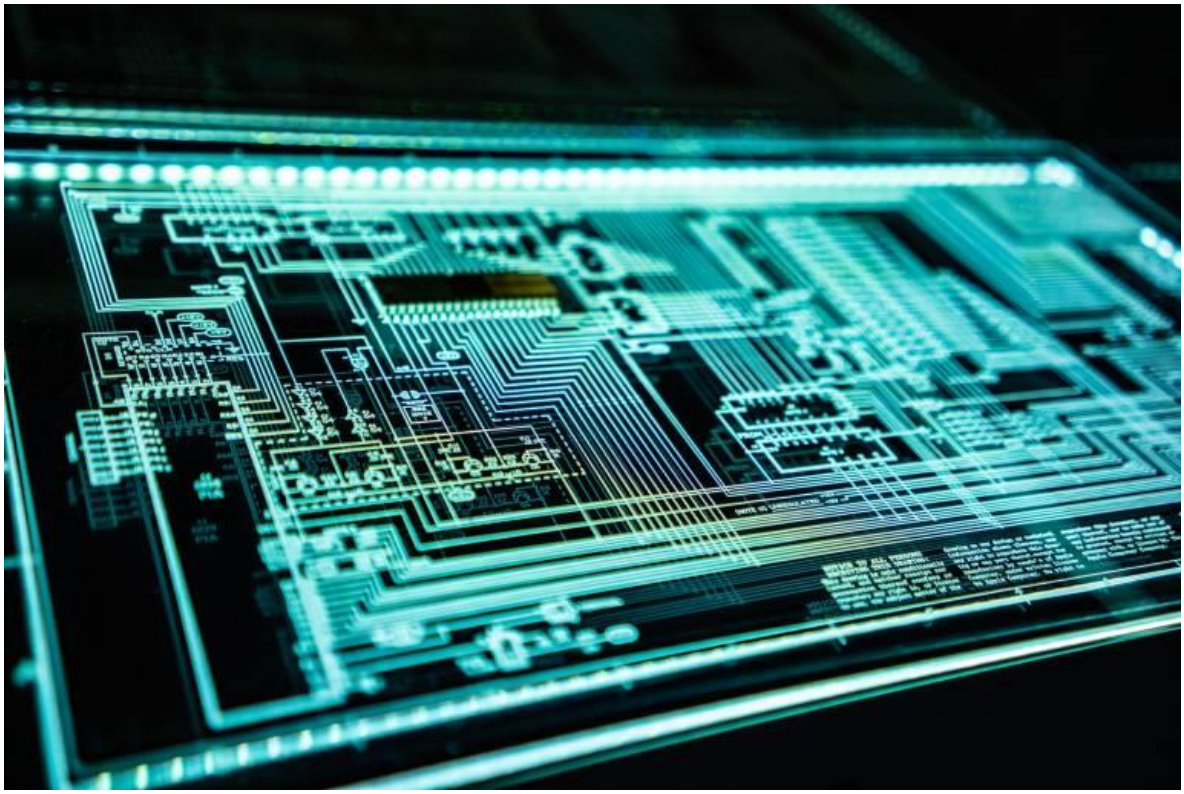
EPSRC CDT in Cyber-Physical Risk

2025 Project summaries



University College London

Version 2

Last updated: 07.3.2025

Introduction

This document provides an overview of the research projects available at the EPSRC Centre for Doctoral Training (CDT) in Cyber-Physical Risk. These projects have been proposed by leading academics from UCL Security and Crime Science, UCL Computer Science, and other partner departments. They offer PhD students the opportunity to work at the forefront of interdisciplinary research, collaborating with academic experts and, in some cases, external partners from industry or government.

The project outlines are organised into four research themes that address critical societal challenges:

- **Futures** – Examining how emerging socio-technical trends shape cyber-physical risks and their geopolitical implications, with research on risk foresight, scenario planning, regulatory challenges, and societal resilience.
- **Cyber-Physical Systems** – Investigating security challenges in cyber-physical systems (CPS) across industries such as healthcare, smart infrastructure, and autonomous transportation, focusing on adversarial machine learning, cyber-situational awareness, and forensic investigations.
- **Online Communication** – Exploring threats from digital platforms, including disinformation, hate speech, and criminal activities, and developing AI-driven content detection and regulatory approaches.
- **Simulation and Interaction** – Using augmented and virtual reality (AR/VR) to study human behaviour in cyber-physical risk scenarios, test security interventions, and refine emergency responses through immersive simulations.

