Cardiff Internet of Things Group

Internet of Things Garage

INTERNET OF THINGS GARAGE
Part of Cardiff IOT Group

Internet of Things Research and Teaching: Vision and Mission

Annual Report (2022)

Charith Perera (MBA, PhD)

Introducing the Internet of Things Garage

We ❤ building connected things that work...also secure, safer, and sustainable

Most of our research is ***build-driven*** and somewhat ***experimental,*** and mostly ***applied***. This means that we build things, systems, and techniques and evaluate them in real-world settings. We aim to demonstrate how they work but focus less on giving theoretical guarantees. Our work aims to produce artefacts (software systems, things) that are useful in the real world. Therefore, we felt that the name ***'IoT Garage'*** is more appropriate and describes our work well than the more traditional name ***'IoT Laboratory'***. We are not alone, see.

History: Established in December 2018 (within School of Computer Science and Informatics, Cardiff University). Previous Annual Reports: 2019, 2020, 2021

Principal Investigator: Charith Perera

As of 2022, the research group comprised 20 PhD students, 2 MPhil students (and 2 affiliated PhD students), and 3 research assistants.

Research Assistants:

Mary Zacharias
Smart Home Testbeds [2021-2022]

Akin Kaki
Forest Observatory [2021-2022]

Osian Morgan
Sensor-driven Anomaly Detection [2021-2022]

PhD Students:

Nada Alhirabi
Designing Privacy by Design IoT Applications [Since OCT 2018]

Lamya Alkhariji
Knowledge-Driven Privacy by Design for IoT [Since DEC 2018]

Areej Alabbas*
Secure Service Placement for IoT [Since JAN 2019]

Bayan Almuhander
Privacy-Aware Smart Home Data Management [Since OCT 2019]

Atheer Jeraisy
Reusable Privacy Components for IoT [Since APR 2019]

Dominic Fonseca
Low-Cost Reliable Multi-Sensor People Counting [Since OCT 2020] [MPhil]

Asma Irfan (PT)
Adapting to Discomfort Towards Sustainable Built Environments [Since JAN 2020]

Hakan Kayan
Context-Aware Security for Cyber-Physical Systems [Since JAN 2020]

Naeima Hamed
Semantic Data Integration For Forest Observatory [Since JAN 2020]

Yasar Majib
Context-Aware Security for Smart Homes [Since OCT 2020]

Reem Aldhafiri
Cyber-Physical Privacy for Ageing [Since OCT 2020]

Mark Butterworth (PT)
Low Power IoT Infrastructure for Harsh Environments [Since OCT 2020]

Wael Alsafery
Layered Framework Towards Resilient Smart Buildings
[Since JAN 2021]

Yaser Awwad
Video Analytics for Anomaly Detection
[Since JUL 2021] [MPhil]

Norah Albazzai
Augmenting Anomaly Detection with Tiny Cameras
[Since JUL 2021]

Suhas Devmane
Talking Buildings: Smart Building Pattern of life
[Since OCT 2021]

Fatmah Alqarni
Learning Privacy and Laws Through AI-Mediated Exploration and Design
[Since APR 2022]

Rayan Binlajdam
Forest Health Index
[Since OCT 2022]

Omar Mousa
End-User Development for Linked-Data Observatories
[Since JAN 2021]

Abdulaziz Aljohani
Self-Configuring Anomaly Detection IoT Architecture
[Since JUL 2021]

Azhar Alsufyani
Context-Aware Knowledge-driven Cyber-Physical Security
[Since OCT 2021]

Mohammed Alosaimi
Evaluation Framework for Anomaly Detection
[Since OCT 2021]

Siyuan Li
Adaptive Mobile Sensing within Buildings
[Since OCT 2022]

* Affiliate PhD students: Omer Rana is the primary supervisor for Areej Alabbas.

Annual Summary for 2022

- The research group is continued to grow around three research themes related to the Internet of Things (IoT) with a significant emphasis on build-driven research methods. Three PhD students have started projects across these themes this year.
- Smart Home lab is now being completed with 170 devices installed and growing. Started deploying a 35-node IoT edge sensing network on the 5th floor of the Abacws with over 15 different types of sensors on each node
- Two BSc students went to DGFC (www.dgfc.life) to gain international experience and deploy, test and gain feedback on their dissertation projects (Vlog). Two PhD students also visited to gain domain understanding and conduct user duties related to their projects (Vlog).
- Created a 6-week-long Cyber-Physical Smart Home dataset in collaboration with Buildings Research Establishment (BRE) funded by PETRAS.
- Got selected and progress through all the phases of the Cyber Security Academic Startup Accelerator Programme (CyberASAP) to explore the commercialisation opportunities of some of the outcomes of the GCHQ National Resilience Fellowship
- Developed a new MSc module on '*Edge Analytics*' in collaboration with IIT Ropar and IIIT Kottayam and received a second round of funding to expand the module in collaboration with Galala University in Egypt, funded by UK-Egypt Trans-National Education (TNE).
- Delivered the IoT module (BSc and MSc) in its full format for the first time, including IoT systems development coursework. Winners did an amazing job (Demo)
- IoT Products Library for research and teaching purposes (YouTube Playlist)
- For the latest publications, visit Google Scholar Research Outputs

Research Vision

Research Interests: Our research primarily focuses on three research questions:

1. How can we build an efficient and effective sensing infrastructure to acquire and use sensor data to better understand and improve ourselves (individuals), our surroundings (homes), our communities, and the world?
2. How can we encourage sensor data sharing in order to achieve (1)?
3. How can we achieve (1) and (2) without compromising safety, privacy or security?

The research group is formulated around research themes as follows:

Figure 1: Primary Research Themes

Privacy Fluid: This theme aims to develop a shared Privacy Mindset through an AI-mediated assistive layer to reduce stakeholder breakdown. The objective is to develop a unified framework and methodology that captures privacy-related information throughout the software development life cycle (i.e., from concept to implementation) and the product life cycle (i.e., from onboarding to disposal). For example, Privacy Fluid will support Privacy by Design (PbD) activities by assisting designers through design tools during the design phase. It will then interact with the developers through development tools to support implementing these privacy-protecting measures. Subsequently, privacy fluid will interact with end-users by assisting them in configuring privacy settings. Such a unified approach can significantly enhance privacy protection due to shared knowledge and provenance.

Data Observatories: This theme aims at developing open data observatories across different domains ranging from smart cities to wildlife conservation to understand how we can make data available for citizen scientists and other end users. We use knowledge-based AI techniques such as Linked-data and semantic web to support end-users to extract knowledge without significant technical expertise while supporting interoperability and provenance.

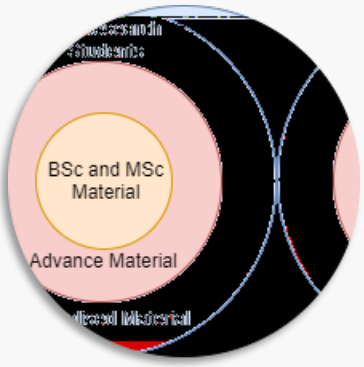
ResilientSensing.AI: This theme explores how we could add layers of resiliency to built environments (and beyond, such as smart city infrastructure) using IoT technologies (e.g., sensors). Smart environments bring both efficiency and convenience; however, they are also vulnerable to attacks and malicious activities due to connectedness. Resilience means the ability and the capacity to recover from cyber-physical attacks (detect, mitigate and recover)

Learning Technologies For the Internet of Things: This theme aims to enhance teaching activities. We aim to understand how to teach IoT to different audiences (from high school to university students and beyond) with different skill levels and innovative tools and techniques. We aim to incorporate conversational AI and personalised learning into teaching and learning experiences to facilitate large student cohorts.

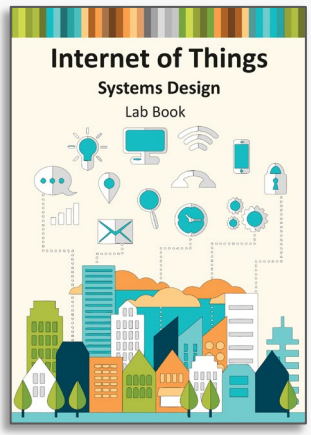
Teaching Vision

At the undergraduate level, the Internet of Things related content is delivered (to second-year students) through a module titled ***CM2306 Communication Networks***. IoT is delivered through a dedicated module titled ***CMT223 Internet of Things: Systems Design*** at the postgraduate level. Both modules are (mostly) identical in terms of delivery and content. However, expectations (from an assessment perspective) are higher at the postgraduate level (link) (YouTube Playlist).

Content: The IoT content is structured under eight themes, namely, (1) *Applications and Use cases*, (2) *Architectures*, (3) *Sensing and Actuation*, (4) *Networking and Communications*, (5) *Data management and analytics*, (6) *Privacy and Security*, (7) *Human Factors and Interactions*, and (8) *Design Strategies and Prototyping*. Each section gets delivered through one or more lectures (which include dedicated slide decks).



Modularity and Complexity: The content under each theme is developed in a modular and layered fashion based on the complexity of the content. This means that each topic has a certain amount of content that delivers the basic information to the students, sufficient to complete both undergraduate and postgraduate modules. However, if a student interested in learning more, they can follow advanced material and learn by themselves. *Advanced materials* are structured and delivered in a similar fashion to the basic material (at times embedded within basic material but are clearly marked) and provide close guidance on following up and self-studying the material. *Specialist materials* are less structured and less organised. they are delivered through either seminars or tutorials (pre-recorded or in-class). Advanced and specialised material may help the students complete the assignments in a much higher quality but not mandatory. Specialised material may be useful for new research students to advance their knowledge.



Labs and Practical: As an applied module, students are expected to complete at least six lab sessions. Students are provided with the lab book explaining each practical session step by step.

Research with BSc and MSc students: Most of the dissertation projects we offer are research-oriented. These projects are usually aligned with existing projects we are working on, at a given time, through either PhD students or research associates. However, we also use these dissertation projects to initiate some high-risk projects or new research directions. Our students are encouraged (and supported) to produce research output (such as conference, workshop paper, and poster).

Dissemination and Community Engagement

IOT Garage TV (bit.ly/2Md8vJE)

YouTube (and similar platforms) has increasingly become a mainstream content distribution stream that provides large audience access. As a build-driven research group, demonstrations are a key part of our dissemination strategy and increase awareness. Therefore, we have created a dedicated YouTube channel to disseminate our work. We believe visual medium can efficiently and effectively motivate our students to complete their high-quality project work. YouTube videos on our channel also act as a gauge for prospective students. For example, video help students decide what kind of project they want to do and the quality of the output they may want to produce. We also use the YouTube channel as a part of our reproducibility and knowledge transfer strategy. We strongly encourage students to create screencasts so that other students can understand what has been done and how. This allows next year's students to take the projects forward. Screencasts also help students provide valuable insights about their projects to their fellow students, which might not be feasible in traditional documentation approaches.

IOT Garage News (@IOTGarageNews)

As complementary to the YouTube channel, Twitter has increasingly become one of the primary ways people consume news updates. We maintain a Twitter account to broadcast updates about our group activities, including research updates, student successes, public engagement, etc.

IOT Garage Code (@IOTGarage)

We take reproducibility and '*building on top of previous work*' very seriously. As a supplement to the screencast, we also encourage organising and sharing their code through Gitlab (or similar). We actively maintain code repositories produced by each student related to each project.

Group Website (iotgarage.net)

We maintain a group website to disseminate the outcomes of different types of projects to a wider audience. Projects can be varied from BSC, MSc, and PhD, to funded projects. We provide all the relevant information under each project, including team members, funder, partners, project demos, publications links, and code repositories.

Funding Support (On-Going / Completed within 2022)

(Principle-Investigator)
Total: 91,279 GBP

A smart home study carried out by NCC Group and the Global Cyber Alliance recorded more than 12,000 attack attempts in a week, including 2,435 specific attempts to maliciously log into the devices. Once they gain control, they can manipulate the device's capabilities, including network communication, actuation, and sensing. Our project aims to develop and commercialise a resilient cyber-physical anomaly detection framework to detect cyber-physical malicious activities within built environments (i.e., smart homes and buildings). We aim to achieve this by combining heterogeneous data streams, such as external sensor data observations and energy consumption data, with traditional Network Traffic Analysis (NTA) techniques. This commercialisation project aims to explore market potential and develop a proof of concept for the proposed approach.

UK – Egypt Trans-National Education (TNE) Grant: Edge Analytics 2.0

(Co-Investigator)
Total: 29,989 GBP

Due to the popularity of edge computing and IoT, there is a huge interest and demand in the market for skilled professionals. The "Edge Computing and Analytics 2.0" module was chosen to fulfil the demand raised by both industry and academia. We will develop course content based on world-class course frameworks between the UK and Egyptian universities. Joint students and researchers will collaborate to face climate change by building edge applications on smart cities, smart grids, and water and energy management to achieve SDG goals. This project aims to extend the module we developed previously.

