

EPSRC CDT in Cyber-Physical Risk

2026 Project summaries



University College London

Version 1

Last updated: 15.12.2025

Introduction	5
The student journey	7
Phase 0: Apply and secure funding	7
Phase 1: Preparing for research (Months 1-6)	9
Phase 2: Advancing knowledge and research (Months 6-24)	13
Phase 3: Becoming an expert (Months 24-43)	13
Phase 4: Unleashing potential (Months 36-48)	13
General entry requirements.....	16
Attendance requirements	16
Part-time mode	16
Research Topics	18
Futures	20
Hybrid Risks in Agentic AI Ecosystems	21
Cyber-Physical Systems	22
Attack and Defence of Cyber-physical Systems relying on Multimodal Foundational Models	23
Adversarial Data Attacks on Machine Learning and Generative AI-based Process Monitoring	25
Hearts on the Cloud: Threat modelling the cybersecurity risks of internet-connected cardiac implants	27
HART: Human-Aware Resilient Teleoperation under Adversarial Threats	29
Computational Threat Assessments: The relationship Between Online Threats and Real- world Action	30

Cyber-Physical Security in Vision-Language-Action Models for Autonomous Systems.....	32
Identifying and Detecting Unauthenticated UAVs to Against Physical and Cyber Sabotage Threats	34
Open-Ended World Modeling for Emergent Multi-Agent Behavior	35
Investigating Digital Supply Chain Attacks in Digital Twins and Developing Solutions	37
Protecting Industrial Control Networks by Disrupting Reconnaissance Through Traffic- Analysis Resistance Techniques.....	39
Risk Assessment and Mitigation of Threats to AI-enabled Devices in Cyber-Physical-Social Systems.....	42
Securing Cyber-physical Systems Against Cyber-attacks: A Hybrid Network Modelling Approach	44
Supporting preparedness and response to cyber-attacks in hospitals	46
Systemic Risk in the Internet of Things	48
Adversarial Attacks on AI-based Control Architectures in Critical Infrastructure Systems	50
Autonomous Cyber-Physical Hazard Mapping	52
SafeGen - Safe Generative AI Models	53
Securing Future Satellite Networks.....	55
Understanding the role of ‘grey zones’ in hybrid attacks against critical national infrastructure in the UK.....	58
When, Not If: Designing for Recovery from Cyber-Physical Breaches.....	60
A Socio-Technical Approach for Responding to Attacks on Critical Infrastructure	62
Adversarial AI for Red-Teaming Cyber-Physical Systems	64
Physics-Inspired ML for Safe Human-Robot Interaction using VLA Models.....	66
Developing Secure and Resilient Smart Building Systems	67

Cyber-Physical Abuse Pathways in Assistive and Home Health Technologies for Patients with Disabilities and Chronic Conditions.....	68
Online Communications.....	70
Protecting Vulnerable Communities Online: Privacy-Preserving Interventions Against Cyber-Enabled Sexual Exploitation on Social Apps	70
Social Media, Crime and the Environment	72
Understanding and Measuring Influence in Harmful Online Conspiracy Theory Communities	74
Simulation and Interaction.....	76
4D Object-Oriented Crime Scene Reconstruction and Simulation using Neural Scene Models.....	77
Cascading Vulnerabilities in AI Agent-Driven Markets	78
LLM supported threat assessment for urban mobility and cross domain response	79
SafeAR: Safety and manipulation risks of diminished reality in everyday life	80
Understanding Crime in Immersive Environments	82
User Agents and Simulations of the Physical World to Predict Cyber-Physical Risks	83
Contact	85

Introduction

The EPSRC Centre for Doctoral Training (CDT) in Cyber-Physical Risk is a national research centre led by [UCL Security and Crime Science](#) and [UCL Computer Science](#). We are dedicated to training the next generation of leaders capable of tackling the complex challenges posed by cyber-physical risks and disinformation in today's hyperconnected societies.



We collaborate with a diverse range of industry partners, academic institutions, and research centres to ensure that our work is relevant and impactful. These collaborations span across engineering, natural sciences, and the social sciences, as well as key UCL departments such as Electronic and Electrical Engineering and the Advanced Research Computing Centre. Based in central London, our research centre serves as a hub for collaboration, innovative thinking, and real-world impact.

We provide our students with a unique, cross-disciplinary research environment that prepares them to tackle critical security challenges. In a world where the boundaries between the digital and physical realms are increasingly blurred, our multidisciplinary approach ensures that our students can tackle cyber-physical security challenges from multiple angles.

This document presents the research projects offered through the EPSRC Centre for Doctoral Training (CDT) in Cyber-Physical Risk. Each topic was proposed by a UCL academic and

reviewed by our recruitment committee to ensure alignment with the CDT's vision and compliance with EPSRC requirements.

We hope you discover a project that resonates with your current skills and future ambitions. If that is not the case, we invite you to connect with relevant UCL academics from our partner departments, discuss your proposed ideas, and determine whether they would be interested in preparing and submitting a new supervisor topic proposal in your area of interest. We will then review it and assess whether it meets the CDT requirements before advertising it on our website.

Prof. Hervé Borriou

EPSRC CDT in cyber-physical risk, Director

The student journey

Our training programme offers a structured environment that balances research excellence, professional development, and leadership opportunities. Students follow a **stepping-stone PhD model**, combining intensive foundational training with independent research. This approach ensures that candidates build the technical expertise, leadership skills, and interdisciplinary awareness needed for careers in academia, industry, and policy.

Phase 0: Apply and secure funding

Every PhD journey starts with the submission of an application. Go to our [website](#) and follow the steps below to apply and secure funding for this innovative programme.

1. Select a project from this document
2. Select a funding plan
3. Develop a research plan
4. Confirm supervisor availability
5. Start your application via the applicant portal
6. Academic transcripts
7. Write your PhD Application Supporting Statement*
8. Nominate two referees
9. Submit your application via the applicant portal

*Competition for funded studentships is strong. Make sure that you use the correct form from our website to write your PhD Application Supporting Statement. If your application is shortlisted, we will invite you to attend an interview where you will have the opportunity to demonstrate your motivation, skills and plans for the PhD.

One of the highlights of the application process was the requirement of a submission of our research plans. As an applicant, this gave me confidence in both the recruitment process and program itself as at the earliest opportunity we were not only able but encouraged to present ideas surrounding such engaging themes. Further to this, the application required the planning of a leadership project, something I initially found unfamiliar. However, the possible opportunities presented within the project, such as hosting and planning conferences where I'm able to invite researchers from previous placement work is such a benefit to my PhD. Coming from a background in natural sciences I am very passionate about the importance of interdisciplinary science, and its presence within the CDT's cohort nature secured my confidence in this being the right program for me."

(Danny, Year 1 student)

Phase 1: Preparing for research (Months 1-6)

The first six months focus on laying the groundwork for student's research and entrepreneurial endeavours. We help them quickly adapt to the new environment and gain a deeper appreciation of cyber-physical risk research

Induction and networking

- A welcome event introduces students to the CDT, UCL's research facilities, and industry partners.
- The Equinox School provides students with hands-on learning with academic and industry partners from around the world exploring global perspectives in real-world security challenges.
- Students are invited to join a cross-departmental platform for students to engage in collaborative activities, organise social events, and network with fellow researchers and industry experts.

"A major benefit of the program is the access to multidisciplinary expertise across academic and industry settings. For example, we were given the opportunity to produce questions for the panel discussion led by Prof. Gloria Laycock, closing this year's International Crime Science Conference. We tailored these to the backgrounds of the panelists, setting the direction for the subsequent discussion on hybrid threats."