Further details about each research theme, including the lead teams and their areas of expertise, visit the website. You can also view the full list of supervisors and their research profiles.

Table of Contents

Futures

There is currently no project available with Futures as a primary theme.

However, many projects include ‘futures’ as a secondary theme:

- Detection and mitigation of ransomware attacks on Industrial Control Systems
- Enhancing Crime Detection and Investigation in the Internet of Things Era
- Risk Assessment and Mitigation of Threats to AI-enabled Devices in Cyber-Physical-Social Systems
- Social media, Crime and the Environment

Cyber-Physical Systems

Attack and Defence of Cyber-physical Systems relying on Multimodal Foundational Models

The supervisory team includes Prof. Mirco Musolesi (UCL Computer Science) and Prof. Stephen Hailes (UCL Computer Science).

What the research is about

Multimodal foundational models, which integrate multiple data modalities such as text, images, audio, and video, have revolutionized various applications, including image captioning, visual question answering, and robotic systems. However, these models are vulnerable to sophisticated attacks that can originate in the virtual world and extend to real-world systems, posing significant security risks.

The project will explore the foundations of the design and implementation of Agentic AI systems based on Multi-modal Foundational Models analysing their vulnerabilities and potential strategies for protection. In particular, the student will: 1) start with the identification and categorization of attacks on multimodal foundational models; 2) develop and evaluate defence mechanisms to mitigate these attacks; 3) analyse the impact of virtual attacks on real-world systems, namely robotic systems and, more in general, systems with physical actuators; 4) develop open source solutions for the community and propose guidelines and best practices for securing multimodal foundational models.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. They will have a background in Computer Science, Engineering, Mathematics or related areas or with a strong quantitative background. A strong interest in machine learning/artificial intelligence (in particular in foundational models and generative AI) is essential given the topic of the project.

Computational Threat Assessments: The relationship Between Online Threats and Real-world Action

The supervisory team includes Prof Paul Gill (UCL Security and Crime Science), specialist in threat assessment, and behavioural analysis.

Theseus Risk Management Ltd is an onboarded partner of the CDT and will provide data related to written communicated threats to industry organisations and their personnel. The data will be supplemented by similar threats (1000+ per year) made to the Royal Family, and Members of Parliament.

What this research is about

Governments and law enforcement agencies rely on effective threat assessment tools to address terrorism, mass shootings, and other forms of violence. Increasingly these tools are asked to assess the likelihood of threats originating in digital spaces translating into real-world violence. However, the science has not kept pace with the rate of change evident in practitioner caseloads. This project has the potential to help authorities pre-emptively identify, triage, prevent and disrupt risk via the testing, and validation of various predictive and computational models. This could advance the state-of-the-art in AI and natural language processing (NLP), especially in sentiment analysis, anomaly detection, and contextual understanding. A necessary aspect of this thesis also involves the exploration of how computational threat assessment tools can be developed and used ethically, avoiding misuse or discrimination.

The chosen student will review studies on psychological and crime science underpinnings of online threats. They will investigate existing computational tools and

algorithms for sentiment analysis, NLP, and threat detection including the use of psycholinguistic dictionaries (e.g. the Grievance dictionary).

Projects could involve temporal examinations of rich case studies where online threats have escalated into real-world incidents. The project will involve the collection, and cleaning of data from multiple stakeholder partners, and the pre-processing of textual data for computational analysis. Empirical analyses could include any mixture of the following: (1) Using machine learning techniques to develop predictive models for identifying credible threats (2) Applying NLP techniques (e.g., sentiment analysis, topic modeling) to assess the content of online posts (3) Integrating behavioural patterns, historical data, and context into the model for better accuracy (4) Testing the model's ability to correlate online activity with real-world actions.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. They should be willing to collaborate across sectors, with strong analytical skills in computational modeling, machine learning, and NLP.