Going Global India – Exploratory Grants – Edge Analytics

(Co-Investigator)

877673574
Total: 20,000 GBP

This project aims to create a postgraduate module focusing on *Edge Analytics*. The course content will be co-created between the UK & India (including engagement of students & industry) based on world-class course assessment frameworks. Course delivery will follow a hybrid offline/online model. Non-credit-bearing content will be trialled with students at the three participating institutions. The proposing team has complementary expertise in - complex systems and IoT (Cardiff), edge computing and sensor networks (IIT Ropar) and IoT analytics (IIIT Kottayam).

EPSRC PETRAS 2 – Internal Strategic Projects and Engagement Fund (ISPEF)

(Co-Investigator)

EP/S035362/1
Total: 74,282 GBP
Cardiff: 74,282 GBP

This project will integrate the outcomes of the PETRAS-funded *Integrity Checking at the Edge (ICE)* project into a prototype operational decision support mechanism at Thales UK. Thales offers an end-to-end Autonomous Logistics Supply that combines an intuitive digital twin interface, unmanned command and control system, and an autonomous, all-terrain Unmanned Ground Vehicle (UGV) system with its own networked communications systems. UGVs are vital for supporting humanitarian rescue and relief efforts in unsafe environments – for example, in regions of natural disaster or conflict settings.

Institutional Sponsorship-International Partnerships

(Principle-Investigator)

(Internal)
Cardiff: 34,349 GBP

EPSRC funds this Institutional Sponsorship award for international partnerships in order to support the pursuit and development of global research partnerships. The primary objective of this project is to develop a better understanding of how to design and deploy a sustainable IoT network in a remote jungle environment with harsh conditions. We want to understand what kind of IoT network would ideally be suited to establish a forest observatory to enable sustainable sensors data collection and wireless communication. We would like to understand potential network design and topology, estimated costs, energy requirements, and other constraining factors that may need to consider when deploying an IoT network in a jungle.

Scalable Circular Supply Chains for the Built Environment

(Co-Investigator)

EP/V042017/1
Total: 368,095 GBP
Cardiff: 424,033 GBP

This project will demonstrate how one of the largest industries in the UK can utilise a digital platform to harness the benefits of a sustainable circular supply chain to reduce waste, increase safety, and promote greater fiscal responsibility. The Architecture, Engineering & Construction (AEC) sector plays a crucial role in the UK economy by employing over 2 million people to deliver civil engineering projects that underpin our economic growth. One of the biggest contributors to GDP, the ACE sector represents commercial activity spanning individual contractors through to multi-national corporations collaborating through complex asset distribution networks that account for over £10 billion of trade. This project is a collaboration between Cardiff University and Newcastle University.

EPSRC PETRAS 2 - Opportunity Fund

(Principle-Investigator)

EP/S035362/1
Cardiff: 133,833 GBP

This secondment aims to explore how to add layers of resilience to built environments in the Internet of Things (IoT) context. Smart-built environments such as smart homes and office buildings heavily depend on IoT systems to reliably sense and monitor their surroundings. Such dependencies also create high risk. Malicious parties could tamper with these IoT devices and systems to report incorrect data to control systems. Such unreliability could create a significant risk (even be catastrophic and fatal) to build environments. Therefore, we need to develop a resilient built environment by creating multiple layers of resiliency.

EPSRC PETRAS 2 (National Centre of Excellence for IoT Systems Cybersecurity)

(Co-Investigator)

EP/S035362/1
Total: 13,850,000 GBP
Cardiff: 290,920 GBP

The PETRAS National Centre of Excellence for IoT Systems Cybersecurity is a consortium that connects 12 research institutions with outstanding expertise

Partners

Awen Collective

Awen Collective develops software for critical infrastructure (water, energy, transport, etc.) and manufacturers to reduce cyber-attacks and cyber-threat.

Altifio

Altifio is an innovation hub. It aims to discover or create innovative technology solutions for companies to help them in a variety of ways, such as competing in a new market, acquiring customers at lower costs, or developing new tools to do work more efficiently.

Airbus Group

Airbus Group Innovations are industry leaders in industrial control system (SCADA) security and have a well-equipped testbed at their Newport site.

Building Research Establishment

The Building Research Establishment (BRE) is a centre of building science in the United Kingdom, owned by a charitable organisation, the BRE Trust. BRE provides research, advice, training, testing, certification and standards for public and private sector organisations in the UK and abroad.

Danu Gurang Field Center

Danau Girang is a collaborative research and training facility managed by Sabah Wildlife Department and Cardiff University.

Exalens

Exalens protects digital manufacturing against downtime and safety incidents through early warning of both system malfunctions and cyber security breaches. Manufacturers enhance their operational resilience with ground-breaking cyber-physical security analyst AI with automated incident detection and response.

Government Communications Headquarters

GCHQ is an intelligence and security organisation responsible for providing signals intelligence and information assurance to the government and armed forces of the United Kingdom.

Innovate Trust

Innovate Trust provides support and guidance to disabled people. In addition, Innovate Trust provides support to the elderly, young, disadvantaged, and vulnerable members of the local community through our Student Volunteer projects.

iPoint

iPoint aims to simplify fleet and data management across the transport industry by unlocking and correlating information from multiple platforms and networks by developing a single transport management platform.

My Data Fix

UK-qualified corporate and finance lawyer with regulatory expertise gained from an international career. My Data Fix specialises in all aspects of data privacy, having worked as the Global Data Protection Officer for an international organisation whose business is personal data.

PETRAS National Centre for Cyber Security

The PETRAS National Centre of Excellence for IoT Systems Cybersecurity is a consortium that connects twelve research institutions with outstanding expertise in securing the connected world.

Thales

Thales has established a new National Digital Exploitation Centre (NDEC) in Blaenau Gwent. The £20 million site will be used for digital and cyber security training and research facilities, providing lab space for SMEs and microbusinesses to test and develop digital concepts.

Vortex IoT

Vortex IoT builds sensors and networks for harsh environments where conditions are hostile, and power supply is limited, AI is needed & data security is critical.

PRIVACY FLUID

Interactive Design Method for Augmenting Software Design Process Toward Privacy-Aware Internet of Things Application Designs

Researcher: Nada Alhirabi (PhD Student-2018-2023)

Internet of Things (IoT) applications development and design process is more complicated than others, such as desktop, web, or mobile. That's because IoT applications need both software and hardware to cooperate across multiple nodes with different capabilities. Moreover, it requires different software engineers with different expertise to cooperate (e.g., frontend, backend, and database). Due to the above complications, non-functional requirements like privacy tend to be overlooked.
Thus far, privacy concerns have not been explicitly considered (i.e., in a unified manner), despite isolated solutions (i.e., a specific technique that address specific privacy challenge) in software engineering processes when designing and developing IoT applications, partly due to a lack of Privacy-by-Design (PbD) methods for the IoT.

This project's primary objective is to efficiently, effectively, and collaboratively develop an interactive design method (facilitated through a tool) that incorporates privacy-preserving techniques into the early phases of the software development life cycle. We envision our tool to be collaboratively used by business analysts, requirement engineers, user experience designers, and software engineers while creating privacy by design IoT application designs. This project composed of three objectives:

- Review the existing design notations, models, languages and tools that facilitate capturing and integrating non-functional requirements (i.e., security and privacy).
- Co-Design an interactive privacy-aware Internet of Things application design methodology towards reducing breakdowns
- Evaluate the efficiency and effectiveness of PRIVACY PARROT (Privacy by Design for the Internet of Things) as a tool for augmenting software engineers' capabilities and enhancing privacy knowledge.

PRIVACY PARROT

Partners and Relevant Projects

PETRAS MY DATA FIX

Outcomes

- **[Journal]** Nada Alhirabi, Omer Rana, and Charith Perera. 2021. **Security and Privacy Requirements for the Internet of Things: A Survey**. ACM Trans. Internet Things 2, 1, Article 6 (February 2021), 37 pages. PDF BIB
- **[Demo]** Nada Alhirabi, Omer Rana, and Charith Perera, **Demo Abstract: PARROT: Privacy by Design Tool for Internet of Things**, In Proceedings of the 2022 IEEE/ACM Seventh International Conference on Internet-of-Things Design and Implementation (IoTDI) 2022, 107-108
  PDF BIB SOURCE VIDEO VIDEO
- **[Poster]** Nada Alhirabi, Stephanie Beaumont, Omer Rana, and Charith Perera, **Privacy-Patterns for IoT Application Developers**, In Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp/ISWC '22) (In-Print)
  PDF BIB SOURCE POSTER

PRIVACY FLUID

Augmenting Software Design Processes by Developing Knowledge-based AI Technique Towards Assisted Privacy-aware Internet of Things Application Designing

Researcher: Lamya Alkhariji (PhD Student-2018-2023)

Internet of Things (IoT) applications need both software and hardware to cooperate across multiple nodes with different capabilities. Moreover, it requires different software engineers with different expertise to cooperate (e.g., frontend, backend, database). Due to the above complications, non-functional requirements like privacy tend to be overlooked. Thus far, privacy concerns have not been explicitly considered (i.e., as unified manner), despite isolated solutions (i.e., a specific technique that address specific privacy challenge) in software engineering processes when designing and developing IoT applications, partly due to a lack of Privacy-by-Design (PbD) methods for the IoT.

This project's primary objective is to develop a Knowledge-based AI technique that assists software engineers by automatically incorporating Privacy by Design (PbD) techniques into a given IoT application design. This project is composed of three main objectives:

- Review and synthesise privacy by design schemes through curating and systematically analysing existing privacy strategies, guidelines, principles, and patterns in the context of IoT.
- Semantically model privacy patterns and IoT systems using knowledge-based AI techniques towards the automated assignment.
- Develop and Evaluate the efficiency and effectiveness of PRIVACY CAPTAIN (Context-Aware Privacy Assistant for the Internet of Things) as a tool for augmenting software engineers' capabilities and enhancing privacy knowledge.

Partners and Relevant Projects

PETRAS

Outcomes

- **[Journal]** Lamya Alkhariji, Nada Alhirabi, Mansour Naser Alraja, Mahmoud Barhamgi, Omer Rana, and Charith Perera. 2021. **Synthesising Privacy by Design Knowledge Toward Explainable Internet of Things Application Designing in Healthcare**. ACM Trans. Multimedia Comput. Commun. Appl. 17, 2s, Article 62 (June 2021), 29 pages. PDF BIB RESOURCES SOURCE
- **[Journal]** Lamya Alkhariji, Suparna De, Omer Rana, Charith Perera, **Semantics-based Privacy by Design for Internet of Things Applications**, Future Generation Computer Systems (FGCS), Volume 138, January 2023 PDF BIB RESOURCES SOURCE
- **[Poster]** Lamya Alkhariji, Suparna De, Omer Rana, and Charith Perera, **Ontology Enabled Chatbot for Applying Privacy by Design in IoT Systems**, In Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22). Association for Computing Machinery, New York, NY, USA, 3323–3325 PDF BIB SOURCE POSTER

PRIVACY FLUID

Augmenting Software Engineers' Capabilities Towards Developing Privacy Law-Friendly Internet of Things Applications using End-User Development Paradigm.

Researcher: Atheer Jeraisy (PhD Student-2019-2023)

Internet of Things (IoT) applications development and design process is more complicated than others, such as desktop, web, or mobile. That's because IoT applications need software and hardware to cooperate across multiple nodes with different capabilities. Moreover, it requires different software engineers with different expertise to cooperate (e.g., frontend, backend, database). Due to the above complications, non-functional requirements like privacy tend to be overlooked.

In order to address this issue, we need to find a way to support and motivate software developers. In this project, we primarily focus on privacy. We aim to address this problem using two methods. First, we need to develop easy-to-use privacy-preserving software components (some form of modules) that developers can incorporate into their IoT application development process. These privacy-preserving components should be reusable and generic enough to be used across multiple domains and applications. Furthermore, these privacy-preserving techniques should be integrated into existing IoT software development tools (i.e., popular IDEs and software frameworks). Secondly, we will use gamification techniques to motivate software developers to incorporate more and more reusable privacy-preserving components within their IoT applications. This gamification framework will also be integrated into popular IoT software development tools. This project has three main objectives:

- Systematically analyse privacy by design schemes to find out how they can be used to satisfy and comply with privacy laws around the world in the context of IoT.
- Explore how different types of privacy by design schemes and elements within them (such as privacy strategies, principles, guidelines, and patterns) can be transformed into reusable privacy-preserving components.
- Based on the above findings, we aim to develop a series of reusable privacy-preserving components that can be easily adapted into the IoT application development process.
- Develop a framework to examine and operationalise each privacy-preserving component in order to quantify them towards developing a gamification-based education method.