(Angelina, Year 1 student)



Poster presentation at the 16th International Crime Science Conference, London (November 2025)

Foundational training

- Five core courses provide a deep understanding of cyber-physical risk, security challenges, and interdisciplinary problem-solving strategies. These include:
 - Introduction to Cyber-Physical Security
 - Crime Modelling and Simulation
 - Security of Cyber-Physical Systems
 - Online Extremism and Hate Crime
 - Integrator module (including research design)
- At the 6-month mark, we assess the students' knowledge and skills in the field of cyber-physical risk through a capstone assessment. The assessment is based on their contribution to a group project.

One of the main highlights of the programme is an "integrator" module in which we learn modelling, simulation, and risk assessment techniques, working as a team to analyse a real-world cyber-physical system and utilising skills gained from other modules. The module organisers are experts in their fields and have strong connections to industry and government. Coming from a programming background, I'm particularly enjoying the computational aspect of the project. We're also building a physical model at UCL's Institute of Making to supplement our work, which is a great experience. We aim to publish our research findings in the end, making this module all the more exciting!

(Wael, Year 1 student)



Year 1 students working on a group project as part of the Integrator module (December 2025)

Phase 2: Advancing knowledge and research (Months 6-24)

This phase marks the beginning of students' doctoral research after completing training in research ethics and securing ethics approval and data.

- Students spend the next 12 months conducting research under supervision, with the flexibility to develop skills and pursue independent work.
- The systematic review or empirical work produced contributes to the first of the thesis's key elements and forms part of the PhD Upgrade assessment at month 18.
- At month 18, students submit a progress report (literature review, empirical study, work plan) and undergo a viva to assess their ability to complete the PhD within the expected timeframe.

Phase 3: Becoming an expert (Months 24-43)

In this phase, students further develop their leadership skills, improve their research dissemination, and drive real-world impact through public engagement and entrepreneurial opportunities.

- Students deepen their research expertise and expand their analytical skills through their doctoral projects, while contributing to external partners.
- They continue to engage in cohort activities and take the Responsible Research Innovation (RRI) module, along with an elective from a pool of courses across six departments.

Phase 4: Unleashing potential (Months 36-48)

The final phase focuses on completing the PhD, maximising research impact, and preparing for the next stage of students' careers whether in academia, industry, or policy.

- Through the Get it Done programme, students receive structured support from their supervisors to complete their thesis to the highest academic standards.

- A series of coaching workshops help them develop resilience, improve writing and organisational skills, and strengthen their academic identity—key factors for successfully finishing a PhD.

The Leadership project

Alongside thesis completion, students deliver a project through which they will have the opportunity to acquire and demonstrate leadership skills.

This can involve translating research insights into practical applications for organisations such as the National Crime Agency; advancing spin-out plans; undertaking a work placement in industry, policy, or a charitable organisation; or completing a research project at Yale University through UCL's exchange programme.

Next Step

Students take part in the Next Step scheme, designed to support their transition into their next career step.

- Guidance and mentorship on career pathways will be offered, as well as professional development workshops covering leadership, networking, CV-building and interview coaching to help secure research positions, policy roles, or industry placements.
- Ongoing entrepreneurial support will be provided for students launching startups or commercial ventures.

By the end of the PhD programme, students will have the research expertise, professional experience, and industry connections to make a meaningful impact in their chosen careers and in the cyber-physical risk field.



Year 1 students designing the induction event for next year's cohort (December 2025)

General entry requirements

To be eligible, you must qualify for 'Home Student' status and have achieved (or be predicted to achieve) a first-class or upper second-class honours undergraduate degree in a relevant subject (or equivalent international qualifications or experience). An appropriate Master's degree is preferred but not essential. You must also meet UCL's English language requirements. If your first language is not English, you will need to provide evidence of proficiency. Further details can be found on [UCL English language requirements web page](#). This programme requires Good level (Level 2) proficiency.

Attendance requirements

To fully engage in core teaching and cohort activities during the initial six months, students must commit to in-person study on campus for two and a half days per week. Please note that the specific required days may fluctuate annually based on scheduling needs. Students must also be available to attend events that are organised on other days, including ad-hoc meetings, seminars and conferences.

Beyond the initial six months, in-person attendance requirements will vary individually depending on the specific project and supervisor arrangements.

Part-time mode

Part-time students complete the same programme and activities* as those enrolled full-time but will typically do so over a period of 7 years. We strongly encourage all students to undertake Phase 1 on a full-time basis over six months; however, where this is not feasible, part time students may be permitted to complete Phase 1 over their first two years. Part-time students will typically be required to complete Phase 2, which includes an upgrade viva (from MPhil to PhD),

within three years of enrolment, and complete Phases 3 and 4 of the programme in the penultimate and final years of registration, respectively.

*Please note that while we aim to offer a flexible schedule, we cannot guarantee that all cohort activities will take place on days when individual part-time students are available.

Research Topics

Leading academics from UCL Security and Crime Science, UCL Computer Science, and other partner departments have proposed over 25 projects. 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](#).

"My chosen project researches the relationship between online extremist content and offline violence to identify, simulate and evaluate countermeasures that can be deployed to prevent at-risk individuals from planning and committing attacks. After working in the online harms sector for several years and retraining as a data scientist, the PhD seemed the perfect next step to further develop both my technical skills and theoretical knowledge. Alongside supportive supervisors and cohort, the first six months are already preparing me for the project with a well-structured programme that includes programming classes in agent-based modelling, lectures on theories in online extremism and crime, and thought-provoking discussions on the broader threat of cyber-physical risks."

(Ellie, Year 1 student)

Futures

Technological changes bring many benefits to society. However, for some technologies, there are unintended consequences that create new threats, which are increasingly cyber-physical. For example, adversarial attacks against the sensors of connected autonomous vehicles might lead to collisions, security vulnerabilities in smart doorbells might facilitate stalking or domestic violence, and attacks against critical national infrastructure can disrupt essential services. Acceleration in the development and adoption of new technologies means that these new threats are often only identified or addressed after attacks have taken place. Alternatively, approaches to their prevention may be insufficient, or only partially implemented before incidents occur. Regulation is an important tool that governments can use to ensure that new technologies and online services meet minimum security requirements. However, new legislation typically takes much longer to develop than does the adoption of new technologies, or the time it takes adversaries to exploit vulnerabilities in a new technology or service. The application of “futures” methods to anticipate new and future threats is hence increasingly important. PhD research conducted in the department has examined future threats facilitated by engineering biology, the internet of things, connected places, and AI. Students have collaborated with industry, law enforcement, government departments and regulators to affect real change, and have secured jobs in academia, government, consultancy, or founded their own company.

Hybrid Risks in Agentic AI Ecosystems

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

What the research is about

This research explores how agentic AI systems (intended as a society/collection of goal-directed AI agents that can act and interact independently) create new forms of hybrid risk that span the digital and physical worlds. As such systems increasingly manage complex decision-making, they can have tangible real-world consequences. Competitive or misaligned agentic AI behaviour can amplify vulnerabilities, creating hybrid risks that span cyber and physical infrastructures. Understanding these dynamics is critical to ensuring the safety, reliability, and accountability of next-generation AI ecosystems.

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 candidates with a background in Computer Science, Engineering, Mathematics and related areas or with a strong quantitative background. A strong research interest in machine learning/artificial intelligence (in particular, multi-agent systems and foundational models) is essential given the topic of the project.