Cyber-Physical Security in Vision-Language-Action Models for Autonomous Systems

The supervisory team includes Assoc. Prof. Chris Xiaoxuan Lu (UCL Computer Science) and a secondary supervisor to be confirmed. The student is also expected to work with self-driving vehicle stakeholders.

What this research is about

The integration of Vision-Language-Action Models (VLAMs) into autonomous systems, such as self-driving vehicles and robotic manipulators, revolutionizes multimodal decision-making but also introduces hybrid cyber-physical vulnerabilities. These systems are increasingly susceptible to adversarial attacks that exploit their reliance on visual and linguistic inputs, posing significant risks to safety-critical applications. This research addresses these challenges by investigating threats at the intersection of cyber and physical domains, aligning directly with the CDT's goal of managing risks that propagate across domains.

By developing comprehensive threat models and defensive strategies, this work contributes to the resilience of critical infrastructures, such as autonomous transportation networks and automated industrial systems. Additionally, it emphasizes the importance of pre-empting unintended consequences, such as cascading failures across cyber-physical interfaces, ensuring safe and ethical integration of AI technologies. The findings will empower stakeholders to anticipate, mitigate, and regulate cyber-physical risks, supporting societal resilience and strengthening defences against hybrid threats.

The student will explore the vulnerabilities of Vision-Language-Action Models (VLAMs) in autonomous systems, focusing on hybrid cyber-physical risks. Their research will involve designing and evaluating adversarial attacks across visual (e.g., adversarial patches, Out-of-Distribution perturbations) and linguistic (e.g., crafted text prompts for jailbreaks) modalities. The student will also investigate hybrid attacks that simultaneously exploit both modalities, leveraging simulators like CARLA and robotic manipulation environments such as VIMA or SimplerEnv.

In parallel, the student will develop and test robust defence mechanisms. This includes adversarial training, multimodal anomaly detection systems, and cross-modality redundancy checks to mitigate attack impacts. These solutions will be validated through rigorous testing in simulated and physical environments.

The project also involves interdisciplinary considerations, such as evaluating how cyber-physical vulnerabilities propagate across domains and addressing ethical concerns in adversarial defences. Through these activities, the student will contribute actionable insights to enhance the security and resilience of critical cyber-physical systems in safety-critical applications.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. We are seeking a motivated and interdisciplinary student with a strong academic background in computer science, robotics, or a related field. The ideal candidate will have experience in machine learning, computer vision, or natural language processing, and a keen interest in cyber-physical systems and

security. Familiarity with ROS, adversarial machine learning, reinforcement learning, or robotics simulation tools (e.g., CARLA, SimplerEnv) is highly desirable.

The student should be proactive, with excellent problem-solving skills and a collaborative mindset. A commitment to addressing hybrid cyber-physical risks and an understanding of ethical implications in AI and cybersecurity will be key to success in this project.

Enhancing Crime Detection and Investigation in the Internet of Things Era

The supervisory team includes Dr. Anna Maria Mandalari (Electronic and Electrical Engineering), Dr. Nilufer Tuptuk (Security and Crime Science), and Dr. Fabio Pierazzi (Computer Science). Dr. Mandalari directs one of the world's most advanced IoT testbeds, providing access to devices for hands-on evaluations within an interdisciplinary research environment. The team has strong UK and international law enforcement connections that could support this work.

What the research is about

This project focuses on developing advanced tools and methodologies to detect and mitigate emerging Internet of Things (IoT)-enabled crimes. The project addresses two main crime categories: IoT-dependent crimes, where IoT devices are directly targeted (e.g., ransomware, botnets, system exfiltration and sabotage), and IoT-facilitated crimes, where IoT device vulnerabilities are exploited to conduct traditional offenses (e.g., cyberstalking, identity theft, technology-facilitated abuse).

Students will gain interdisciplinary expertise in IoT security, digital forensics, and crime prevention strategies. They will contribute to the development of a dynamic knowledge base of IoT vulnerabilities, advanced forensic tools, and standardised testbeds for cyber-physical vulnerability testing and crime simulation. Additionally, they will explore the legal and ethical dimensions of digital evidence handling. This project will equip students with the skills to address complex security challenges and leverage IoT devices as both evidence and intelligence sources, preparing them for careers at the intersection of technology and security.