Partners and Relevant Projects

Outcomes

- **[Journal]** Atheer Aljeraisy, Masoud Barati, Omer Rana, and Charith Perera. 2021. **Privacy Laws and Privacy by Design Schemes for the Internet of Things: A Developer's Perspective**. ACM Comput. Surv. 54, 5, Article 102 (June 2022), 38 pages. PDF BIB RESOURCES SOURCE
- **[Demo]** Atheer Aljeraisy, Omer Rana, Charith Perera, **Canella: Privacy-Aware End-to-End Integrated IoT Development Ecosystem**, Work-in-Progress VIDEO

PRIVACY FLUID

Interaction Methods for Privacy Preferences Management in Shared Spaces

Researcher: Bayan Almuhander (PhD Student-2019-2023)

The balance between protecting users' privacy while providing cost-effective functional and usable devices is a key challenge in the Internet of Things (IoT) industry. The primary user interface in traditional desktop and mobile contexts is a screen. However, in IoT, screens are rare or very small, which invalidate most traditional interaction approaches (i.e., popup notifications).

We examine how end-users interact with IoT products and how the IoT devices convey information back to the users, particularly regarding their data (i.e., How IoT devices manage data about end-users). We explore how individuals with a non-technical background can be notified about the privacy-related information of the spaces they inhabit in an easily understandable way.

This project's primary objective is to develop a tangible device that facilitates interactive privacy preferences management of IoT devices in shared spaces such as smart homes. We envision our 'PrivacyCube' as an enhanced privacy notice for the IoT devices and assist people in making informed privacy decisions and increasing privacy awareness. This project has three objectives:

- Review the various methods available to notify the end-users while considering the factors that should be involved in the notification alerts within the physical domain.
- Develop a tangible interactive device that serves as a privacy notice and visualises how IoT devices manage data in shared spaces such as smart homes.
- Evaluate the effectiveness of the PrivacyCube towards increasing privacy awareness and privacy preference management in shared spaces such as smart homes.

Partners and Relevant Projects

PETRAS

Outcomes

- **[Technical Report]** Bayan Al Muhander, Jason Wiese, Omer Rana, Charith Perera, **Interactive Privacy Management: Towards Enhancing Privacy Awareness and Control in Internet of Things,** Technical Report, 2022 PDF BIB
- **[Demo]** Bayan Al Muhander, Omer Rana, Nalin Arachchilage, and Charith Perera, **Demo Abstract: PrivacyCube: A Tangible Device for Improving Privacy Awareness in IoT**, In Proceedings of the 2022 IEEE/ACM Seventh International Conference on Internet-of-Things Design and Implementation (IoTDI) 2022, 109-110 PDF BIB SOURCE VIDEO VIDEO

PRIVACY FLUID

Privacy Considerations when Designing Smart Home Systems to Facilitate Independent Living for Ageing

Researcher: Reem Aldhafiri (PhD Student-2020-2024)

We live in the revolution of smart home devices such as smart speakers, lighting and thermostats, which are rapidly developed and adopted by different people. Those devices collect, process, and disseminate end-user data to facilitate different functionalities, such as recommendations and automation. These functionalities are typically convenient and efficient to the environments in which they are being deployed. For example, a smart lighting system may automatically configure its setting to reduce energy consumption while providing optimal service to the end users. However, such functionalities require them to monitor end-user behaviour and track their whereabouts, moods, preference, etc.

Smart devices are connected and share data to achieve a common goal. We can consider that some devices have sensitive data, such as the house's location, that can negatively affect the household's life. People (especially older adults and vulnerable people) face violating their privacy if data collection practices deviate. Some studies show that the elderly have privacy concerns and avoid using smart devices to monitor them. Privacy concerns are one of the most significant barriers to using the monitoring device in a smart home. Older adults, especially those with mild cognitive impairment, are vulnerable to privacy violations as they may not configure their privacy preferences.

This project focuses on privacy and data protection in smart homes and users of vulnerable communities by using physical artefacts. We focus on augmenting existing smart home systems and their privacy configuration mechanisms to improve privacy and data protection among vulnerable groups and help them better configure their privacy and data protection requirements. The main objectives of the project are:

- Review existing work of designing privacy-aware Internet of Things for older people
- Co-Designing privacy needs of Internet of Things systems to support successful ageing
- Rethinking Privacy-Awareness in Connected Homes: Design and Evaluation of a Privacy Toolkit Towards Augmenting Older People

Partners and Relevant Projects

Innovate Trust PETRAS

Outcomes

- [Demo] Reem Aldhafiri, Georgina Powell, Elizabeth Smith, Charith Perera, **Voice-Enabled Privacy Assistant Towards Facilitating Successful Ageing in Smart Homes,** *(Work-in-Progress)* VIDEO

PRIVACY FLUID

Facilitating Novice Software Engineers to Learn Privacy by Design and Privacy Laws through AI-Mediated Exploration and Design

Researcher: Fatmah Alqarni (PhD Student-2022-2026)

Internet of Things (IoT) applications development and design process is more complicated than others, such as desktop, web, or mobile. IoT applications need software and hardware to cooperate across multiple nodes with different capabilities. Moreover, it requires different software engineers with different expertise to cooperate (e.g., frontend, backend, database). Due to the above complications, non-functional requirements like privacy tend to be overlooked. One way to address this problem is to better educate novice software engineers about applying privacy-preserving measures in their IoT systems design process. Currently, universities are mostly focusing on teaching cyber security than privacy. Therefore, novice software engineers have very limited knowledge of designing a privacy-aware system, especially when collecting sensitive information using sensors in IoT applications.

In this project, our focus is to develop a technique (formulate as a tool) driven by AI to help novice software engineers to learn privacy and privacy laws by using design activities. The novice engineers will use our tool to implicitly and explicitly help them understand how to incorporate privacy-preserving design features into their IoT system. It is important to note that our focus is on enhancing novice engineers' teaching and learning experience. We do not aim for the proposed technique to be used in the context of industrial software engineering. However, we believe that the knowledge that noise of ranging is gained from interacting with our tool will enable them to apply privacy-preserving measures in an industrial setting. As a community, privacy researchers have developed a large number of privacy-preserving measures identified by various names, such as privacy principles, guidelines, strategies, goals, and patterns (which add up to 168 privacy-preserving measures in total).

Each of these privacy-preserving measures is varied in granularity; some are very high level, and others are low level (close to implementation). Despite investing a significant amount of resources over many years (e.g., privacypatterns.org, privacypatterns.eu), there aren't any unified mechanisms at the moment to help novice software engineers learn how to apply those privacy-preserving measures in their designs in a meaningful way. One of our key objectives is to encapsulate all this knowledge into a tool where novice engineers will incrementally learn how these heterogeneous sets of privacy-preserving measures could be potentially used to preserve privacy and comply with privacy laws. This project has three main objectives:

- Conduct literature review on intelligent design tools and how they implicitly and explicitly contribute to enhancing the understanding of the domain knowledge of the end users.
- Develop a technique that highlights privacy risks and explains what kind of privacy preserving meassure could be applied in a given context (i.e., IoT system Design) while allowing novice software engineers to learn more about privacy and legal compliance better.
- Evaluate the efficiency and effectiveness of the proposed techniques as well as scalability, extendibility, usability and so on.

Partners and Relevant Projects

PETRAS

Integrity Checking at the Edge

Team: Matthew Nunes, Pete Burnap, Charith Perera (2019-2022)

Industrial Control Systems (ICS) is the all-encompassing term to describe Distributed Control Systems (DCS) and Supervisory Control And Data Acquisition (SCADA). DCS referred to the systems connecting sensors and actuators and is controlled locally at a plant. In contrast, SCADA refers to systems used to control and manage geographically remote communication systems. Security has not traditionally been given much attention within ICS environments since they have not faced many threats because they have not been connected to the Internet in the past. Besides, a wide range of proprietary protocols is used in ICS environments that are not as well known, thereby giving the illusion of security. It is still a very relevant topic, as attacks against ICS environments increased by 110% as of 2016.

When designing an IDS for an ICS environment, the most important factor is its impact on the overall performance. As ICS environments tend to be hard real-time environments, even the smallest delay introduced by an IDS can have catastrophic effects. Therefore, particular care must be taken when determining how the IDS should intercept data, as delays render the solution unusable. Additionally, despite their widespread use within regular IT environments, Signature-based IDS are largely obsolete within environments. This is due to the wide range of devices and protocols used within ICS. Digital Bond provides the most well-known set of IDS rules for SCADA. Its support is limited by the type of devices and protocols it recognises, which is far from exhaustive.

To help with the uptake of IDS solutions within an ICS environment, it is important that operators can trust the system. To gain their trust and make actionable decisions, it is essential that they clearly understand the IDS solution operates and what informs its decisions. To this end, we review visualisation solutions of both network traffic and ML algorithms to understand the best way to communicate information about them. This will allow us to create a holistic solution that can (i) recognise malicious behaviour and pass on the information to an administrator in a manner that will give the administrator confidence in its conclusions and (ii) provide relevant detail about the malicious activity so the administrator can determine the most appropriate course of action for remediation.

- Develop an explainable IDS for ICS in an OT context that would enable security operations teams to drill into an alert and identify security concerns and suitable mitigation solutions.
- Develop a dynamic method that allows an analyst to interrogate it in real-time. The method chosen will be open-source.
- Develop an explainable solution complementary to the leading algorithm(s) for ML-based attack detection in OT.

Partners and Relevant Projects

PETRAS

Outcomes

- [Demo] Matthew Nunes, Pete Burnap, Charith Perera, **ICS-ViZ: Integrity Checking at the Edge: Visualising cyber Attacks towards enhanced Decision-Making Experience** (*Work-in-progress*)
  VIDEO

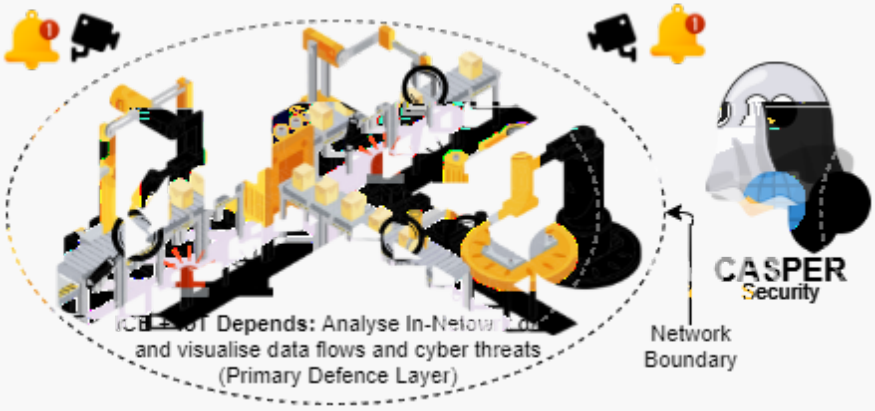
RESILIENTSENSING.AI

Context-Aware Security for Industrial Cyber-Physical Edge Resources

Researcher: Hakan Kayan (PhD Student 2020-2024)

Industrial cyber-physical systems (ICPSs) manage critical infrastructures by controlling the processes based on the "physics" data gathered by edge sensor networks. Recent innovations in ubiquitous computing and communication technologies have prompted the rapid integration of highly interconnected systems to ICPSs. Hence, the "security by obscurity" principle provided by air-gapping is no longer followed. As the interconnectivity in ICPSs increases, so does the attack surface. Industrial vulnerability assessment reports have shown that a variety of new vulnerabilities have occurred due to this transition, leading to an increase in the targeting of ICPSs. Key findings from Verizon's 2020 data breach report show that 381 data breaches (10% of the total) are against industrial systems, not all target OT equipment.

We aim to develop a context-aware anomaly detection mechanism/model that physically observes ICPS edge devices to detect cyberattacks. The followings are the main objectives of the project:



- Review the current ICPSs from a cybersecurity perspective.
- Develop end-to-end reconfigurable IoT sensing infrastructure for training and deploying analytics at scale.
- Augment cyberattack detection through physical, behavioural monitoring in ICPSs.
- Evaluate the efficiency of a Context-aware, Dynamically Adaptive IoT Edge Network for Cyber Attack Detection in Industrial Control Systems (CASPER) through experimental evaluations.

Partners and Relevant Projects

PETRAS GCHQ EXALENS®

Outcomes

- **[Journal]** Hakan Kayan, Yasar Majib, Wael Alsafery, Mahmoud Barhamgi, Charith Perera, **AnoML-IoT: An End-To-End Reconfigurable Multi-Protocol Anomaly Detection Pipeline for Internet of Things**, Elsevier Internet of Things (Elsevier IOT), Volume 16, 100437, December 2021 PDF BIB SOURCE CODE CODE CODE
- **[Journal]** Hakan Kayan, Matthew Nunes, Omer Rana, Pete Burnap, Charith Perera, **Cybersecurity of Industrial Cyber-Physical Systems: A Review**, ACM Computing Surveys (ACM CSUR), Volume 54, Issue 11s, January 2022, Article No.: 229, pp 1–35 PDF BIB SOURCE
- **[Demo]** Hakan Kayan, Omer Rana, Pete Burnap, Charith Perera, **CASPER: Context-Aware Anomaly Detection System for Industrial Robotic Arms** (*Work-in-progress*) VIDEO

Context-Aware Security for Smart Homes using Cyber-Physical Behavioural Data Analysis

Researcher: Yasar Majib (PhD Student-2020-2024)

The rapid growth of the Internet of Things (IoT) requires a deep look into security and privacy challenges. This growth is changing our contemporary world, which is now connected in novel ways and poses new challenges. Nowadays, these tiny little devices (IoTs) routinely communicate with each other on behalf of humans. As we move into this AI era, the world needs assurance that this fabric of interconnected things is not vulnerable to traditional or cyber-physical security threats. The entire spectrum of IoT fabric includes; devices/things, connectivity, storage, and applications – all of which are potentially vulnerable. In addition to traditional connectivity channels, IoTs are exposed to physical channels such as temperature, humidity, air quality, illumination, sound, and many more. A single vulnerable IoT can be a gateway to break into a secure smart home system by exploiting a cyber vulnerability or a physical channel(s).