Cyber-Physical Systems

Cyber Physical Systems (CPS) are engineered systems in which computational components monitor, control, and interact with physical processes. They include systems found in critical national infrastructure such as power plants, energy networks, and water treatment facilities, as well as industrial automation, consumer Internet of Things, and emerging intelligent systems such as autonomous vehicles. Ensuring the security, safety, and resilience of CPS is paramount, as cyber attacks or disruptions can directly affect individuals, public health, economic stability, and national security.

PhD students in this theme will investigate the challenges of securing and managing CPS, developing research that strengthens these systems against evolving threats. Research topics include artificial intelligence for security, including adversarial machine learning, cyber situational awareness, incident response, forensic investigations, governance and regulation, human and organisational factors, and socio-technical modelling and simulation.

You will join a vibrant research community with expertise from the Department of Computer Science and Security and Crime Science, covering areas such as crime analysis, situational crime prevention, formal verification, secure systems design, artificial intelligence, security modelling, and networks and systems. Students will have access to CPS equipment at UCL and will be supported in developing their own testbeds. You will also be part of a supportive CDT community of students tackling complex, related challenges, sharing knowledge, and working together to develop secure and resilient systems.

CPS security is a growing field with strong career prospects and a clear need for skilled researchers and practitioners.

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.

Adversarial Data Attacks on Machine Learning and Generative AI-based Process Monitoring

The supervisory team includes Dr Eike Cramer (UCL Chemical Engineering) and Dr Ingolf Becker (UCL Security and Crime Science).

What the research about

Adversarial attacks are deliberate modifications of data that aim to deteriorate the function of machine learning models, e.g. inducing false classifications. In engineering systems, there are numerous cases of data-based decision-making, particularly for process monitoring and safety. These include fault and anomaly detection systems for safety measures and to assure product quality. With the increasing usage of machine learning and artificial intelligence, the risk of attacks on such models becomes a threat to the operation of industrial processes. Adversarial attacks may be distinguished as white-box and black-box attacks based on the knowledge of the attackers. Manufacturing companies generally protect their models and their deployment, but there are instances in which companies must interact with external entities, e.g., in resource procurement. For instance, early studies have shown that adversaries can learn the company's trading behaviour on electricity markets and launch targeted attacks that lead to significant losses. Furthermore, human factors in cyber-physical system interaction frequently present vulnerable interfaces. Ultimately, the project aims to support organisations and industrial manufacturers to detect their vulnerabilities and defend against data-based attacks.

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 applicant should have a keen interest in model-based decision-making (both data-driven and

physics-based). Experience in coding and a background in engineering or other STEM fields is beneficial but not mandatory.

Hearts on the Cloud: Threat modelling the cybersecurity risks of internet-connected cardiac implants

The supervisory team includes Dr Isabel Straw (UCL Institute of Health Informatics) and 2nd supervisors Prof. Steven Murdoch (UCL Computer Science) & Dr Anish Bhuvra (UCL Institute of Cardiovascular Science)

What is the research about

Implanted cardiac devices such as pacemakers are increasingly common in our ageing population. Growing technological innovation has expanded cloud connectivity and remote monitoring across much larger patient cohorts, creating shared digital dependencies, and with them, new cyber-physical risks that directly interface with human biology. These cloud-connected ecosystems support early detection of life-threatening events but also vastly expands the cyber-attack surface in the healthcare sector. Exploits across these networks could place thousands of patients at simultaneous risk, posing novel technical and biological risks.

Existing safety regulation focuses on attack surface of the implanted device, not the end-to-end digital pathway that encompasses the full ecosystem. This research will pioneer a full-stack threat model, from the pacemaker inside the body, to the home monitoring gateway, hospital network, and cloud infrastructure, to understand how cyber vulnerabilities could manifest as physical harm. It will generate new evidence to inform next-generation safety engineering and NHS readiness.

Is this PhD project for you?

The ideal applicant will be enthusiastic about working at the frontier of cybersecurity, medicine

and safety-critical engineering and gaining broad interdisciplinary knowledge of cyber-physical risk in healthcare. You should be curious about how cyber threats translate into real physiological consequences and motivated by the responsibility of protecting human life in the most direct sense. This project would suit someone with a background in cybersecurity, computer science, engineering or a related technical discipline, with a desire to gain deep exposure to real clinical environments. No prior medical training is required, only a willingness to learn, collaborate with clinicians, and think rigorously across both digital and biological system

HART: Human-Aware Resilient Teleoperation under Adversarial

Threats

The supervisory team includes Dr Mark Colley (UCL Computer Science) and Prof. Tom Carlson (UCL Division of Surgery and Interventional Science).

What is the research about

Remote driving and remote assistance of partially automated vehicles introduces new attack surfaces: prompts can steer operator user interfaces (UIs); networks can be degraded or spoofed; command pipelines can be tampered with. This project studies adversarial inputs against human-centred teleoperation, quantifies oversight failure modes under time pressure and uncertainty, and designs interaction safeguards and provenance checks for end-to-end command chains. Mixed reality testbeds link simulated traffic and real control devices to evaluate safety, trust calibration, and recovery. Outputs include tested operator interfaces, provenance and attestation mechanisms for commands, and measurable guidance for when to hand over, slow down, or stop. Prior work on remote control concepts, digital twins, uncertainty displays, and eye tracking provides a strong base.

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. You can code in Python or C#, you want to run user studies with operators, you understand experimental design and statistics, and you think clearly about failure and recovery. Experience with Unity, transport simulation, computer vision, networking, or security is a plus. You will work with partners, follow trusted research practices, and ship usable artefacts.

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 modelling) 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 modelling, 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.

Identifying and Detecting Unauthenticated UAVs to Against Physical and Cyber Sabotage Threats

The supervisory team includes Associate Prof. Bo Tan (UCL, EEE, ICCS) and Prof. Kevin Chetty (UCL Security and Crime Science). The project is supported by Associate Prof. Matt Ritchie (UCL, EEE, Radar) with the newly developed ARESTOR multiple channel radio transceivers system and Prof. Mikko Valkama (Tampere Wireless Research Centre, Finland).

What this research is about

The rapid evolution of Unmanned Aerial Vehicles (UAVs) technology over the past decade has led to their increasing exploitation by criminals and hostile parties for cyber, physical, and hybrid attacks, threatening both national security and public safety. Physically, UAVs are used to carry and distribute explosives, hazardous chemicals, or contraband, facilitating organized crime. On the cyber front, UAVs enable unauthorized surveillance, signal jamming, and eavesdropping. Hybrid cyber-physical attacks, such as UAV-driven sabotage of communication networks, pose significant risks to security agencies and law enforcement operations. This research leverages machine learning-enhanced radio sensing and signals intelligence (SIGINT) to counter these emerging threats. It focuses on the early detection and classification of both autonomous and connected UAVs involved in criminal or malicious activities. By identifying unauthenticated UAVs in real time, the project enhances security responses, mitigates risks to critical infrastructure, and aids crime prevention efforts, supporting law enforcement and national defence strategies.

Is this PhD project for you?

The successful applicant will have a strong interest in machine learning, radio signal processing, and cyber-physical security, as well as expertise and experience in doing theoretical modelling, simulation, and experimental works. Prior experience in programming

(Python, Matlab, and/or C) and a willingness to explore innovative methods for complex problems are essential. Also, the candidate is expected to have strong writing and presenting skills in the technical and academic context. 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.

Open-Ended World Modeling for Emergent Multi-Agent Behavior

The supervisory team includes Prof. Tim Rocktäschel (UCL Computer Science) and Dr Roberta Raileanu (UCL Computer Science).