The student will conduct in-depth research on IoT vulnerabilities and their exploitation in emerging cyber-physical crimes. They will design and implement advanced security and forensic tools for vulnerability detection, evidence gathering, and crime prevention. Using UCL's state-of-the-art IoT testbed, the student will perform vulnerability testing and simulate cybercrime scenarios to evaluate security solutions. They will also analyse real-world data to identify threat patterns, collaborate with law enforcement agencies to align research with investigative needs, and contribute to developing a dynamic knowledge base on IoT security risks. The student will engage with legal and ethical considerations surrounding digital evidence handling, ensuring compliance with privacy and data protection standards. This project offers hands-on experience in cybersecurity research, IoT systems, and forensic science, preparing the student for impactful roles in cybersecurity and crime prevention.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. We are seeking a highly motivated student with a strong academic background in computer science, cybersecurity, electronic engineering, or a related field. Ideal candidates should have experience in IoT systems, network security, and/or digital forensics. Proficiency in programming (e.g., Python, C/C++), machine learning, and familiarity with network penetration and testing, and other cybersecurity tools are desirable. The student should possess strong analytical and problem-solving skills, be capable of independent and collaborative research, and have an interest in interdisciplinary work that combines technology, security, and crime science.

Investigating Digital Supply Chain Attacks in Digital Twins and Developing Solutions

The supervisory team includes Prof. James Hetherington (expert in large-scale modelling and simulation), Dr Nilufer Tuptuk (Director of the Operational Technology Lab, specialising in cyber-physical systems security and digital twins), and Prof. Stephen Hailes (expert in networking, system security, and AI). The team has good connections with relevant industries.

What this research is about

Digital twins are increasingly integrated into critical infrastructure sectors, including utilities, energy, manufacturing, and transportation. Built on advanced technologies like industrial control systems, sensor networks, cloud computing, and AI-driven analytics, they remain continuously connected to their physical counterparts. This constant connection makes digital twins highly susceptible to cyberattacks, which can lead to physical consequences.

The complexity of the supply chain, involving diverse technologies such as operational technology, information and communication technology, and AI-based analytics, significantly increases the attack surface. Cyberattacks targeting digital twins in critical systems can have devastating effects, including economic losses, operational disruptions, safety hazards like pipeline explosions or factory shutdowns, and the failure of essential services such as electricity and water.

This PhD topic is both timely and underexplored, addressing a critical research gap. It will enable the secure and safer adoption of digital twins in critical sectors, ensuring their reliability and resilience against evolving cyber threats.

The student will undertake practical research to investigate supply chain attacks on digital twins. The research will involve analysing supply chain vulnerabilities and threats, simulating and modelling realistic systems, and developing functional digital twin to serve as a testbed. The student will create detailed threat models and simulate cyberattacks targeting the digital twin, enabling the assessment of their impacts on critical systems. The goal is to propose effective, robust solutions to mitigate these threats and improve the security and resilience of digital twins used in critical sectors.

To support this research, the student will have access to the Operational Technology Lab, equipped with technologies widely used in the energy, manufacturing and water sectors. This hands-on access to industry-relevant equipment will provide a realistic environment for testing and validating the developed solutions, ensuring their practicality and relevance to real-world applications.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. Applicant must possess outstanding academic achievements and a background in Computer Science, Mechanical Engineering, Manufacturing Engineering, Robotics, Mathematics, or a related field. The applicant should have a strong interest in cyber-physical systems security, as well as expertise in simulation and modelling tools and AI (machine learning and deep learning). They should have strong skills in (e.g., Python, MATLAB/Simulink, C/C++). Previous experience in digital simulation and industrial computing, such as PLC instrumentation, would be an added advantage.