Currently, available solutions are mostly focused on traditional Network Traffic Analysis (NTA) for detecting anomalies in cyber systems (Intrusion Detection or Intrusion Prevention), which is insufficient in the IoT scenario.

This project focuses on cyber-physical behaviour, where we aim to detect cyber attacks by detecting anomalies in smart homes through cyber-physical behavioural data analysis. We aim to develop low-cost multi-purpose sensor nodes which can detect anomalies in a smart home by analysing cyber-physical data. In another scenario, imagine a malicious party switch on a toaster at midnight while spoofing the smart plug and preventing it from reporting to the smart home hub. The multi-purpose sensor network can detect such anomaly events by physically observing temperature, vibration, light, or sound even though the malicious party may have compromised the smart plug and the smart home hub preventing it from generating NTA-based anomaly. The project has the following objectives:

- Review the existing cyber-physical anomaly detection techniques in smart homes.
- Learn typical smart home behavioural patterns using a distributed multi-purpose sensors network by understanding the signatures (e.g., which devices are typically activated in sequence due to household behaviours) of smart home devices.
- Detect anomalies by observing behavioural patterns by combining and correlating network traffic analysis and observational data from an independent IoT sensor network.

Partners and Relevant Projects

PETRAS bre GCHQ

Outcomes

- **[Journal]** Yasar Majib, Mahmoud Barhamgi, Behzad Momahed Heravi, Sharadha Kariyawasam, Charith Perera Detecting Anomalies within Smart Buildings using Do-It-Yourself Internet of Things, Journal of Ambient Intelligence and Humanized Computing (JAIHC), September 2022
  PDF BIB DATA SET SOURCE CODE

Video Analytics towards Anomaly Detection on Edge for Smart Cities

Researcher: Yaser Abu Awwad (MPhil Student-2021-2023)

Cameras are widely used in the smart city domain to monitor and supervise environments such as road traffic, office buildings, smart homes, etc. However, most commercial (off-the-shelf) camera systems can only detect a few sets of predefined objects (e.g., persons and vehicles) and behaviours. Most of these camera systems are designed for streaming video to the cloud. In a limited number of systems, cameras may do minor edge processing tasks such as detecting people and vehicles. Such primitive capabilities are insufficient to facilitate the more complex use cases below. Further, sending video streams to the cloud without processing may not be useful and require significant network bandwidth, especially when the systems need to be scaled for thousands of cameras. Further, not all video frames are worth processing in-depth.

Farms in Monmouthshire want to prevent/detect crime and safeguard lone workers. The objective is to prevent thefts of machinery and livestock and monitor farmers to ensure their safety, particularly whilst working alone at remote locations on the farm. Raglan Castle in Monmouthshire wants to detect vandalism and ensure children's safety by monitoring any children climbing walls or performing any dangerous activities so that the local staff can intervene in a timely manner. Blaenau Gwent wants to monitor their car parks to understand how they are being used and how to better incentivise public transport (e.g., monitor how many people get off from a vehicle). Another important aspect is to detect anti-social behaviour using bus stop cameras. All the use cases require some level of anomaly detection capabilities beyond what off-the-shelf systems can provide.

This project combines pre-trained object detection and computer vision models to detect complex anomaly behaviours using cameras. Each pre-trained model plays a crucial role in a particular scene, extracting information and actions to be incorporated to detect different types of anomalies. Moreover, this project is not focused on processing a full video in real-time. It aims to pick up signals of potential anomalies through lightweight edge processing (e.g., a farm animal moving towards an unusual area). Once the signals are detected, systems will conduct an in-depth analysis using their full capabilities by feeding the selected frames into several different pre-trained computer vision models. The objective of this project is as follows:

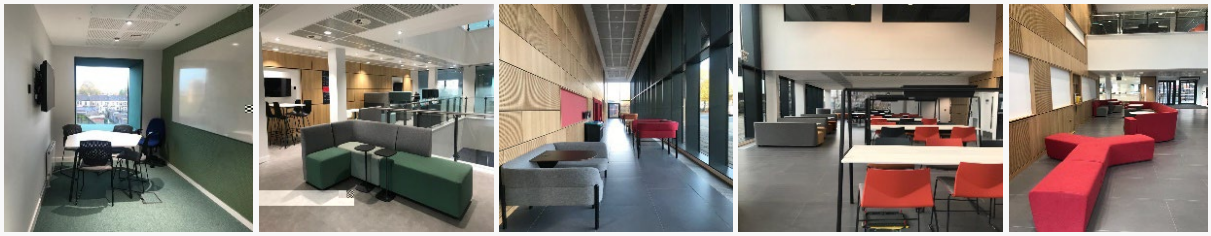
- Conduct a literature review on anomaly detection from camera feeds image/video and explore the types of anomalies.
- Measure trade-offs of processing videos fully and partially on edge and in the cloud.
- Explore the factors that impact anomaly detection performance in terms of environment and in-the-wild challenges, such as lighting, angle, camera resolution and other factors.
- Develop a system that can automatically adapt and decide which anomalies should be monitored on edge to pick the early signals related to a given use case (i.e., deployment) to improve overall performance and accuracy.
- Develop a technique to detect anomalies using off-the-shelf cameras by applying existing deep learning and computer vision techniques.

Partners and Relevant Projects

Sensing as a Service within Buildings Towards Data-Driven Collaborative Service Design

Researcher: Wael Alsafery (PhD Student-2021-2025)

University buildings are unique, given they partly act as office buildings and partly as semi-public buildings. They may use the building for a variety of different activities such as individual work, group work, meetings, socialising, and so on. Modern university buildings are built having these requirements in mind at the design stage by incorporating different types of spaces to facilitate these students and their needs. However, there aren't any follow-ups being carried out to measure these spaces' actual utilisation after the building is commissioned and handed over to the University.



This project is conducted within the Abacws, the newly built home for Computer Science and Mathematics students at Cardiff University. It contains a number of different types of spaces that are dedicated to facilitating taught students, research students, and staff members. Not all modern buildings, including Abacws, are augmented with sensors due to additional costs and lack of perceived value and understanding. First, we aim to develop, deploy, and understand which IoT technology is best suited to measure occupant behaviour and usage patterns related to different types of study spaces. Secondly, we aim to understand how to utilise IoT technology to facilitate occupants and the building service design team to communicate better and make collaborative and informed decisions using data-driven approaches to study space utilisation. Even though this project primarily focuses on Abacws, the technology we develop could apply to any building with similar characteristics and requirements (e.g., multi-purpose heterogenous open spaces to facilitate temporary occupants). This project is composed of several objectives:

- Conduct a literature review on how sensor technologies are used within indoor environments to monitor occupancy and usage of spaces to improve service delivery.
- Design, develop and deploy IoT sensing technologies to monitor a variety of heterogeneous study spaces and investigate which technologies work best for each space by measuring their performances and trade-offs.
- Develop data-driven approaches to facilitate/mediate informed communication between occupants and the building's service design team towards improving the overall quality of service

Partners and Relevant Projects

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Outcomes

- **[Technical Report]** Wael Alsafery, Omer Rana, Charith Perera, **Sensing within Smart Buildings: Survey,** Technical Report, 2022 PDF BIB

Self-Configuring Internet of Things Architecture for Context-Aware Anomaly Detection

Researcher: Abdulaziz Aljohani (PhD Student-2021-2025)

Anomaly detection is identifying unexpected items or events in data sets that differ from the norm. It is a well-investigated area within research communities; however, anomaly detection using IoT sensor data is comparatively unexplored. In order to develop IoT sensor-based anomaly detection solutions, engineers require significant technical knowledge (e.g., which algorithms to use, how to set parameters, etc.) and domain knowledge (e.g., agriculture, built environments, usual patterns within a given context, etc.). Recently, some commercial solutions (e.g., Microsoft Anomaly Detector) are being developed to simplify the development process by allowing engineers to use black-boxed anomaly detection algorithms with few configurable parameters (i.e., sensitivity, max window size, max anomaly ratio).

We believe that much more complicated contributing factors need to be considered when deploying anomaly detection systems. Further, even though we may know some of the contributing factors during design time, we may not know how to configure a system until we deploy the anomy detection system in a given context. For example, IoT devices have limited resources (e.g., energy, memory, computing resources) and may have shared responsibilities (i.e., not dedicated to anomaly detection). As a result, which devices would be available to perform anomaly detection may not be known beforehand. Further, the heterogeneity of IoT application scenarios makes it infeasible to find one generalised anomaly detection technique that works for every possible IoT architecture.
Additionally, competing requirements such as privacy vs performance could need to be managed. We believe that the best way to handle these challenges is to develop a self-configurable anomaly detection system configuring the above-mentioned parameters at runtime and adapting to the given context. In this project, we propose FedBio-IoT, a federated self-configuring IoT architecture for context-aware anomaly detection. FedBio-IoT is based on nature-inspired algorithms that use the concept of evolutionary algorithms and swarm intelligence to monitor, configure, adapt, and change the federated IoT architecture according to the population's behaviour and biological evolution from one generation to the next. We aim to investigate how to reduce the technical and domain expertise engineers require and reduce the trial-and-error guesswork required during the development stage. This project is composed of several objectives:

- Conduct a literature review on anomaly-detection techniques, their characteristics and configurable properties.
- Study the capabilities of a wide range of swarm-intelligence algorithms that can be used in self-configuring IoT architecture and examine their strengths and weaknesses.
- Evaluate the performance of self-configuring IoT architecture for context-aware anomaly detection based on swarm intelligence through experimental evaluations in different IoT application scenarios.

Partners and Relevant Projects

PETRAS GCHQ

Explore the Role of Tiny Cameras Towards Augmenting Anomaly Detection within Built Environments

Researcher: Norah Albazzai (PhD Student-2021-2025)

In the cyber world, anomalies are detected by analysing network packets. However, the cyber-physical world requires a different approach to monitor both network and physical worlds. An anomaly is an observation that does not conform to a normal pattern. Anomalies within built environments include intrusion, fire, variation in power consumption, unusual activation of smart devices, abnormal living patterns and so on. Traditional physical anomaly detection systems (e.g. temperature sensor monitoring afire through temperature variations) use simple sensors (temperature, humidity, vibration, motion). For example, an open window has been detected using a temperature sensor. However, as the complexity of the anomalies increases, the achieved results become less accurate. In addition, traditional sensors can be affected by noises produced by the surrounding environment. Another limitation of traditional sensors is that they can only detect measurable properties, and simple sensors cannot detect some parameters. Cameras are an advanced type of sensor that has been used mainly in surveillance tasks. Historically, in anomaly detection, the utilisation of camera sensors is limited due to multiple factors such as increased costs, comparatively larger, and privacy issues. However, tiny cameras are becoming cheaper and less than 1 inch in length.

This project investigates how to augment sensor-based anomaly detection systems with tiny cameras in a privacy-aware manner. For example, to reduce privacy invasion, camera sensors will only be activated to observe a scene if another sensor (e.g. temperature, motion) produces an abnormal result. Further, we believe tiny cameras can be used to train other sensors over time to improve their anomaly detection capabilities and reduce the involvement of tiny cameras in decision-making, therefore reducing privacy concerns. This project use pre-trained object detection and computer vision models to detect anomalies and correlate them with other sensor data to improve the overall performance of the anomaly detection system. The project has the following main objectives:

- Conduct a literature review on camera systems to explore the role of the camera as a sensor in the context of anomaly detection in built environments.
- Investigate how integrating sensor-based anomaly detection with low-cost cameras can affect the overall performance.
- Identify the capabilities and limitations of the tiny camera sensor and the deployment challenges and investigate how the tiny camera can be used to (re)train other sensors over time and enhance their performance.