What this research is about

This project explores how generative simulation environments can give rise to adaptive, self-organizing strategies among intelligent agents. The goal is to develop open-ended world models capable of generating an unbounded stream of dynamic, multi-agent scenarios where autonomous systems learn to cooperate, compete, and evolve. These environments serve as powerful testbeds for studying emergent behavior relevant to real-world challenges in cyber-physical and defense contexts—such as swarm coordination, adversarial adaptation, and resilient strategy formation under uncertainty. Core research questions include how world models can support continual scenario generation and model-predictive reasoning, which self-play or unsupervised curricula best drive innovation, and how emergent multi-agent behaviors can reliably transfer from simulation to reality. This interdisciplinary project, in collaboration with

UCL's DARK Lab and Iconic Interactive, bridges game AI, reinforcement learning, and cyber defense to advance the science of open-ended intelligence.

In addition, we are working with a stealth 'adversarial intelligence' company. The intention of this project is to generate open-ended environments that help to train, test, and harden autonomous systems for real-world environments. Simulated environments will act as the synthetic training and proving ground for real-world capabilities. For instance, similar to how Wayve is using Gaia to train robust generalist drivers, this research will allow for a World model suited for training systems to model and intercept incoming objects at different MAG speeds, diverse conditions, swarms of systems etc

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. You should have a strong background in computer science, machine learning, with an interest in emergent intelligence, simulation, and open-ended AI. Curiosity, creativity, and a drive to explore uncharted research directions are essential. You should also bring experience with reinforcement learning, multi-agent systems, or generative modeling, and have published research before. This PhD is ideal for those who want to conduct high-impact, exploratory research bridging AI safety, cyber defense, and open-ended world modeling.

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 Reconnaissance 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.

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 has been studied in many digital systems security scenarios (e.g., malware detection, network intrusion detection), it has been less explored in hybrid cyber and physical systems, where attacks may have different risk and impact. For example, systems including home automations, 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 policy makers and regulators.

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 Borrión (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 downtime, 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.

Systemic Risk in the Internet of Things

The supervisory team will include Dr Tristan Caulfield (UCL Computer Science) and Dr Ingolf Becker (UCL Crime Science).

What the research is about

More and more connected devices are being built and installed in our environment. Securing these devices can be challenging as they may have long service lives and may be difficult to update or remove once they are deployed. The rise of IoT botnets such as Mirai shows that whole classes of vulnerable devices can be exploited. We don't yet have a good understanding of the systemic risks posed by the IoT ecosystem, and we don't know the best ways to manage or remediate them. For example, if a large number of connected home appliances in a city can be made to switch themselves on or off at the same time, it has the potential to destabilise the electricity grid. This project seeks to understand these types of risks and the roles and responsibilities various stakeholders (for example, users, ISPs, manufacturers, regulators) have in securing the IoT ecosystem.

This is a multi-faceted project that seeks to first gain a better understanding of the systemic risks in the IoT ecosystem and then understand the best way to address these. This will involve looking at technical aspects of IoT devices and protocols, their security policies and features, and how these interact with the wider environment and the incentives of various stakeholders. You will use a number of approaches throughout the project, including modelling and simulation to explore risk, as well as studies (such as interviews, surveys, or workshops) with stakeholders to understand their roles.

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. They will have a strong interest in security and in IoT. A good understanding of IoT environments (software, devices, networking) is desirable and expertise in programming, which is required for building simulations, is essential. Applicants should also be interested in learning how to conduct user studies and be familiar with statistical analysis methods.

Adversarial Attacks on AI-based Control Architectures in Critical Infrastructure Systems

The supervisory team will include Dr Francesca Boem (UCL Electronic and Electrical Engineering) and Prof. Steve Hailes (UCL Computer Science).

What the research is about

Machine Learning and AI are revolutionising our lives. They have attracted a lot of attention in the Control Systems community too, where Reinforcement Learning and Learning-based control architectures promise enhanced performance, adaptability to varying/uncertain situations and ease of deployment. However, can these methodologies be reliably deployed for the control of safety-critical systems, such as the electric grid or transportation systems? What happens if the information they use is not accurate, or even worse, if some attackers intentionally modify the data to create a physical damage without being detected? The implications of attacks in AI technologies for control applications can be tremendous but have not been properly investigated yet. It is vital to assess cyber risks and vulnerabilities of learning-based control architectures that can have an impact in the physical world. The goal is to design control schemes robust to these cyber-physical attacks, so to inform policy makers and make AI-enabled critical infrastructure systems reliable, secure and resilient.

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 candidate will also have expertise in Control Engineering and/or strong mathematical background. Experience with machine learning and optimisation will be appreciated. The candidate will have expertise in modelling and simulation and be committed to

explore interdisciplinary research at the overlap between Control Theory, Computer Science and an application field (such as energy, water, transportation, healthcare,..). Prior experience of coding (ideally Python/Matlab) is essential

Autonomous Cyber-Physical Hazard Mapping

The supervisory team will include Prof. Dimitrios Kanoulas (UCL Computer Science) and Prof. Steve Hailes (UCL Computer Science).

What the research is about

This research focuses on developing an intelligent four-legged robot capable of autonomously assessing and mapping both physical and cyber hazards in complex, high-risk environments such as disaster zones or critical infrastructure sites. Unlike conventional risk assessment methods that are slow and expose humans to danger, the proposed system integrates advanced perception, navigation, and cyber threat detection to generate a unified “hazard map.” This map combines information about physical obstacles and structural instability with data on cyber vulnerabilities, such as communication blackouts or compromised networks. By fusing these dimensions into a single, real-time situational overview, the system will enable faster, safer, and more informed decision-making for security and emergency response teams. The research is important because it addresses the growing interdependence of cyber and physical systems, providing a foundation for resilient robotic inspection technologies that enhance safety, situational awareness, and response effectiveness in critical environments.

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. You should have a strong interest in robotics, artificial intelligence, and cybersecurity, and be motivated to apply these disciplines to real-world safety and resilience challenges. This PhD is suited for candidates who enjoy both theoretical and hands-on research, including algorithm development, system integration, and experimental validation. Experience in robotic perception, control, or network security would be advantageous. If you are

driven to create autonomous systems that enhance safety, awareness, and decision-making in hazardous environments, this PhD is for you.

SafeGen - Safe Generative AI Models

The supervisory team will include Dr. Jagmohan Chauhan (UCL Computer Science) and Prof. Miguel Rodrigues (UCL Electronic and Electrical Engineering).

What the research is about

This research addresses the rising Cyber-Physical Risk posed by modern genAI models such as text-to-image and text-to-video generative models. These models can produce photorealistic synthetic media at scale, leading to major safety challenges for cyber physical systems. Integrity failures arise when deepfakes corrupt automated perception pipelines in domains such as surveillance, industrial monitoring, or robotics; synthetic media misinterpreted as real can trigger unsafe or catastrophic physical actions. Also, content-safety failures occur when harmful, biased, or violent outputs is produced that can lead to physical risks. To address these issues, this research will develop novel framework and algorithms to embed proactive safeguards directly into generative models. The framework will introduce intrinsic watermarking to guarantee deepfake detectability and controlled safety via generation aligned to safety norms. By integrating integrity and safety at generation time, this research builds foundational safeguards that preserve physical security, operational trust, and system resilience across CPS reliant on visual data

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. This PhD is well-suited for candidates passionate about building safe, intelligent algorithms at the intersection of machine learning and cyber-physical systems. You should have strong quantitative skills and an interest in cyber physical security and digital integrity. Good programming skills in Python and experience in frameworks (TensorFlow, PyTorch) are essential for developing ML pipelines. Demonstrated experience in ML is required. Knowledge in adversarial ML is helpful; curiosity and willingness to learn are essential. If you enjoy combining theory with practical experimentation to deliver safety-critical innovation, this project is an excellent fit.