Protecting Industrial Control Networks by Disrupting Reconnaisance Through Traffic-Analysis Resistance Techniques

The primary supervisor will be Dr Steven Murdoch (UCL Computer Science), specialised in secure network design and traffic-analysis resistance techniques. The secondary supervisor is to be confirmed. Experts in the development of high-assurance traffic analysis techniques, including the developers of Arti, a memory-safe low-latency anonymous communication system as well as organisations using and developing industrial control networks.

What this research is about

Industrial control networks must be well protected, since adversaries with access to them can cause harm to equipment connected to the system as well as to the wider public. When critical national infrastructure is compromised the damage resulting from a successful attack could be substantial. However, industrial control networks are challenging to protect because they rely on legacy technologies and concerns about safety create obstacles to upgrading systems to adopt modern security approaches. Attacks can only be effective if they are well-targeted and so their initial stage is reconnaissance (as discussed both in MITRE ATT&CK and Lockheed Martin cyber kill chain). In this stage the attacker will observe networks to identify the targets necessary to achieve objectives. Even if network data is encrypted traffic-analysis is effective at identifying targets. Therefore, in this project we will apply traffic-analysis resistance techniques to disrupt the reconnaissance stage, preventing attacks before they take place.

The student will develop techniques for creating secure overlay networks designed for industrial control systems that use encryption and traffic-analysis resistance techniques to disrupt attacks at the reconnaissance stage. This will include adapting existing technologies such as Arti to support the requirements of industrial control systems including guarantees on latency, fault-tolerance and compatibility with network protocols using in industrial control networks. Furthermore, the project will develop techniques for network visibility to allow the operator of the network gain assurance of correct operation and detect attacks while simultaneously preventing attacks from being able to initiate attacks. To give assurance of secure software development, techniques from the EPSRC Digital Security by Design project will be used including the CHERI-IoT platform.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. The project needs a student with strong computer science skills, including programming, and at least an interest in cybersecurity. A good background in mathematics will be important for analysing the security of the prototypes developed. The student should be willing to learn about cyber-physical systems and particularly develop an excellent understanding of the risks faced by industrial control systems and the challenges of securing them.

RF Awareness and Fingerprinting + Detection (RF-FD) for countering cyber-physical attack applications

The supervisory team includes Dr. Matthew Ritchie (EEE) and Prof. Kevin Chetty (UCL Security and Crime Science). This project will look to work with both government and industry partner to better understand the threats + technical challenges and move towards a co-designed innovative solution to this currently and future real world problem.

What the research is about

This research is about detecting and classifying the presence of Radio Frequency (RF) signals that are linked to criminal/terrorist activities. Adversaries may use a variety of electronic devices to perpetrate crimes but mask their identities by avoiding affiliations with them (e.g. burner phones). Traditional cyber techniques may look to use the network layer information but counters to this exist. The RF signature of a given device is not something that can be masked or spoofed making the potential applied research outcomes highly desirable for end users.

The research will also tackle challenges in securing critical national infrastructure (CNI). Portable IoT based devices can be placed within CNI in order to disrupt or deny its operation. RF fingerprinting will be able to detect and uniquely identify these devices via physical security deployments, neutralising their capabilities and providing forensic information that can be used to prosecute those that were involved.

The student will work on tasks that range from modelling RF signals, planning experimentations, running captures in controlled environments to outdoor trials with equipment (that is available in the EEE dept. and can be exploited in this PhD). An adaptive RF prototype solution (ARESTOR) developed in EEE can be leveraged within this work that will enable rapid and impactful research outcomes from the work. Once real datasets are captured the student can then process the results, apply machine learning algorithms (bespoke designed for RF signals) and demonstrate how RF fingerprinting can be applied in a cyber-physical security environment.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. The successful applicant will demonstrate good signal processing experience, knowledge of applying practical machine learning algorithms on signals and an interest in performing real world experiments. Any prior experience processing Radio Frequency signals is helpful. They will have the ability to take on complex problems, develop a research plan and delivery outcomes in the form of publications, presentations and outputs to the project partners. It is likely that this student will need to be a UK national willing to undertake background security checks.

Risk Assessment and Mitigation of Threats to AI-enabled Devices in Cyber-Physical-Social Systems

The supervisory team includes Dr. Fabio Pierazzi (UCL Computer Science) and Dr. Anna Maria Mandalari (UCL Electronic and Electrical Engineering). Dr. Mandalari is also Director of the SafeNetIoT lab in EEE, which will give access to real AI-enabled devices for evaluations.

