paper)</b> , Ezzeldin Shereen, Dan Ristea, Burak Hasircioglu, <b>Shae McFadden</b> , Vasilios Mavroudis, Chris Hicks.	MAGMaR co-located with ACL
2024	<b>Impact of "TESSERACT: Eliminating Experimental Bias in Malware Classification across Space and Time"</b> , <b>Shae McFadden</b> , Zeliang Kan, Daniel Arp, Feargus Pendlebury, Roberto Jordaney, Johannes Kinder, Fabio Pierazzi, Lorenzo Cavallaro. [paper], [code].	ACSAC Artifacts Competition

## 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.	
<b>Deep Learning Security &amp; Privacy Workshop (DLSP) co-located with IEEE S&amp;P</b>	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.	
<b>Machine Learning and Cyber Security Symposium @ Imperial</b>	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.	
<b>WorldCUR*BCUR</b>	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	<b>European Defence Tech Hackathon</b> , MARSim Team Lead.	London, UK
2025	<b>LLMail-Inject: Adaptive Prompt Injection Challenge</b> , Mindrake Team Member, <i>IEEE SaTML</i> .	Copenhagen, DEN
2024	<b>Cybersecurity Artifacts Competition</b> , Finalist, ACSAC.	Hawaii, USA
2024	<b>AIxCC Semifinal Competition</b> , Mindrake Team Member, DEFCON 32	Las Vegas, USA

### AWARDS

2025	<b>Top Reviewer</b> , awarded to recognize a reviewer's professionalism and diligence, <i>ACM AISeC</i> .	Taipei, Taiwan
2024	<b>Top Reviewer</b> , awarded to recognize a reviewer's professionalism and diligence, <i>ACM AISeC</i> .	Salt Lake City, USA
2023	<b>Layton Research Award</b> , given not merely for success in passing examinations but for the best promise of aptitude and genius for original scientific work, <i>King's College London</i> .	London, UK
2023	<b>Alan Fairbourn Memorial Prize</b> , awarded to the student who produces the most meritorious final year project in the Department of Informatics, <i>King's College London</i> .	London, UK
2023	<b>Associateship of King's College</b> , elected by the Academic Board of King's College London as an 'Associate of King's College' (AKC), <i>King's College London</i> .	London, UK
2020	<b>Governor General Academic Medal</b> , awarded to the student graduating with the highest grade point average from a Canadian high school, <i>University of Winnipeg Collegiate</i> .	Winnipeg, CAN

## Technical Skills

### Python Programming

PROJECTS USING PYTHON:

- Published several papers which used Python to perform machine learning evaluations.
- Current maintainer of four open-source Python Github repositories ( DRMD [link], 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.

## Academic Service

---

2026	<b>33rd ACM Conference on Computing and Communication Security (CCS)</b> , PC Member	Netherlands
2026	<b>40th AAAI Conference on Artificial Intelligence</b> , PC Member	Singapore
2025	<b>4th Machine Learning and Cyber Security Symposium (ML-CSS)</b> , Organizer	London, UK
2025	<b>18th ACM Workshop on Artificial Intelligence and Security (AISeC)</b> , PC Member	Taipei, Taiwan
2025	<b>4th Workshop on Rethinking Malware Analysis (WoRMA)</b> , PC Member	Venice, Italy
2025	<b>8th Deep Learning Security and Privacy Workshop (DLSP)</b> , PC Member	San Francisco, USA
2024	<b>17th ACM Workshop on Artificial Intelligence and Security (AISeC)</b> , PC Member	Salt Lake City, USA
2024	<b>21st Conference on Detection of Intrusions and Malware &amp; Vulnerability Assessment (DIMVA)</b> ,	Lausanne, Switzerland
2024	Sub-reviewer	
2024	<b>33rd USENIX Security Symposium</b> , Sub-reviewer	Philadelphia, USA

## Extracurricular Activity

---

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

London, UK

#### CAPTAIN

2022 - 2024

#### 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.