Partners and Relevant Projects

Outcomes

- **[Technical Report]** Norah Albazzai, Omer Rana, Charith Perera, **Camera as a Sensor Towards Augmenting Anomaly Detection in Internet of Things Systems: A Survey,** Technical Report, 2022 PDF BIB

Context-Aware Knowledge-Driven Cyber-Physical Security at the Edge for Smart Homes

Researcher: Azhar Alsufyani (PhD Student-2021-2025)

Smart devices are heterogeneous, and each has a different set of capabilities in sensing and actuation. To unlock the true potential of self-adaptive smart spaces, these devices should work and collaborate by sharing their capabilities to achieve a given goal. These smart IoT devices should evolve automatically, depending on users' needs, and adapt to new contexts/conditions. While smart spaces are advantageous and desirable in many ways, they may be hacked, exposing privacy and security, or rendering the entire area a hostile environment where ordinary tasks are impossible. Therefore, securing smart spaces can be challenging due to device heterogeneity, continuous changes in context, and limited device resources.

This project aims to develop techniques that can dynamically configure a given smart space (i.e., self-adapting) to achieve a goal (i.e., ensuring security and safety of the cyber-physical system) without needing of cloud services (i.e., edge computing). To achieve this, we adopt Monitor-Analyze-Plan-Execute-Knowledge (MAPE-k) method. Some of the investigations we need to carry out are as follows. First, we need to capture information that MAPE-k requires. Some key pieces of static information are smart device capabilities and limitations. For example, devices such as smart vacuum cleaners can move. Another example is webcams which have the capability of taking images. Other important information needs to be continuously updated (e.g., device locations, weather, environmental conditions, calendar information). Some updates could be simple as downloading a calendar, whereas others require data analytics (e.g., detecting a window open by analysing temperature variations near the window). We expect this knowledge base to be modelled around well-known ontologies (e.g., W3C SSN, W3C BOT). Next, we aim to assess and select open-source frameworks that can analyse a given context and plan the right course of action to achieve the given goal. We aim to combine rule-based systems, e.g., Drools/OpenHAB-Rules and AI planning techniques, e.g., Optaplanner, to implement parts of MAPE-k. Currently, smart home security solutions focus on network traffic analysis to detect cyber-physical threats using ML/DL techniques. This project aims to demonstrate knowledge-based systems' utility in smart home security.

- Conduct a literature review on knowledge-based techniques that are being developed and deployed within the smart home domain with a special focus on cyber-physical security
- Develop a knowledge model to capture all the relevant information required by Monitor-Analyze-Plan-Execute-Knowledge (MAPE-k) loop to enable self-adaptive cyber-physical security.
- Investigate, select and implement the best techniques for each phase within MAPE-k while utilising open-source APIs/frameworks as much as possible.
- Measure the trade-offs of competing techniques and make recommendations for their use
- Develop a series of demonstrators to showcase how knowledge-based self-adaptive systems work in the wild in the context of smart homes.

Partners and Relevant Projects

Talking Buildings: Making Buildings Talk using Adaptable Data Analytics

Researcher: Suhas Devmane (PhD Student-2021-2025)

Modern smart buildings are equipped with IoT sensors to facilitate efficient and effective maintenance of buildings. These IoT sensors can be used to measure quite valuable aspects of buildings such as structural health, occupant behaviours, occupant health, and many more towards increasing functionality, comfort, safety, and reducing running costs. Even though much academic work has been done to generate these insights from sensor data, deploying them in the real world is quite challenging due to the simplistic assumption made within academic work. A more viable option is to buy very expensive off-the-shelf solutions from companies specialising in Buildings Management Systems (BMS) or Buildings AI solutions providers. The downside is that these solutions are often highly restrictive in terms of capabilities, extendability and adaptability. For example, we will be required to deploy their sensors exactly as prescribed and require a lot of manual labour to adapt them to new building types and layouts. Further, most of these BMS and AI solutions are designed to be used by domain experts.

In this project, we aim to address two key issues highlighted above. First, we will investigate how we could develop a semantic interoperability layer between IoT sensors and data analytics so the analytics could be adaptable for a given building's configuration and layout. We aim to embed the domain knowledge into the system we are building so non-domain experts can use the system to understand better how the buildings are performing. To make the system more accessible, we aim to utilise conversational AI techniques to mediate the communication between the building and the non-experts. By doing this, we aim to give a voice to the buildings so they can communicate with humans in natural language and express how it feels. We envision a future in which the buildings can answer performance-related questions (e.g., Building Research Establishment Environmental Assessment Method (BREEAM)) with the help of IoT sensors. This project has the following objectives:

- Conduct a literature review on the relationship between useful insights and what data types and analytics

Developing an Evaluation Framework for Anomaly Detection within Built Environments

Researcher: Mohammed Alosaimi (PhD Student-2021-2025)

Smart Built Environments are composed of physical and digital infrastructure and aim to improve data-driven decision-making and provide faster and cheaper operation and maintenance (e.g., better whole-life value). They are increasingly more vulnerable to cyber-physical attacks. Anomaly detection techniques are traditionally used to detect any abnormal behaviours. Anomaly detection is a broad field with a rich history where many different techniques have been developed. Out of those, a subset of techniques is focused on real-time anomaly detection. Another subset of techniques focuses on sensor data based on real-time anomaly detection. A key challenge of anomaly detection in the context of built environments is that they are heterogeneous in nature and produced by different sensing devices in an unorderly fashion. Some data types are sensor values (e.g. temperature 23C). Other data types could be status or commands (e.g., ON/OFF, 0-1). Some data types could be energy consumption. There are also encrypted data where the actual content is unknown but the metadata available (e.g., packet destination, packet size, frequency of communication). Developing anomaly detection techniques within such a context requires comprehensive testbeds (or at least datasets collected from a comprehensive testbed). However, no significant emphasis has been put on developing testbeds that can be used to develop, evaluate and compare anomaly detection techniques.

Developing a testbed has always been treated as a secondary task, as the development of anomaly detection takes priority. The impact of a testbed's characteristics and properties towards the anomaly detection techniques developed using them is largely unknown and less studied. The fundamental problem with generating synthetic environments is that in order to be realistic, a large amount of data must be generated in order to provide a convincing pattern of life for the simulated network, as well as give the appearance of longevity (the network must not appear to have been recently generated). Further, anomaly detection techniques are challenging to evaluate, especially when developed using different testbeds and conditions. This project aims to develop a comprehensive framework to evaluate the capabilities of a given anomaly detection technique. The project objectives are:

- Conduct a literature review to determine how testbeds are built to evaluate IoT-based anomaly detection techniques.
- Identify characteristics and properties of smart home testbeds that impact the quality of the anomaly detection techniques developed using them.
- Develop techniques to capture, annotate and model data from smart home testbeds to support comparable anomaly detection techniques development.
- Develop techniques to generate realistic synthetic datasets compared to real-time live anomaly detection and measure the trade-offs of both approaches.

Partners and Relevant Projects

Integrity Checking at the Edge (ICE) for Operational Decision Support (ICE-ODS)

Team: Matthew Nunes, Pete Burnap, Charith Perera, Neetesh Saxena (2021-2022)

This project will integrate the outcomes of the PETRAS-funded "Integrity Checking at the Edge (ICE)" project into a prototype operational decision support mechanism at Thales UK. Thales offers an end-to-end Autonomous Logistics Supply that combines an intuitive digital twin interface, unmanned command and control system, and an autonomous, all-terrain Unmanned Ground Vehicle (UGV) system with its own networked communications systems. UGVs are vital for supporting humanitarian rescue and relief efforts in unsafe environments – for example, in natural disaster regions or conflict settings.

However, UGVs and their communication systems are vulnerable to cyber-attacks which could disrupt such missions. Most existing cyber attack detection systems only flag the presence of an attack. The current ICE project is working with Thales to understand when and how an attack is occurring and support making decisions such as when and how to act to ensure the continuity of the mission. This is vital as such missions are not as simple as stopping communication or turning the devices off during an attack – everything must keep operating, and human interaction in the most strategic and minimal-impact way is key. A key outcome of the current ICE project is a visual and interactive method to "dig into" data collected from edge networks – to flag anomalies and potential cyber-attacks – and to enable security operations analysts to collaborate with business continuity teams in a way that enables cyberattacks to be responded to while taking into account the impact of any incident response decisions on the wider system. This is unusual, especially in operational technology settings where the focus is on the safety of the system. The aims of this project are:

- To co-create a practical and impactful toolkit with Thales that translates the outcomes of the ICE project into a living, breathing context that has major life-critical safety and cybersecurity implications. Thales supports integration through access to real UGV testbeds and leading workshops with CU to capture insider knowledge on possible responses to cyber threats.
- To validate the outcomes of the ICE project in a real-world setting and gain end-user feedback in a real-world scenario.
- To create a permanent showcase demonstrator of the outcomes of PETRAS-funded research at a major private sector R&I investment location – Thales Ebbw Vale.
- To identify and mitigate vulnerabilities to cyber-attacks within the showcase demonstrator, thereby ensuring that the demonstrator represents a benchmark testbed for cyber risk visualisation and assessment.

Partners and Relevant Projects

THALES PETRAS

Outcomes

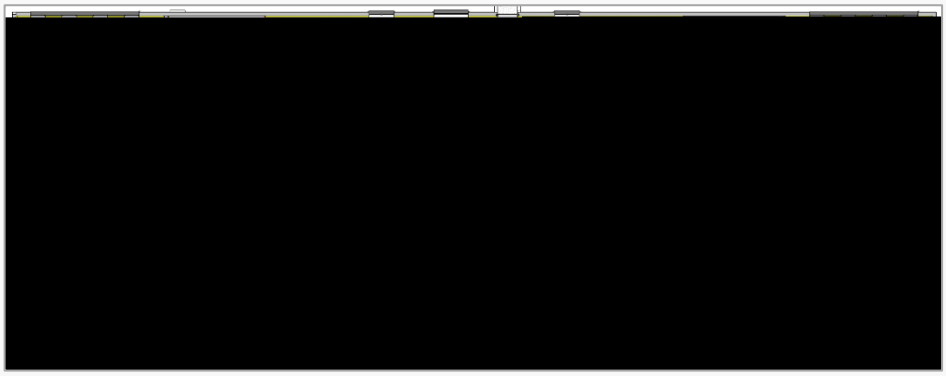
- [Demo] Matthew Nunes, Pete Burnap, Charith Perera, Jason Dykes, **Exploiting User-Centred Design to SecureIndustrial Control Systems** (*Work-in-progress*) VIDEO

Low-Cost Adaptive Mobile Sensing within Buildings towards Augmenting Smart Buildings

Researcher: Siyuan Li (Cecilia) (PhD Student-2022-2026)

Nowadays, people spend most of their time in indoor environments. The indoor environment has become a key factor in people's health and productivity. At the same time, smart building systems have emerged and are playing a role in building management to improve the health and productivity of occupants. Sensors are placed in the building to collect data to generate insights. Smart building systems use AI to understand the occupants' habits from the data and use actuators to improve occupant experience and overall building efficiency. There are several sub-systems in the smart building system, for example, the air quality system, the lighting system, the thermal comfort system, and the HVAC system (heating, ventilation, and air conditioning). Depending on the size and use of the building, the number of different sub-systems and sensors varies. The number of sensors and the way they are placed becomes a challenge. A large number of sensors need to be deployed throughout the building to generate useful results and insights. Such deployments, especially if retrofitting, would lead to increased deployment costs/effort/time and maintenance costs.

This project aims to design a smart building system for an existing building that reduces the complexity and consumption of the system without compromising accuracy and functionality by using mobile sensors instead of static sensors. The system has four sub-systems, an air quality system, a lighting system, a thermal system, and an activity recognition system. There is an additional control centre to integrate the various systems and make them interoperable. The activity recognition system is also the only subsystem that collects human data. It will detect the number of people in the design area and send it to the control centre. Except for the activity recognition system, all other systems collect environmental data. The main objectives of this project are:



- Conduct a literature review of autonomous/mobile systems used in smart building systems and smart environments and classify and compare their utility, characteristics, and capabilities.
- Design a mobile sensor system based on IoT technology to reduce the number of sensors used in a smart building system and try to achieve equal or (close enough results).
- Evaluate the performance of the mobile sensors system and examine the trade-offs between static/stationary sensor systems and hybrid systems in the context of various anomaly detection tasks (e.g., violation of sustainability standards, comfort and wellbeing preferences, etc.)

Partners and Relevant Projects

bre

Resilient Build Environments (CASPER Shield)

Team: Charith Perera, Hakan Kayan, Yasar Majib (2022-2023)

Motivation and Business Need: NCC Group and the Global Cyber Alliance recorded more than 12,000 attacks to maliciously log into smart home devices. Recent statistics show that over 200 million smart homes can be subjected to these attacks. Conventional security systems are either focused on network traffic monitoring (e.g., firewalls) or physical environment monitoring (e.g., CCTV or sensors), but not both. These systems fail to detect sophisticated attacks/intrusions (e.g., advanced persistent threats, zero-day) that can cause physically behavioural changes (e.g., an increase in room temperature due to hacked smart air conditioner). Thus, there is a need for an advanced Cyber-Physical security system for homes that can detect those abnormalities. You can think of our product as 'anti-virus software for smart homes' that protect you and your connected things from cyber criminals and physical intrusion, making the connected living spaces more secure and safer.