Securing Future Satellite Networks

The supervisory team will include Dr Stefano Vissicchio (UCL Computer Science) and Prof. Mark Handley (UCL Computer Science).

What the research is about

Low-Earth orbit satellite networks (LSNs) that are being built are expected to be a key future asset and critical infrastructure for nation states, including the UK. This is also testified by the recently increasing investments in LSN technologies (see for example <https://www.gov.uk/government/news/16-million-for-new-projects-to-boost-uk-benefits-of-satellite-constellations>). LSNs have indeed the potential to deliver much more performant, ubiquitous and diversified network services than the terrestrial Internet. Soon, LSNs' security will therefore be crucial for future digital ecosystems, with possible consequences on individuals, industry and nation states. This project focuses on simulating, designing and rigorously assessing secure LSNs for a range of future use cases. This will encompass consideration of vulnerabilities and interactions of both physical (e.g., satellites' hardware, antennas and radio transmissions, etc.) and virtual (e.g., network control logic) components of LSNs.

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 a talented, highly motivated candidate, excited by system research, and not afraid of hard challenges. The ideal candidate would demonstrate problem solving and programming skills. Knowledge of technologies, methodologies and research work in networking, security and LEO constellations are a plus, but not strictly required.

The Balance between Efficiency, Performance and Resilience of Embedded AI Systems under Cyber-Physical Risk

The supervisory team will include Dr He Wang (UCL Computer Science), second supervisor TBC.

What this research is about

AI models are in the foundation of many systems, e.g. AVs, CCTVs and robots. However, they are vulnerable to cyber-attacks (e.g. adding purposely computed perturbation to the data) and physical attacks (e.g. specially designed stickers to traffic signs to fool self-driving systems, garments with patterns to fool person identification systems).

There is active research in improving the model resilience against the afore-mentioned attacks. However, the current solutions often lead to large models, long training time, and frequent re-training. This creates a practical problem for Embedded AI systems in e.g. robots, AVs, which have limited compute resources and require fast responses. Model compression is a promising solution, but it severely compromises the performance and the resilience.

The project will investigate the triangle of efficiency, performance and resilience of such systems by (1) theoretically analysing of the trade-offs of the triangle, and (2) building a practical deployment-centric framework for adaptive model compression.

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 applicant will be both mathematically minded and practically savvy. They will be eager to conduct rigorous research, but at the same time root their research in real-world applications. The applicant will have good programming skills and deep understanding of computer science. The applicant will be adaptive and outcome-driven, capable of managing pressure of deadlines, ambitious for producing high quality and impactful research.

Understanding the role of ‘grey zones’ in hybrid attacks against critical national infrastructure in the UK

The supervisory team will include Prof. Kate Bowers (UCL Security and Crime Science) and Prof. Matt Ritchie (UCL Electronic and Electrical Engineering).

What the research is about

‘Grey spaces’ are the areas immediately bordering critical national infrastructure. The NPSA define these areas as those within the vicinity of a venue but not under its direct control and give the example of the large foyer area adjacent to the Manchester arena in the 2017 attack². These areas, whilst beyond the perimeter and jurisdiction of organisations, have a direct impact on their security risks. This is a parallel idea to ‘buffer zone’ areas in crime science- where crime at a place is strongly linked with that in the immediate surrounding environment. In grey zones, the ownership of risk and security problems is often vague, and place management can be lacking, leading to areas that are ripe for state actor espionage, reconnaissance and attack. It is likely that such actors will consider using cyber, physical and/or cyber-physical means (‘attack vectors’) to undertake activities in grey zones at different stages during a national security incident crime commission process (crime script). Understanding the attack vectors used in grey zones should help with more tailored prevention and mitigation within these spaces.

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 applicant will need to have a strong foundation in research methods such as systematic review, experimental design, data collection, collation and statistical analysis. A

knowledge of R or Python is strongly recommended. Favoured candidates would have direct experience of the National Security Sector. The PhD student will need to be a British National.

When, Not If: Designing for Recovery from Cyber-Physical Breaches

The supervisory team includes Dr Ingolf Becker (UCL Security and Crime Science), second supervisor TBC.

What this research is about

Modern organisations operate cyber-physical systems where digital infrastructure underpins virtually all productive activity. Traditional cybersecurity focuses on prevention, yet breaches are inevitable. When Jaguar Land Rover suffered a ransomware attack, the company couldn't produce a single vehicle for over a month, costing billions and cascading through supply chains. The attack itself was unremarkable; the catastrophic impact came from inadequate recovery capability.

This research challenges the prevailing assumption that cyber resilience means stronger defences. Instead, it asks: how do we design systems and organisations that can rapidly recover from successful attacks? Current business continuity approaches inadequately address cyber-physical integration. We need architectures, processes, and organisational structures that assume breach and prioritise recovery speed. This isn't about damage limitation; it's about maintaining societal and economic function when attacks succeed. As critical infrastructure becomes increasingly digitised, recovery capability becomes a matter of national resilience.

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. This project welcomes diverse backgrounds: system architects and software engineers interested in designing resilient infrastructures, or social scientists and psychologists keen to understand organisational responses to crises. You should be curious about

challenging conventional cybersecurity thinking that prioritises prevention over recovery. Strong analytical skills are essential, whether applied to system design or qualitative research. You'll need intellectual flexibility to work across technical and organisational domains, and comfort with ambiguity as the project direction evolves based on your interests and emerging findings.

A Socio-Technical Approach for Responding to Attacks on Critical Infrastructure

The supervisory team includes Dr Nilufer Tuptuk (UCL Security and Crime Science) and Prof. Steve Hailes (UCL Computer Science). Dr Vasilios Mavroudis from the Research Centre at Alan Turing Institute is also part of the supervisory team.

What this research is about

Critical infrastructure systems, such as electricity grids, water treatment facilities, manufacturing systems and transportation networks, are increasingly targeted by malicious actors, threatening societal safety and access to essential services. While existing research focuses on detecting cyber-physical attacks, detection alone is insufficient: systems must also respond and recover in real time. Designing such mechanisms poses technical challenges and human-centred considerations, including compliance with internal and external policies, accountability, transparency, unintended consequences and the role of human oversight. Despite this critical need, safe and secure autonomous recovery mechanisms remain underexplored. This PhD addresses the gap by developing response and recovery strategies that integrate technical robustness with human-centred considerations, along with metrics to evaluate recovery success. The research aims to ensure that critical infrastructure systems operate continuously, reliably and responsibly under existing and future threats, bridging detection, autonomous response and socio-technical safety.

Is this PhD project for you?

The ideal candidate will have a strong interest in cyber-physical systems found in critical infrastructure or in system security, and a solid background in Computer Science, Mathematics, or a related subject. The applicant should have a strong interest and a willingness to engage in multidisciplinary research, including collaboration with industry, conducting simulations of

attacks, and the development and testing of solutions. Prior experience of AI (including machine learning and deep learning) would be highly desirable.