What the research is about

Artificial Intelligence (AI) and Machine Learning (ML) have been widely adopted both for automated decision-making (e.g., smart home automation) and for threat detection. However, this creates new opportunities for malicious actors to conduct “adversarial attacks” to compromise the security and privacy of AI-enabled device users. While the security and robustness of AI have been studied in many digital systems security scenarios (e.g., malware detection, network intrusion detection), they have been less explored in hybrid cyber and physical systems, where attacks may have different risks and impacts. For example, systems including home automation and industrial IoT devices that interact with humans. This project will design new risk assessment methodologies of AI-enabled devices in Cyber-Physical-Social Systems (CPSS), which is crucial to understand how AI security intertwines with the physical world, what hybrid mitigations need to be put in place, and how to produce evidence to inform AI policymakers and regulators.

This project involves the analysis and impact of adversarial attacks on AI-enabled devices in cyber-physical-social scenarios. For example, smart home devices running AI algorithms for task automation may be compromised or malfunction as a result of

adversarial ML attacks. The student will design novel algorithms and risk assessment methodologies to (semi-)automatically identify threats in AI-enabled devices that may also impact the physical world, towards designing mitigations based on both AI hardening methodologies (e.g., adversarial training, embedding expert knowledge with model-driven AI), and potentially suggesting physical mitigations (e.g., adding more sensors). Since physical mitigations are more costly, the candidate will propose techniques for estimating the likelihood of attacks. The candidate will identify and focus on a primary application scenario, based on their prior expertise and in discussion with their supervisors. The evaluations will be both with real AI-driven devices and on simulated physical systems with human interaction.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. The successful applicant will have technical expertise in machine learning, cybersecurity, and cyber-physical systems (e.g., IoT, simulation environments, RL). The candidate needs to be willing to create rigorous methodologies for automated risk assessment and mitigation of threats in the real world and be open to working both with real devices and in simulated environments. The candidate may also need to be open to conducting user studies and surveys for the effectiveness and relevance of the proposed risk assessments. Coding experience and expertise is essential (ideally using Python).

Securing Cyber-physical Systems Against Cyber-attacks: A Hybrid Network Modelling Approach

The supervisory team includes Prof. Mirco Musolesi (UCL Computer Science) and Prof. Stephen Hailes (UCL Computer Science).

What this research is about

Cyber-physical systems integrate computational elements with physical processes, creating a seamless interaction between the digital and physical worlds. These systems are prevalent in critical infrastructure, industrial plants, and robotic systems, making them essential to modern society. However, their connectivity to communication networks and the internet exposes them to significant vulnerabilities and potential cyber-attacks. Our society relies on such systems for their functioning, and, given the current geopolitical landscape, these cyber-physical risks are one of our primary security concerns.

The project will first focus on the study of vulnerability of cyber-physical systems, considering aspects concerning robustness and resilience from a modelling and simulation point of view. The idea is to study this problem through mathematical and computational models of this class of hybrid systems, which comprise physical networks (composed, for example, by sensors, actuators, physical processes, etc.) and digital networks (composed, for example, by computational elements, communication infrastructure, etc.). The student will then apply these theoretical findings to the design of practical proof-of-concept implementations.

In terms of theoretical analysis, machine learning (e.g., graph neural networks), complex network theory, and game theory provide a set of powerful tools for analysing and predicting interactions in hybrid systems. The idea is to model interactions between attackers and defenders as a strategic game, where each player aims to maximise their payoff. The student will also focus on problem of resource allocation, analysing how resources can be optimally allocated to enhance system security and performance.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. Students with a background in Computer Science, Engineering, Mathematics and related areas or with a strong quantitative background. Strong research interest in (network) modelling, machine learning/artificial intelligence, and game theory is essential given the topic and methodologies of the project.