CASPER Shield

Technical Challenge: To secure smart homes, we promise a Cyber-Physical anomaly detection system that uses AI/ML technology to identify the 'normal' of the home and detect the 'abnormal'. We look for abnormal occurrences in a smart home using cyber and physical data rather than relying only on insecurely transmitted, manipulable cyber/network data. Our techniques will discover real-time anomalies based on the behaviours of devices and occupants. A key technical challenge is to figure out how to integrate cyber and physical data best to detect anomalies in smart homes.

Market Opportunity and Competition: The smart home market is worth £87.61 billion, expected to grow to £136.34 billion in 5 years. The smart home security market is £0.69 billion and is expected to reach £1.08 billion in 2026. The smart home security market could be our target market. Initially, we will focus on the independent living and remote healthcare market, where smart devices are increasingly being adopted. This project is part of the CyberASAP programme.

The Cyber Security Academic Startup Accelerator Programme is a 1-year programme (FEB-2022-2023) separated into three stages; only the best teams progress to the next stage after an evaluation by a panel of senior academics, investors, industry experts and startup advisors. CyberASAP (funded by DCMS) is a national competition to identify the most promising commercial opportunities.

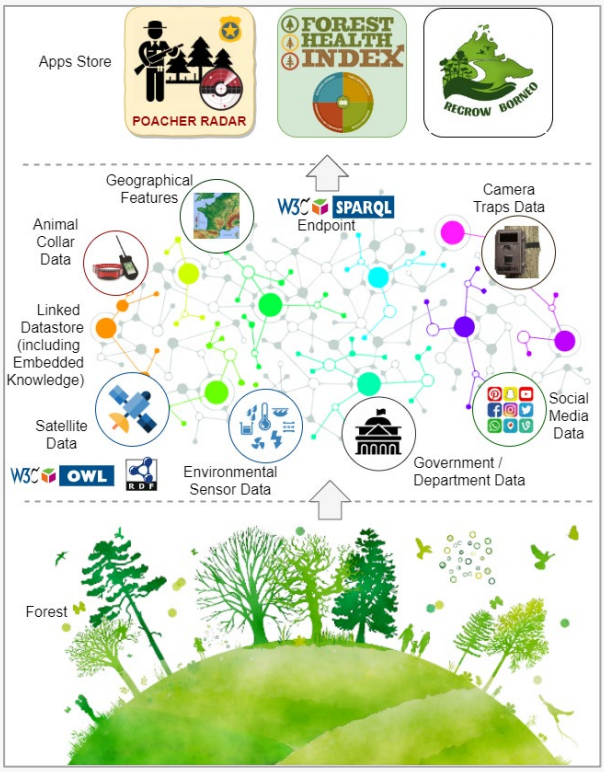
CYBER + ASAP | Academic Startup Accelerator Programme

DATA OBSERVATORIES

Semantic Data Integration Towards Forest Observatory-based App Ecosystem

Researcher: Naeima Hamed (PhD Student-2020-2024)

Poaching and animal trafficking are significant challenges around the world. Anti-poaching efforts are always underfunded and under-resourced. Law enforcement officers cannot keep up with the large number of poachers trying to kill and capture animals. Due to limited manpower, they cannot patrol and protect vast areas of land. We will semantically integrate data gathered by Bio-science researchers and environmental scientists to predict where the poaching activities will occur in the future. Our data-driven prediction models will tell areas and time frames that are highly likely to have poaching incidents. Therefore, law enforcement agencies can deploy their limited resources into those areas. This project will focus on the Lower Kinabatangan Wildlife Sanctuary, Sabah, Malaysia. This project collaborates between the School of Computer Science and the School of Biosciences (and its Danau Girang Field Centre; DGFC) at Cardiff University.



Our approach is to develop a Forest Observatory and develop data-driven predictive analytics to predict poaching incidents. Forest Observatory is a Linked Datastore that integrates heterogeneous data. Collecting data in forests is much more challenging than in cities due to the lack of infrastructure. However, while we expect to deploy an Internet of Things (IoT) infrastructure to enable poaching monitoring, we aim to utilise already collected data sets to develop predictive poaching models. For example, DGFC has data sets collected by researchers for wildlife species monitoring over the last decade, such as animal collar data, camera traps, satellite imagery, LiDAR and environmental data, with each data set generated using different time frames, durations, geographic areas etc.

Partners and Relevant Projects

Outcomes

- **[Technical Report]** Naeima Hamed, Omer Rana, Pablo Orozco-terWengel, Benoît Goossens, Charith Perera, **Open Data Observatories**, Technical Report, 2021 PDF
- **[Technical Report]** Naeima Hamed, Omer Rana, Pablo Orozco-terWengel, Benoît Goossens, Charith Perera, **Forest Observatory: A Resource Of Integrated Wildlife Data**, Technical Report, 2022 PDF VIDEO RESOURCES

DATA OBSERVATORIES

Dynamically Orchestrate-able Low Power Internet of Things Infrastructure for Sustainable Wildlife Conservation

Researcher: Mark Butterworth (PhD Student-2020-2026) [PT]

This project aims to develop a reliable communications technique to monitor animal traps remotely. Low power digital transmission techniques encounter many hurdles when operating in harsh/dense jungle environments. Traditionally the problem can be overcome using higher power transmissions; however, in this case, it is not possible as devices need to operate for long periods autonomously and cannot afford the increased burden of regular battery changes. This research project examines frequencies and develops protocols that allow secure, reliable communication across dense jungle environments using low-power digital transmission protocols.

The research aims to deliver a fully functional concept demonstrator based upon communications theory; the key objective is to be able to monitor sensing infrastructure in the Kinabatangan wildlife sanctuary without the need to visit each sensor.




Trap activation detection – Most traps operate using weight-based or bait based activation triggers. Smaller animals could accidentally become ensnared, meaning cages must be visited regularly to ensure animal safety. Any sensor monitoring system must be reliable and fail-safe to ensure wildlife welfare.

Poacher tracking – While poachers and vehicles' accurate pursuit is not practical without deployed sensors on the person or vehicle, it would be possible to monitor poachers' activities and movements. Sensors could monitor people passing through pinch points and congregating at meeting points. The data from these sensors could provide information to other data science projects to help elicit information on poacher behaviour and help predict everyday activities.

Poacher detection – Vehicles are not allowed in the sanctuary after 19:00, so sensors deployed to detect these vehicles could use vibration sensors, Automatic Number Plate Recognition (ANPR), or sound as it is reasonable to assume that vehicles in the wildlife reserve after 19:00 are unauthorised.

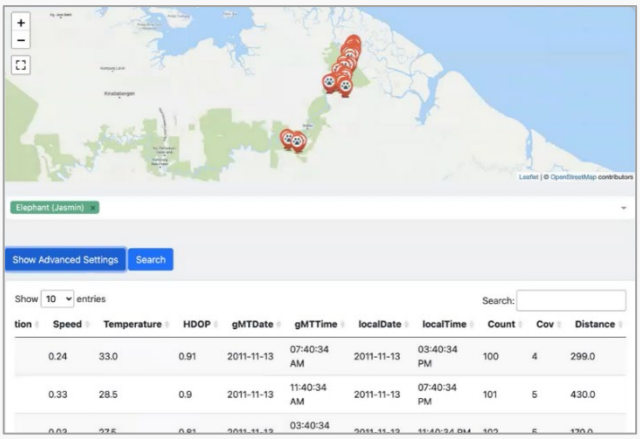
Remote camera trap battery monitoring – Messages for monitoring battery life can be tiny and not time-critical. Message updates can be provided on a predetermined cycle, such as hourly or daily. This tradecraft would reduce the number of messages sent and enhance battery life. User-definable heartbeats would allow users to define a refresh timeframe with which they are comfortable.

Partners and Relevant Projects

Making Linked Data Accessible through End-User Development for Bioscience Researchers in the Context of Micro Observatories

Researcher: Omar Mussa (PhD Student-2021-2025)

Linked Data is a set of design principles to structuring data in an interconnected system to make them accessible and machine-readable. When the data gets linked, it becomes traversable, and nodes will be linked through relationships. Linked Data breaks down the information silos that exist between various formats and brings down the fences between various sources. It facilitates the extension of the data models and allows easy updates. As a result, data integration and browsing through complex data become easier and more efficient. In addition, Linked-Data follows a specific schema that makes it easily understood by machines and humans alike. Unfortunately, even though the data is human-readable, it is challenging for non-expert users to retrieve it because Linked-Data will need a good understanding of Semantic queries. Learning Semantic query (i.e., SPARQL Query Language) is not easy for non-expert users, and end-users will unlikely use it.



This project makes the Linked-Data more accessible and allows the non-technical end-user (e.g., Bioscience Researchers, and wildlife conservationists) to perform their job more efficiently through developing novel interfaces. More specifically, we aim to combine GUIs with conversational AI techniques to facilitate efficient and effective linked data retrieval for non-technical users. The naive user will not need to have any experience using SPARQL or any other query language to retrieve the data. Besides, expert users will perform their job easier in less time. This project composes three main objectives:

- Review existing end-user development techniques on how to make Linked-Data accessible.
- Design and develop a hybrid interface that combines traditional WebGUI with a conversational chatbot to allow non-technical users to efficiently and effectively express data requirements.
- Enhance the user experience by developing novel data visualisations techniques and context-aware predictions that enable end-users to explore data more efficiently and effectively.

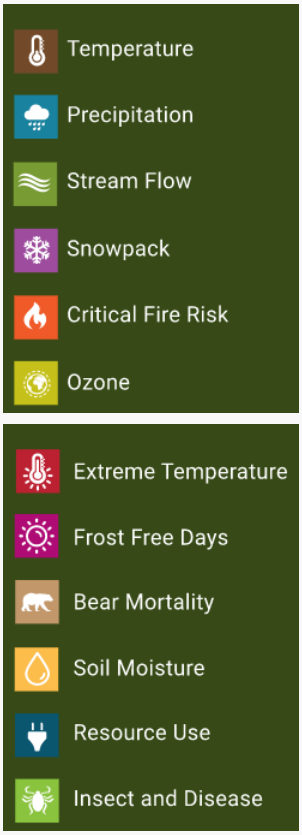
Partners and Relevant Projects

Outcomes

- **[Technical Report]** Omer Mussa, Omer Rana, Pablo Orozco-terWengel, Benoît Goossens, Charith Perera, **Making Linked-Data Accessible: A Review**, Technical Report, 2022 PDF
- **[Demo]** Omar Mussa, Omer Rana, Benoît Goossens, Pablo Orozco-terWengel and Charith Perera, **ForestQB: An Adaptive Query Builder to Support Wildlife Research**, In Proceedings of the 12th International Semantic Web Conference (Posters & Demonstrations Track), Hangzhou, China, October 23-27, 2022 PDF BIB SOURCE VIDEO

Integrate a Heterogeneous set of Data towards Developing a Forest Health Index

Researcher: Rayan Binlajdam (PhD Student-2022-2026)



Internet of Things technologies has been used in the context of Wildlife Conservation in various ways. One such area of interest is forest health. Forest health has been defined from many different perspectives. Forest health can be defined by the production of forest conditions which directly satisfy human needs and by resilience, recurrence, persistence, and biophysical processes which lead to sustainable ecological conditions. Forest health is a qualitative term that refers to the general condition of a forest. A healthy forest is relatively free of insect infestations, diseases, exotic weeds, and air pollution. A healthy forest can resist damage from catastrophic events like acute insect and disease attacks, fire, wind, and flooding and fully recover from these perturbations to continue its life history functions over decades, centuries, or millennia. However, there aren't any well-accepted methodologies to measure forest health. Some frameworks highlight (foresthealthindex.org) so that need to consider when measuring forest health. These are high-level frameworks without any specific formula to operationalize and measure a given forest area.

IoT sensors are useful due to their ability to accurately measure environmental parameters on the ground level. However, deploying large-scale IoT networks in forest environments are quite challenging as sensors can only monitor a small geographical area. Therefore, the only way to scale up is to deploy a large number of sensor nodes, which is difficult due to hardware costs, deployment costs, difficulty in providing energy and developing network communication in harsh environments with high humidity and a lot of physical obstacles and so on.

Another alternative mechanism is to use drones to observe a forest area. However, drones have limitations where they only see first from the top and cannot directly sense what is happening deep inside the jungles at the soil level. If we were to use drones, we would need to use proxy measurements to determine what is happening on the ground by looking at the quality of the trees and other characteristics that can be remotely measured through cameras. In this project, our objective is to combine both IoT and drone imaging data. Drones bring the scalability aspect, and IoT technologies bring the accuracy aspect. To combine, we will deploy both IoT and drones in a test forest environment to train a model capable of using drone images to predict the outcomes of IoT sensors and, subsequently, the forest health index. We might also use satellite images to complement drone imaging and IoT sensor data. Our ultimate objective is to develop an AI that could use drone images to produce a forest health index without needing IoT hardware sensors deployed at scale, perhaps only with a limited number of IoT sensors to capture the ground truth and calibration purposes.