Adversarial AI for Red-Teaming Cyber-Physical Systems

The supervisory team includes Prof. Steve Hailes (UCL Computer Science), Dr Nilufer Tuptuk (UCL Security and Crime Science) and Dr Francesca Boem (UCL Electronic and Electrical Engineering)

What this research is about

This research addresses the growing cybersecurity risks in critical national infrastructure (CNI) arising from the convergence of operational technology (OT) and IT networks. Cyber-physical systems (CPS), including power grids, water treatment plants, and industrial automation facilities, are increasingly connected and AI-integrated, expanding their attack surfaces and making them targets for sophisticated cyber threats. While AI is widely adopted in cybersecurity, limited research exists on AI-driven adversaries that can autonomously generate and refine attack strategies against CPS. Current red-teaming approaches rely on manual testing, static rules, or pre-programmed simulations, which are less adaptable than AI-based methods. This PhD project will aim to develop an AI-driven red-team agent using deep reinforcement learning with causal reasoning, to autonomously identify and exploit vulnerabilities in CPS systems such as SCADA-controlled systems. By evaluating its effectiveness against state-of-the-art CPS security defences, the research will benchmark traditional protections against adaptive AI-driven attacks, providing critical insights for improving the resilience and security of essential infrastructure systems.

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 candidate will have a strong interest in cyber-physical systems within critical infrastructure or in system security, and a solid background in Computer Science,

Mathematics, or a related discipline. Experience with machine learning, particularly reinforcement learning, or a strong willingness to learn these techniques, is essential for the project. The candidate should also have expertise in modelling and simulation and be eager to engage in interdisciplinary research. Prior coding experience, ideally in Python or Matlab, is required.

Physics-Inspired ML for Safe Human-Robot Interaction using VLA Models

The supervisory team includes Dr Jagmohan Chauhan (UCL Computer Science) and Prof. Dimitrios Kanoulas (UCL Computer Science)

What this research is about

This research will tackle the safety challenges of deploying large Vision-Language-Action (VLA) models in close-proximity human-robot interaction (HRI). While VLAs can interpret visual cues and natural language to perform general tasks, their training on unstructured data means they may generate actions that violate physical laws or HRI safety constraints (e.g., excessive force or abrupt motion). To address this, this project will design a Physics-Inspired Machine Learning (PiML) framework, integrating kinematic and dynamic constraints into VLA training and execution so that all robot behaviours remain physically feasible and safe. This project will also explore Formal Verification as a high assurance to benchmark the reliability of the PiML approach. This work will enable intrinsic safety, reduces data requirements by leveraging physics priors, and limits cyber-physical risks by ensuring motion invariants are respected. Ultimately, it enhances trust, predictability, and regulatory readiness for real-world robotic collaboration.

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. This PhD is well-suited for candidates passionate about building safe, intelligent robotic systems at the intersection of machine learning and cyber-physical systems. You should have strong quantitative skills and an interest in integrating physics, control, and data-driven models. Ideal applicants will be motivated by tackling open problems in Vision-Language-Action

modelling, real-time safety filtering, and robust HRI. Good programming skills in Python and C/C++ are essential for developing ML pipelines and real-time robotic control. Demonstrated experience in ML and/or robotics is required. Knowledge in optimisation, or simulation is helpful; curiosity and willingness to learn are essential. If you enjoy combining theory with practical experimentation on real or simulated robots to deliver safety-critical innovation, this project is an excellent fit.

Developing Secure and Resilient Smart Building Systems

The supervisory team includes Dr Nilufer Tuptuk (UCL Security and Crime Science), Prof. Steve Hailes (UCL Computer Science) & Prof Jeremy Watson (UCL STEaPP).

What this research is about

This research focuses on the security and resilience of Building Management Systems (BMS), which control critical building services such as power, water, ventilation, and cooling. Failures or cyberattacks on these systems can have severe consequences, from uninhabitable offices and universities to data centre outages or life-threatening interruptions in hospitals. Despite their critical role, BMSs are often housed in major plant rooms, which are less secure than other building areas and frequently accessed by external trades for maintenance. Control panels and the wider BMS network are therefore exposed, presenting opportunities for attackers to disrupt building operations or gain access to wider IT networks. This project aims to identify and investigate the vulnerabilities of BMS technologies, analyse adversarial behaviours using deceptive honeypots, and develop a solution with best practices to integrate security and resilience into building management systems, thereby improving the protection of buildings and occupant safety against evolving cyber and physical threats.

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 strong interest in the security of building and automation systems, and a solid background in Computer Science, Mathematics, or a related subject. They should have an interest in collaboration with industry, conducting simulations of attacks, and developing and testing solutions. Prior experience in machine learning and network security, as well as familiarity with penetration testing tools and programming would be highly desirable.

Cyber-Physical Abuse Pathways in Assistive and Home Health Technologies for Patients with Disabilities and Chronic Conditions

The supervisory team includes Dr Isabel Straw (UCL Institute of Health Informatics), Dr Leonie Tanczer (Associate Professor in International Security and Emerging Technologies and UKRI Future Leaders Fellow), third supervisor TBC.

What this research is about

Patients with disabilities and long-term health conditions increasingly rely on personal healthcare technologies and “*Health at Home*” services for basic bodily safety and daily function. Common examples include hearing aids, continuous glucose sensors (CGMs), home oxygen machines, wearable cardiac tracers and fall detectors. These assistive and home health technologies enable independence, but they also create new cyber-physical risks when exploited in the context of domestic abuse, coercive control or stalking.

Historically, medical device cybersecurity research has focused on remote, anonymous attackers and technical exploits such as breaking encryption or bypassing authentication. It has rarely accounted for intimate adversaries, such as individuals with legitimate proximity, shared

devices, or embedded trust (e.g. partner, parent or carer). In such contexts, both passive misuse (e.g. passive tracking through Bluetooth emissions) and active interference (e.g. manipulating therapy delivery) are possible, without requiring advanced hacking skills, only access and intention [1]. This PhD directly fills the existing research gap by developing the first cyber-physical threat models grounded in *domestic abuse, dependency and proximity-based harm* - a critical but currently neglected priority for NHS digital safety and national safeguarding strategy.

[1] 'Suspended Sentence for Woman Who Delivered Insulin Overdose to Her Partner'. *Diabetes*, 11 Mar. 2019, <https://www.diabetes.co.uk/news/2019/mar/suspended-sentence-for-woman-who-delivered-insulin-overdose-to-her-partner-97764852.html>.

Is this PhD project for you?

The ideal student will be motivated by the challenge of protecting vulnerable people from real cyber-physical harm, especially in situations where digital technology intersects with disability, care and domestic or power-based risk. You should be excited by technically rigorous cybersecurity research, but also open to engaging with human, social and safeguarding realities that traditional security models overlook. This project would suit someone with a background in cybersecurity, computer science, engineering or a related technical discipline, who wants to work ethically on technology that directly affects patient safety. No clinical or social science background is required, only a willingness to learn from clinicians, safeguarding experts and lived experience stakeholders, and to think with both technical precision and real-world responsibility.

Online Communications

Digital ecosystems are reshaping how harmful behaviours develop. In addition to the dark web, messaging apps, social media platforms, and gaming environments are increasingly exploited to disseminate disinformation, amplify hate, and support a wide range of criminal activity. These threats are expected to grow with the rise of decentralised networks and rapid advances in large language models. Projects in the Online Communications theme will address these challenges by drawing on interdisciplinary theory and methods and contributing rigorous evidence to a rapidly evolving policy domain. Projects will help understand the scale and scope of emerging offline threats that are facilitated or enabled by online communications, such as online harms groups (The Com) that encourage arson campaigns. Additionally, projects will explore effective countermeasures. For instance, although social media platforms are deploying automated systems to detect and remove problematic content, the effect of these interventions on users' behaviour has yet to be shown, including the unintended consequences—users may migrate to other platforms or shift their behaviour offline in ways that sustain risks. The interactions between digital and offline countermeasures, including different security technologies, must be examined as well.