Supporting preparedness and response to cyber-attacks in hospitals

The supervisory team will include Dr Luca Grieco and Professor Christina Pagel (Clinical Operational Research Unit, University College London) and Professor Hervé Borrion (Security and Crime Science, UCL). Dr Saira Ghafur (Institute of Global Health Innovation, Imperial College London) will be a project advisor. You will conduct this project in tight collaboration with the Emergency preparedness, resilience and response (EPRR) unit at UCLH.).

What the research is about

Recently, the number and severity of cyber-attacks against healthcare organisations has increased significantly. According to ITPro (2024), healthcare was among the top three most targeted sectors in 2023, with 1,500 weekly attacks on average. In May 2021, the Irish health system experienced a serious ransomware attack, where access to electronic systems and data was blocked, severely impacting critical services such as gynaecology, maternity, cancer care and children's care. In June 2024, the ransomware cyber-attack against pathology services provider Synnovis targeted hospitals in London, resulting in the postponement of over 9,000 acute outpatient appointments and over 1,500 elective procedures. Following system downs, hospitals face situations where key clinical information stored in digital systems is momentarily lost, while patients still need to receive possibly urgent care. Identifying potential disruptions caused by cyber-incidents of different types/sizes and establishing procedures to minimise such disruption are crucial for the delivery of healthcare at highest standard.

The project will consist of: i) identifying disruptions caused by cyber-incidents in hospitals, establishing links between different types/sizes of incidents and the type and amount of disruption caused; ii) exploring network effects if several hospitals are affected; iii) exploring preparedness and response procedures and quantifying their potential for mitigation of the above disruptions. The student will work in close collaboration with the partner organisation to gain a full understanding of hospital operations regarding the digital systems in use and to identify implementable mitigating procedures at the hospital. The latter might involve, for instance, a combination of manual record procedures, paper back-up strategies, patient prioritisation rules, etc. The student will then develop software implementing simulation-optimisation algorithms to test the above combinations of procedures and inform the partner organisation about their potential effectiveness. Envisaged algorithms would consist of Discrete Event Simulation or Agent-Based Modelling approaches incorporating Stochastic Optimisation and/or Game Theory procedures.

Is this PhD project for you?

The successful applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. We are seeking a student who has a strong interest in the application of operational research and data science techniques to healthcare settings, as well as a willingness to explore the development of simulation-optimisation algorithms tailored for the problem at hand. Prior experience of coding (e.g. Python) is desirable.

Transfer Learning for Threat Detection and Mitigation in Cybersecurity

The supervisory team includes Prof Benjamin Guedj (Computer Science). His team specialises in theoretical machine learning, generalisation theory, and transfer learning, among other topics. They will also leverage their network of collaborators in the UK, France, and many other countries.

What the research is about

The increasing sophistication of cyberattacks, such as zero-day exploits and advanced persistent threats (APTs), poses significant challenges to traditional security measures. This PhD project will focus on leveraging advanced machine learning algorithms, and more specifically transfer learning, to design, implement, and evaluate intelligent systems for real-time threat detection, prediction, and mitigation in cybersecurity. Potential applications include intrusion detection systems (IDS), malware analysis, and proactive response frameworks for protecting critical infrastructure. The work will bridge machine learning research with practical cybersecurity challenges, enhancing the resilience of digital systems against emerging threats.

Transfer learning will be central to this project by enabling models to adapt efficiently to cybersecurity challenges with limited labeled data or evolving threats. Pre-trained models, which capture broad patterns from large datasets, will be fine-tuned on domain-specific cybersecurity data, such as network logs or malware signatures. This approach significantly reduces training time and computational costs while improving performance. Additionally, transfer learning will enhance the detection of novel or rare threats, such as zero-day attacks, by requiring fewer examples to generalize effectively.

It will also facilitate knowledge transfer across different threat scenarios, allowing insights from one type of attack to inform the detection of others. By leveraging transfer learning, the project will develop scalable, adaptable AI solutions for real-time threat detection and mitigation in cybersecurity.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. We are looking for an exceptionally bright student with a mathematics and/or computer science background, with excellent maths (statistics, machine learning, probability theory) and coding (ideally in Python) skills, and a genuine interest in machine learning and cybersecurity.