Partners and Relevant Projects

Scalable Circular Supply Chains for the Built Environment

Team: Yingli Wang (PI), Jon Gosling, Omer Rana, Pete Burnap, Charith Perera, Yacine Rezui, Qian Li, Rajiv Ranjan, Aad van Moorsel, Graham Morgan, Ellis Solaiman (2021-2024)

The outcome of this multi-disciplinary industry/academic co-development effort will be to create a scalable, decentralised blockchain environment to enable tracking of reusable materials, parts/components or services to support a circular supply chain for the Architecture, Engineering and Construction (AEC) sector. The academic team in the project will create a digital (software) platform (supporting supply chain tracking and data analytics) to facilitate 5 "R" features, which are: (i) Reuse and Redistribute (ii) Refurbish and Remanufacture; and (iii) Recycle. The outcome will be validated with sector-leaders in AEC, such as HS2, Arup, Celsa Steel and with an SME (SeroHomes). The key transformational contribution of this project is an establishment and assessment of a highly connected circular supply chain that contributes to radical whole lifecycle decarbonisation and waste reduction within an AEC project. We believe the digital (software) platform will also have the potential for commercialisation and possible integration with systems from other AEC suppliers – such as the "Pathway to Zero" tool from SeroHomes (to achieve zero carbon outcomes within a retrofit context). Our strategy to improve impact and usage will involve close consultation and co-development with our industry partners – who will provide use cases and actively work with us to design and realise the software platform and new digitally enabled supply chain models.

We explore four research questions (RQs):

- RQ1: How do we design, execute and govern a circular supply chain (CSC) that is environmentally sustainable and economically viable for the AEC sector?
- RQ2: How do we incentivise organisational actors to participate in a CSC, particularly by increasing transparency of operations within a CSC?
- RQ3: How do we automatically label and subsequently track the whole life-cycle activities of materials/ components/ assets and services to a) provide better insights into their composition and effectiveness at end-of-use, and b) sustain and preserve existing building stock?
- RQ4: How can economic models be developed to support the digital infrastructure that is essential to drive a CSC and sustain it over the lifecycle of an AEC asset?

The main outcomes of the research will be: a) a digital (software) platform that harnesses the potential of multi-layered blockchain (often referred to as "parachain", e.g. in Ethereum Plasma and Polkadot.Network) and the concept of 'material & service passport' to show the circularity potential of materials/components/ assets/ services and enable stakeholders (designers, main contractors, manufacturers and clients) to assess the likelihood for 5R. b) a road map based on the co-developed (with industry) digital platform, to incentivise repeated use and integration of "circularity" within industry-based systems – aimed at influencing a behaviour change in the way that the 5R are considered in the AEC sector and to enable collaborative partnerships to support CSC.

Partners and Relevant Projects

DATA OBSERVATORIES

Forest Observatory

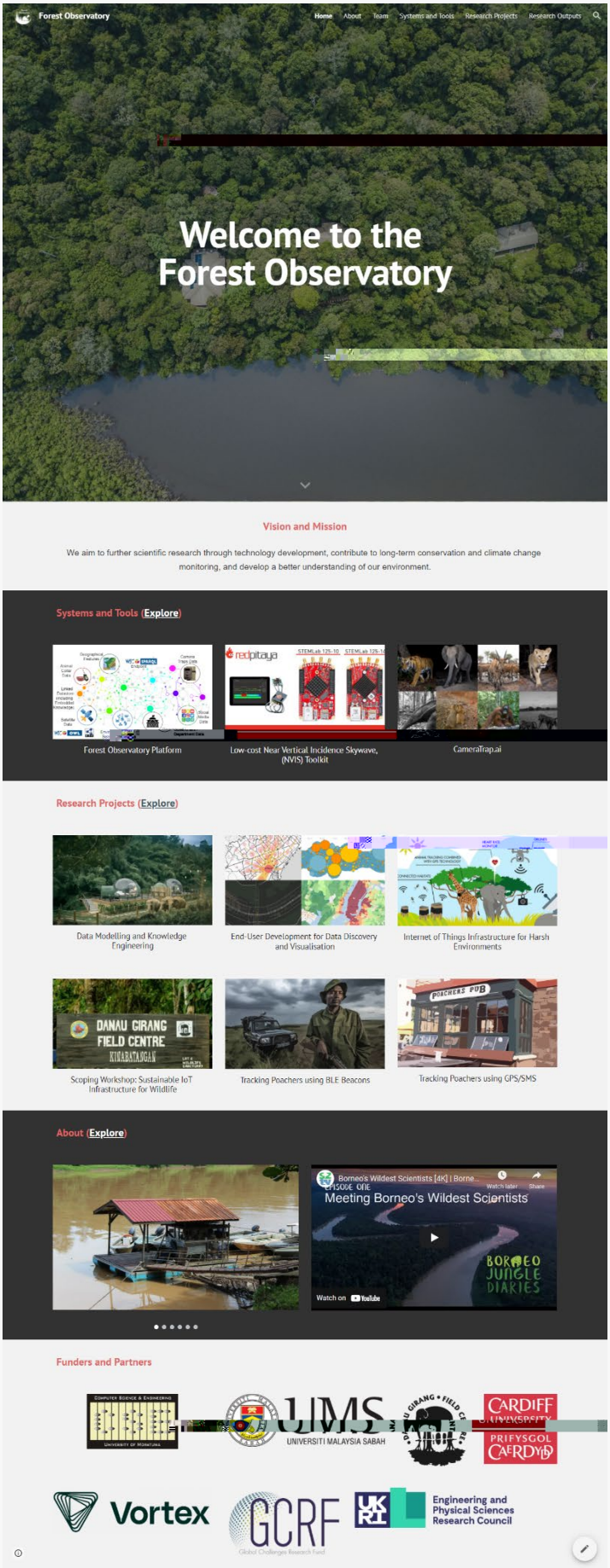
Team: Charith Perera, Omer Rana, Benoit Goossens, Pablo Orozco-TerWengel, Oktay Karakus (2021-Present)

This is a research program seed-funded through different sources over the last few years, such as GCRF Facilitation, EPSRC International Partnerships, and several Cardiff University summer student projects and travel funding. This research program aims to bring faculty and related projects and strategic recourses underutilised and scattered across the University under a coherent theme that would enhance collaborative interdisciplinary research. The program focuses on fundamental and applied research, resulting in usable tools and systems.

Forest Observatory is a Linked Datastore that integrates heterogeneous data. We consider Forest Observatory as an extension of Urban Observatories, aiming to gather real-time urban data across cities. Collecting data in forests is much more challenging than in cities due to the lack of infrastructure. However, while we expect to deploy an IoT infrastructure to enable poaching monitoring, we should utilise already collected data sets to develop predictive models to better track poaching activities.

To develop a Forest Observatory, we aim to integrate various data sets collected by the bioscience researchers at DGFC into a unified linked data store. We use semantic data integration techniques while conforming to the data modelling standards (e.g., ontologies) and needs of bioscience research --towards developing a model and novel tools that are exportable to other world areas where poaching is a threat to wildlife conservation.

Visit: forest-observatory.org



DATA OBSERVATORIES

Internet of Things Network for Forest Observatory

Team: Charith Perera, Omer Rana, Benoit Goossens, Pablo Orozco-TerWengel, (2022-Present)

The primary objective of this project is to develop a better understanding of how to design and deploy a sustainable IoT network in a remote jungle environment with harsh conditions. We deployed an IoT network with three sensors and three network extension mesh routers supported by a gateway to push data to the cloud. We wanted to understand what kind of IoT network would ideally be suited to establish a forest observatory to enable sustainable sensors data collection and wireless communication. We would like to understand potential network design and topology, estimated costs, energy requirements, and other constraining factors that may need to consider when deploying an IoT network in a jungle. Our long-term plan is to develop a forest observatory that has the capability to observe animals and the environment through heterogeneous sensors at scale to facilitate bioscience research and wildlife conservation activities. This project aims to collect data over 24 months period of time.

Danau Girang Field Centre (DGFC) and surrounded area. Researchers conduct research usually 2-4 miles from the river bank and 20 miles each side along the river (5.430443299150367, 118.0396091749387)

(From Left to Right) both the rechargeable battery pack and the sensor attached to a tree, rechargeable battery pack, sensor installed on top of the DGFC main building roof

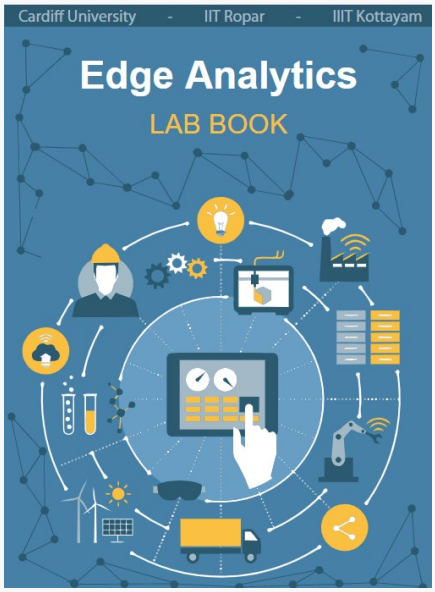
Course Development in Edge Computing and Analytics

Team: Omer Rana, Charith Perera, Nitin Auluck, Sujata Pal, Shajulin Benedict (2021-2022)

Due to the increasing popularity of cloud and edge computing, edge analytics focuses on using machine learning & AI on user-owned devices. Short-term benefits: engineers qualified to work on modern edge computing applications in the UK and India, economic benefit & student exchange, and micro-credentials offered online. Long-term benefits include research collaboration between participating institutes, research-based course content, and dedicated online labs. The course content will be co-created between the UK & India (including engagement of students & industry) based on world-class course assessment frameworks. Course delivery will follow a hybrid offline/online model. Non-credit-bearing content will be trialled with students at the three participating institutions. The proposing team has complementary expertise in - complex systems and IoT (Cardiff), edge computing and sensor networks (IIT Ropar) and IoT analytics (IIIT Kottayam).

The course titled “Edge Analytics” was chosen to acknowledge the proposal's demand and requirements raised by the participating organisations. In fact, the market analysis manifested the necessity of organising such a course due to the evolving demand from varied research domains such as healthcare, education, smart cities, government, social goodness, and so forth of the Indian/UK market. Before validating the statement, the potential students willing to travel to the UK/India were assessed for their preferences.

Based on the carefully identified partners for the proposal, we are self-assessed to deliver the best among the many other competitors for the proposed course. Due to the team's expertise in proposing the course and the huge market demand, we are confident that the course will generate significant interest among students and working professionals.



- Develop a series of lectures and accompanying slides to facilitate the delivery of the proposed module.
- Develop a series of lab tutorials to be used in teaching
- Focus on increasing the critical thinking capability of the students in the course plan.
- We aim to make this course very hands-on with a good amount of laboratory activity and projects. Some of these will involve designing and implementing edge analytics solutions for state-of-the-art real-life applications.

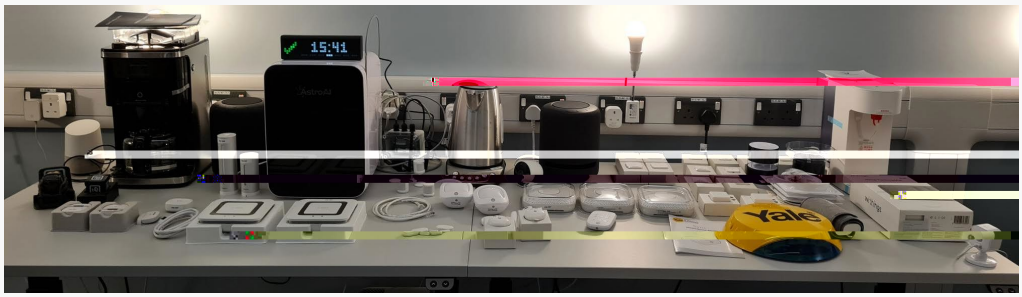
Partners and Relevant Projects

RESEARCH INFRASTRUCTURE

Smart Home Lab

Team: Charith Perera, Mary Zacharias, Mohamed Alosaimi, Yasar Majib (2021-)

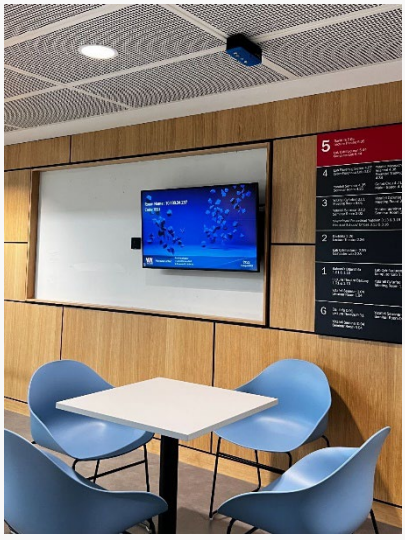
The Smart Home Lab (SHL) is a key research facility designed, developed and managed by the IoT group. It is a physical space comprised of over 170 networked smart home devices. It has six hotdesk working areas where researchers (staff and students) can conduct research with smart home devices. It comprises a wide range of devices such as smart TV, environmental monitoring kettles, various smart speakers, robotic vacuum cleaners, smart metre bells, door locks and many more. We use OpenHab and Home Assistant to capture semantic data and Wireshark to capture network traffic and related network communications and behaviours. The smart home lab also comprises a video network connected to six different types of smart home and consumer CCTV cameras to a network video recorder that allows connecting video analysis research.