Protecting Vulnerable Communities Online: Privacy-Preserving Interventions Against Cyber-Enabled Sexual Exploitation on Social Apps

The supervisory team includes Dr Martin Dechant (Department of Clinical Educational Health Psychology (CEHP) and UCL Interaction Centre) Dr Mark Warner (UCL Computer Science).

What this research is about

Criminal actors increasingly use dating and location-based social platforms to coordinate drug-facilitated sexual encounters (e.g., chemsex). These cyber-enabled interactions can lead to physical risks of overdose, sexual assault, coercion, and blackmail. Typically, users do not initially seek drugs online; drugs are introduced by others on the platforms. Offenders may exploit psychosocial vulnerabilities including social anxiety and loneliness to target users who rely on digital spaces for intimacy and connection. These factors reduce confidence in identifying danger and in seeking support. Moreover, prevention services operate almost entirely offline, yet the behaviour that anchors people to harm begins online — the critical gap this project addresses. Therefore, this project investigates how hybrid threats propagate across the cyber, social, and physical layers, and how malicious actors adapt to cause harm. The aim is to develop privacy-preserving digital safety strategies that reduce physical harms while protecting the rights, dignity, and autonomy of LGBTQ+ communities and other at-risk users.

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. You will be motivated to understand how cybercrime exploits human behaviour and how platform design can reduce real-world harm, particularly for vulnerable communities targeted by hybrid threats. You will bring experience in cybersecurity and human-centred design, including user research and prototyping of web or mobile applications, alongside a strong commitment to ethical engineering. This PhD suits someone who wants their work to deliver direct social impact, advancing digital rights and practical safety mechanisms.

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.

Understanding and Measuring Influence in Harmful Online Conspiracy Theory Communities

The supervisory team includes Dr Mark Warner (UCL Computer Science) and Dr Tristan Caulfield (UCL Computer Science).

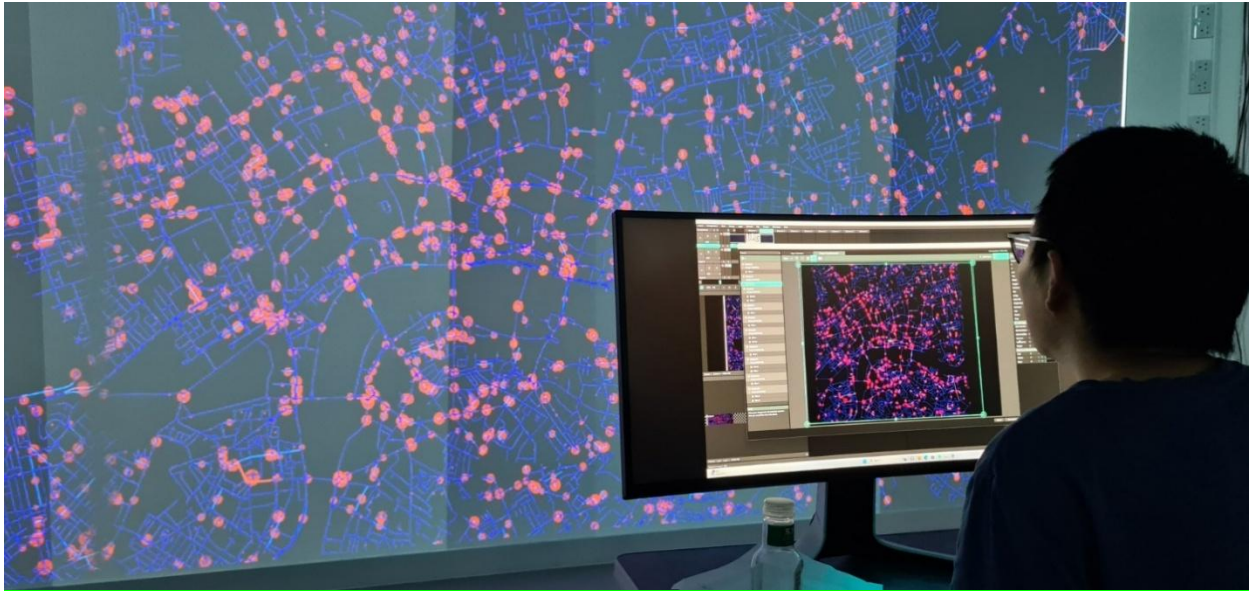
What is this research about

We have witnessed overlaps in online conspiracy theorist communities and offline violence, from QAnon and the 2021 storming of the US capitol building to 5G COVID-19 conspiracy theories and the attacking of critical communication infrastructure in the UK. Whilst prior research has investigated language use in these harmful online communities, analysis has typically been conducted either at a higher-level such as discourse analysis (i.e., what was said), or has been conducted at a community-level (i.e., understanding language within a community such as toxicity levels over time). Little prior work has focused on linguistic patterns of individual members and their power relationship within communities over time. This project will investigate the role of language as a marker of power and influence within harmful online communities. It will compare changing linguistic patterns (e.g., linguistic style matching, and linguistic variations and norms) of influential and non-influential community members and compare these to both community-level and individual-level linguistic patterns. Through the identification of members of influence, we can better understand how they influence and shape communities, and how their influence impacts risks of harm that develop from these online communities. Moreover, if early indicators of influence can be understood and measured, interventions could be developed to monitor and/or mitigate risks associated with these users.

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 applicant will have strong interest in understanding online social interactions, group relationship dynamics, and online harms. They will need to be open to applying qualitative, ethnographic methods (e.g., observations, interviews, content analysis) to help them understand the communities being studied. They must also have prior experience of coding (ideally using Python) and have either prior experience of using natural language processing tools or a desire to develop these skills.

Simulation and Interaction



Simulation and Interaction explores how immersive technologies can help us understand and improve responses to cyber-physical risks. This theme unites researchers from crime science, computer science, psychology, engineering, and human-computer interaction to investigate threats in ways that are both analytically rigorous and highly creative. Rather than relying solely on abstract models, students experiment with dynamic, lifelike, virtual environments that examine how people actually behave when digital and physical systems collide.

You will have access to UCL's Crime Science Immersive Lab, a unique four-wall projection space for building interactive scenarios, as well as VR headsets, 360-degree cameras, and a range of simulation methods including agent-based modelling and system dynamics. These tools allow you to test ideas that would be impossible, unsafe or very challenging to explore in the real world, whether examining emergency responses or designing new interventions.

Graduates from this theme develop skills that map directly onto careers in security research, technology development, emergency planning, and policy innovation. You would join a lively community of CDT students collaborating on complementary projects, contributing to a shared effort to shape safer and more resilient digital-physical systems.

4D Object-Oriented Crime Scene Reconstruction and Simulation using Neural Scene Models

The supervisory team includes Dr. Sajad Saeedi (UCL Computer Science), and the second supervisor is Prof. Hervé Borrión (UCL Security and Crime Science).

What is this research about

Reconstructing, and analysing crime scenes requires integrating evidence from both the physical and digital domains. This project proposes to develop a 4D (3D + time) object-oriented mapping framework that fuses multi-sensor data (e.g., images, LiDAR, videos, and digital traces) using advanced neural representations such as neural radiance fields and graph-based scene reasoning. The resulting “living digital twin” of a crime scene will allow investigators to interactively explore spatio-temporal hypotheses, simulate human behaviour, and test physics-based object interactions under various scenarios. The system will enable users to examine causal relationships and reconstruct events consistent with available evidence. This research will help investigators manage cyber-physical risk by linking digital evidence (e.g., CCTV, IoT data) with physical actions, ultimately enhancing situational awareness, forensic reconstruction, prevention, and decision-making in complex hybrid crime scenes.