Online Communications

Social Media, Crime and the Environment

The supervisory team includes Prof. Hervé Borrion (CDT Director) and Prof. Ben Bradford (UCL Security and Crime Science).

What this research is about

Environmental policies and legislation are increasingly emerging as significant sources of political division. Activists on both sides of the debate have engaged in anti-social behaviour and criminal acts, including the destruction of ULEZ cameras, vandalism of museum paintings, and clashes between drivers and cyclists. This project seeks to explore whether and how digital communication tools, such as social media, facilitate, enable, and legitimize environmentally related crime and anti-social behaviour. The insights gained will help guide the development of tools and strategies to anticipate and mitigate such risks.

After identifying specific crime issues that are directly or indirectly linked to environmental policies and legislation, you will work to gain a deeper understanding of these issues, including their nature and trends. Subsequently, you will review relevant social media posts and threads to propose targeted research questions about how social media contributes to facilitating, enabling, and legitimising environmentally related anti-social behaviour and crime. To address these questions, you will conduct three or four empirical studies during the period of your PhD. The analysis will likely involve a mixed-methods approach.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. We are seeking a highly capable student with a strong academic background and a desire to specialize in analysing crime risks that span both digital and physical domains. The ideal candidate will have a keen interest in integrating and applying theories and methods from diverse disciplines, including computer science, environmental criminology, psychology, and media studies. These methods may encompass crime scripting, narrative analysis, video analysis, surveys, interviews, and computational techniques for data extraction and analysis (e.g., natural language processing). Proficiency in statistical analysis and programming (C/C++, R, or Python) is essential. While prior experience with machine learning and API usage is preferred, it is not mandatory.

Talk is Cheap? Assessing How Extremist Content Online Can Promote Violence Offline to Identify Countermeasures

The first supervisor is Dr Sandy Schumann (UCL, Security and Crime Science, Lecturer). The second supervisor is Dr Tristan Caulfield (UCL, Computer Science, Associate Professor).

What this research is about

Extremist (i.e., anti-democratic) content flourishes online. Policymakers and practitioners have invested in various measures (e.g., detection/removal; counter-narratives) based on the premise that such content has detrimental consequences, notably, facilitating one-off incidents or campaigns of collective violence against minority groups and government representatives offline. However, comprehensive empirical (causal) evidence supporting these 'exposure effects' is scarce. One reason for this gap in the literature is a lack of ecologically valid analytical frameworks that draw on mixed methods. Additionally, theoretical insights from disparate disciplines (e.g., psychology, information science) are not applied to their full potential as they have not been integrated. As a result, resource-intensive countermeasures online have not been evaluated systematically, that is, capturing, beyond the impact on attitudes and activities online, reduced risk of violence offline. Crucially, although the attack scenarios could span digital and physical settings, inter-relations between online and offline (e.g., SCP) countermeasures remain unexplored.

The student will, first, systematise interdisciplinary evidence that documents *whether* and *how* (e.g., desensitisation) exposure to extremist content online promotes violence targeting government and minorities offline. Next, the student will develop a framework

to analyse dynamics that are expected to be observable in text, combining longitudinal (secondary) data from relevant social media platforms/services (e.g., EDL Telegram channels prior to the 2024 riots; Stormfront forum) and open-source information about respective incidents/campaigns of violence offline. The student will explore how tools like GPT can be exploited for computational linguistic analysis that forms the core method of this study. Ecologically valid human-participant experiments will consolidate key findings and investigate dynamics not easily observable in text (e.g., cognitions).

All insights will inform the design of an agent-based model employed to examine how distinct interventions online (e.g., counter-narratives) and offline (e.g., SCP) affect (together) the probability of one-off incidents and campaigns of collective violence offline.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk and contributing to the development of the field within a cohort-based environment. The ideal student will show affinity for interdisciplinary scholarship, especially such that combines insights from social science domains and computer science. Advanced knowledge of R, analytical skills enabling them to (learn how to) conduct computational linguistic analyses and agent-based modelling, and a readiness to implement reproducible/open science practices are required.

Simulation and Interaction

There is currently no project available with Simulation and Interaction as a primary theme.

However, many projects involve performing computational simulation.