Is this PhD project for you?

The ideal applicant will be enthusiastic about acquiring broad interdisciplinary knowledge of cyber-physical risk, AI, and digital forensics within a collaborative, cohort-based environment. You should have a strong interest in 3D computer vision, neural scene representation, and simulation technologies, together with solid programming skills (e.g., Python, PyTorch, or ROS). Experience with robotics, digital twins, or

physics-based modelling would be advantageous. This PhD is well-suited to candidates who enjoy combining technical depth with societal relevance—developing trustworthy AI systems that support law enforcement and enhance resilience against hybrid threats spanning the digital and physical worlds.

Cascading Vulnerabilities in AI Agent-Driven Markets

The supervisory team includes Dr. R Maria del Rio Chanona (UCL Computer Science), and the second supervisor is Prof. Tomaso Aste (UCL Computer Science).

What is this research about

AI agents are increasingly deployed in financial and energy markets, creating new cyber-physical attack surfaces. Evidence shows trading AIs have reduced strategic diversity (del Rio-Chanona et al., 2025) and heightened susceptibility to misinformation. These vulnerabilities let adversaries inject misinformation or targeted signals that trigger synchronised AI responses that destabilise markets within seconds. This can echo the 2010 Flash Crash, but here AI behaviour functions as an attack vector, not a mere market anomaly, via deliberate manipulation. In electricity markets, coordinated AI selling could force plants offline and trigger rolling blackouts; in natural-gas markets, “pipeline sabotage” misinformation injection could cut heating; in just-in-time manufacturing logistics, market manipulations could halt production lines.

We will build agent-based models of AI-mediated markets with varied network structures and behaviours to trace shock propagation, attack pathways, and tipping points. Insights will inform

defences—diversity-enhancing market rules, authenticated/verified data feeds, and automated “circuit-breakers”—to strengthen critical infrastructures.

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. You may come from quantitative social sciences (economics, finance, computational social science) or quantitative disciplines (computer science, physics, mathematics, engineering). You should be passionate about policy-relevant research that contributes to societal welfare and excited about complexity economics and emergent systems behaviour. We value candidates who prioritize doing rigorous, impactful research over simply publishing epsilon papers—those motivated by getting things right, understanding deep mechanisms, and producing work that genuinely informs better decision-making.

LLM supported threat assessment for urban mobility and cross domain response

The supervisory team includes Dr Mark Colley (UCL Computer Science), and the second supervisor is Prof. Hervé Borrión (UCL Security and Crime Science).

What is this research about

Hybrid threats against mobility mix online planning, coordinated narratives, and cyber events that degrade physical transportation networks. The project develops retrieval-augmented LLM pipelines that integrate OSINT (Open-Source Intelligence) and transport telemetry to forecast

potential attack scenarios on road networks in London and beyond. We leverage LLMs as proxies for real-world human behaviour, both as attackers and as simulations of human behaviour in the traffic network. It extracts event chains and simulates counterfactuals in a city twin, leveraging previous accomplishments using Simulation of Urban MObility (SUMO). Red and blue narratives co-evolve using real and LLM-based attackers to probe and harden defences.

Is this PhD 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. You are comfortable with Python, data engineering, and simulation. You think critically about evaluation, uncertainty, and safety/security. You can run human studies with operators and translate findings into system changes. Experience with reinforcement learning, retrieval augmented generation, LLM finetuning, or SUMO is helpful.

SafeAR: Safety and manipulation risks of diminished reality in everyday life

The supervisory team includes Dr Mark Colley (UCL Computer Science), and the second supervisor is Prof. Anthony Steed (UCL Computer Science).

What is this research about

Diminished reality (DR; removing objects from sight) and altered reality (AR; modifying object appearance) create new safety and manipulation risks across everyday contexts.

Malicious actors or faulty systems could hide hazards, disguise threats as benign objects, alter warning signs, or conceal people. Beyond mobility, these risks extend to workplaces, public spaces, and social settings where AR mediation affects safety-critical decisions. This project will systematically test how DR/AR manipulations impact user behaviour across multiple scenarios, then develop countermeasures that protect both users and bystanders. The research will produce a threat taxonomy, validated behavioural impact data, technical defences including cryptographic provenance and anomaly detection, and design patterns for safer AR systems.

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. You enjoy XR prototyping, safety testing, and human subjects research. You can implement computer vision filters, measure behaviour, and design clear UI responses to failure. Experience with Unity, Augmented Reality SDKs, or HCI research helps.

Understanding Crime in Immersive Environments

The supervisory team includes Prof. Hervé Borrión (UCL Security and Crime Science), Dr Dai Jiang (UCL Electronic and Electrical Engineering) & Prof Stephen Hilton (UCL School of Pharmacy)

What is this research about

This research topic is intentionally broad to allow prospective applicants the flexibility to develop a project aligned with their own interests:

This doctoral research will examine whether current theories about crime (e.g., offender decision-making, risk perception and fear of crime) and crime prevention principles (e.g., Situational Crime Prevention, Crime Prevention Through Environmental Design) are applicable to digital and hybrid spaces (e.g., the Metaverse). Part of the project will involve investigating how immersive and biosensing technologies (e.g. eye tracking, brain wave, heart rate, respiration, skin impedance) can be used to better understand how individuals perceive, interpret, and respond to events and situations in virtual spaces.

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 be curious, motivated, and comfortable working across disciplines such as crime science, psychology, computer science, and data analytics. A strong interest in experimental research, immersive technologies, and human-centred data collection is essential, along with a willingness to develop skills in quantitative analysis and programming (C, Python, Matlab). The successful candidate will enjoy collaborative research, be open to learning new methods, and be keen to apply innovative technological approaches to real-world crime and

security challenges. Prior experience of developing VR games and/or biosensing, in particular EEG, would be advantageous but is not essential, as training opportunities will be provided as part of the project.

User Agents and Simulations of the Physical World to Predict Cyber-Physical Risks

The supervisory team includes Professor Emine Yilmaz, and the second supervisor is TBD.

What this research is about

As AI is increasingly being used in the development of real-world systems such as robots, autonomous vehicles and other cyber-physical systems, it comes with additional risks caused by sophisticated (possibly AI driven) cyber-attacks that could pose risks to critical systems in the physical world. While there has been significant research on simulating and predicting possible cyber-attacks, simulating and predicting cyber-physical attacks involving different modalities is a relatively less explored area. With the recent advancements on multimodal foundational models and vision-language-action models, it is now possible to develop accurate simulators that model the physical world in which the system will be used/attacked, together with agents that simulate the different types of users who may be using (or attacking) the system *before* the system has been used in reality, which could lead to much bigger risks.

In this PhD project we propose to develop LLM based simulators which can be used to detect the possible attacks a cyber-physical system may face *before* the system has been used in reality. As part of this work, we propose to develop two different types of simulators: Initially we will focus on developing simulators that simulate a typical environment in which the system will be used, together with how a real user would interact with the system in such an environment, and detect whether there are any cyber-physical risks associated with the system in a typical environment when a regular user of the system is using it. We will further focus on developing adversarial simulators, where the goal of the simulator would be to detect any possible risks coming from adversarial attacks.

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 student is expected to be interested in cyber-physical systems and should have previous experience in working with large language models. The student should have a good understanding of foundational models, ideally including multi-modal and vision-language-action models. Good programming skills and some background in machine learning is necessary.

Contact

CDT Administrator: scs.phd@ucl.ac.uk