RESEARCH INFRASTRUCTURE

Abacws Smart Building Testbed

Team: Charith Perera, Suhas Devmane (2022-)



The Abacws Smart Building Testbed is one of our latest projects in the design and development phase. We are developing and deploying sensor nodes across two floors in the Abacws building at high density, which capture a wide range of sensor data and enable the researcher to contact various research. We collect over 16 different types of sensor parameters at a high frequency, enabling us to monitor the building at high resolution. The nodes will be deployed in shared spaces and do not capture any privacy-sensitive information

Incubator Projects

Feasibility of Detecting Door Slamming towards Monitoring Early Signs of Domestic Violence

Researcher: Osian Morgan (MSc Student-2022)

With the ever-increasing availability of low-cost micro- controllers and other computing devices, and advances in more lightweight machine learning techniques, it is becoming increasingly viable to make many of the everyday objects found in our homes smarter. By using low-cost microcontrollers and TinyML, we investigate the feasibility of detecting potential early warning signs of domestic violence and other anti-social behaviors within the home. We created a machine learning model to determine if a door was closed aggressively by analyzing audio data and feeding this into a convolutional neural network to classify the sample. Under test conditions, with no background noise, an accuracy of 88.89% was achieved, declining to 87.50% when assorted background noises were mixed in at a relative volume of 0.5 times that of the sample. The model is then deployed on an Arduino Nano BLE 33 Sense attached to the door, and only begins sampling once an acceleration greater than a predefined threshold acceleration is detected. The predictions made by the model can then be sent via BLE to another device, such as a smartphone of Raspberry Pi.

The COVID-19 pandemic resulted in many changes and restrictions for our daily lives, most notable being mandates to work from home where possible, as well as legal requirements to socially isolate. Between March 2020 and June 2020, police in England and Wales recorded a 7% increase in offences flagged as domestic abuse related, with the ONS noting a general increase in demand for domestic abuse victim support services (including a 65% increase in calls to the National Domestic Abuse Helpline between April and June 2020, compared to the previous quarter). Overall, the entire 12- month period between March 2020 and March 2021 saw an overall increase of 6% in domestic abuse related crimes. This follows general increases seral increasen.261{1.3 an(in Tnaseeilco (s)-1.3 (flag-3 (s)-623a)

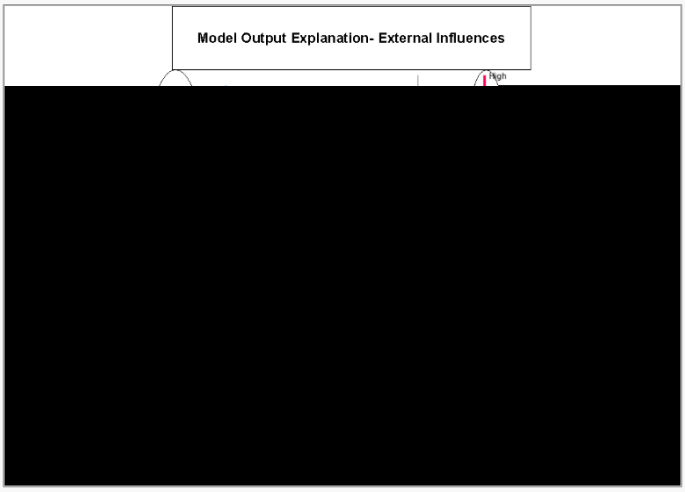
Explainable Sensor Data-Driven Anomaly Detection in Internet of Things Systems

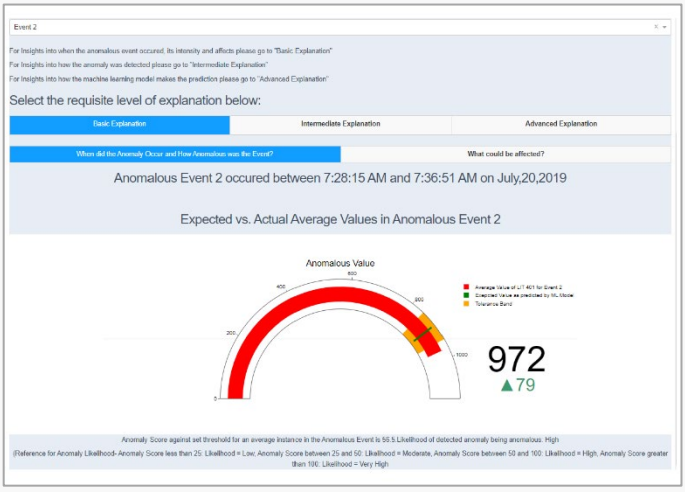
MSc

Researcher: Moaz Tajammal Hussain (MSc Student-2022)

Deep learning or black-box models are widely used for anomaly detection in Internet of Things (IoT) data streams. We propose a technique to explain the output of a deep learning model used to detect anomalies in an IoT based industrial process. The proposed technique employs dual surrogate models to deliver black box model explanation. The dashboard integrates our proposed deep learning explanation technique with historical logs to explain the detected anomaly for personas with different backgrounds.

Deep neural networks have shown robust anomaly detection capabilities. They are capable of capturing temporal and multimodal dependencies. Moreover, they allow for minimal manual feature engineering and domain knowledge independent data pre-processing. Conversely, deep learning or 'black box' models, are difficult to explain. This work presents a technique to explain the output of a unsupervised deep learning model. The well-known IoT dataset of Secure Water Treatment or SWaT has been used for model training and anomaly detection.





Anomalies are detected by monitoring reconstruction errors of LSTM Auto-encoder. LSTM Auto-encoder's (LSTM-AE) output for detected anomalies is then attempted to be explained by training a duo of Random Forest regression models. The surrogate models are trained to replicate the output of the LSTM-AE. We then use SHAP plots to explain the output of the surrogate models.

Each surrogate model captures unique dependencies of the deep learning model for the probed output and is decrypted using TreeSHAP. Finally, dashboard is designed to answer the questions (when, how, what, and why) associated with the detected anomaly for different personas.

Outcomes

- [Poster] Moaz Tajammal Hussain, and Charith Perera, **Explainable Sensor Data-Driven Anomaly Detection in Internet of Things Systems**, In Proceedings of the 2022 IEEE/ACM Seventh International Conference on Internet-of-Things Design and Implementation (IoTDI) 2022, pp. 80-81 PDF BIB SOURCE POSTER VIDEO

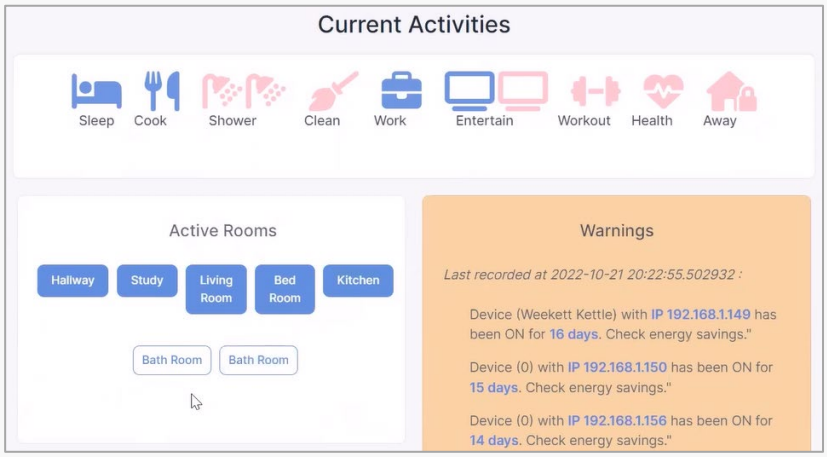
Smart Home Activity Inference using Network Data

Researcher: Mary Zacharias (MSc Student-2022)

MSc

This project seeks to infer physical events in real-time from an ensemble of off-the-shelf smart home devices using only encrypted network data. A "smart-home style network" is set up to develop the proposed system. A multi-stage classification process is then used - a set of classifiers are trained to first 'fingerprint' connected devices using passive techniques and then to detect binary and more specific differences in the states of these connected devices. These classifiers are then deployed to validate the system within a smart home 'diagnostics' application. This layman-friendly solution offers users real-time visibility and correlated insights into their smart home devices' current and historic physical activities -without any manual set-up or device integration necessary.

This project achieves a 99.3% fl score in detecting a connected device as 'IoT', 96.0% fl score in fingerprinting the device, and 90.1 % fl score in identifying the state of the targeted smart home device.



Through these results, we seek to demonstrate that network data could be used to supplant more complex device data collection techniques within diagnostics or even simple Human Activity Recognition tools in smart home settings.

The primary motivation for the proposed system is to assist diverse user groups in getting access to insights on their home devices' activities and, by extension, their own activities through simple, non-invasive, hardware-agnostic means. As such, a fundamental requirement is to support in answering questions like "What are the current states of all connected devices?", "Are there any seasonal activity patterns?" and even "What was happening in the house up to and during a particular event?".

While historically, research and development in activity recognition catered to specific user groups like healthcare assistants and 3rd-party researchers, a general-purpose activity recognition application could prove useful to a much wider audience of smart homeowners, care home managers, IoT enthusiasts, students, and residents of the smart home themselves.

Outcomes

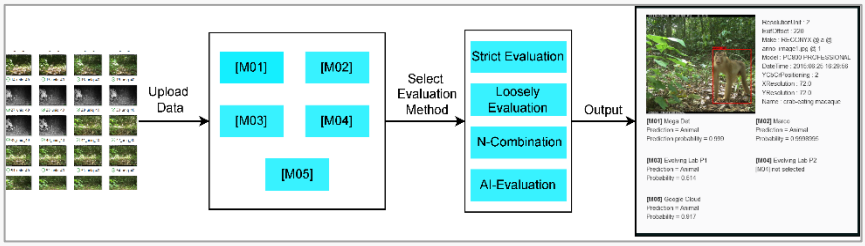
- **[Demo]** Mary Zacharias, Charith Perera, **Smart Home Activity Inference using Network Data**, Work-in-Progress VIDEO CODE

CamTrap.AI: Semi-Automatic Wildlife Image Classification Using Ensemble Learning

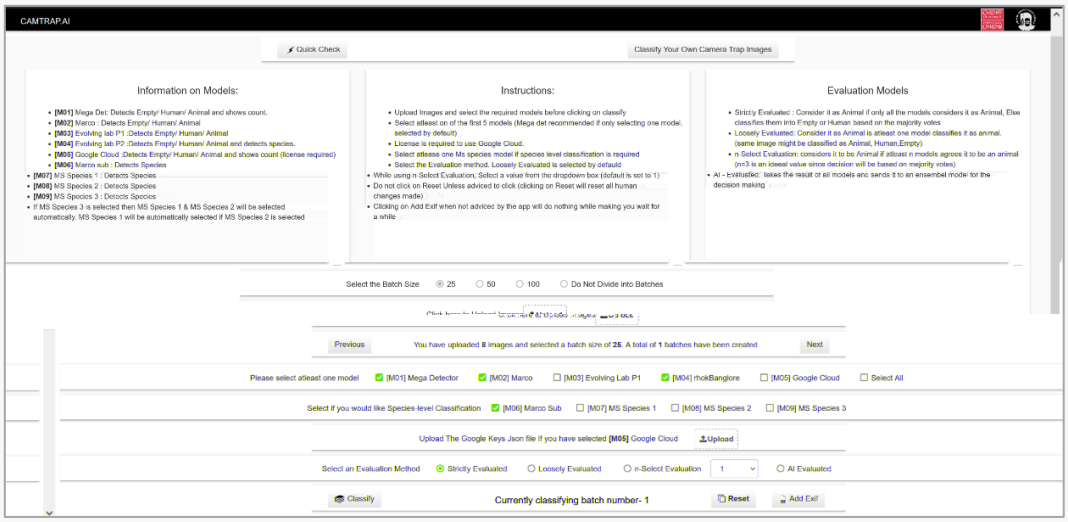
MSc

Researcher: Akin Kaki (MSc Student-2022)

Bioscientists and wildlife conservationists regularly deploy trail cameras (Camera Traps) to monitor species and their habitats. These cameras capture moving objects, including animals, using motion sensors. Depending on the project's purpose, the camera trap deployment might take a few weeks to several months. During this time, each camera would capture thousands of images. However, most of them could be false positives due to the sensitivity of the embedded sensors (i.e., environmental factors such as wind and rain could trigger them). Consequently, camera operators must filter out images manually. Classification and recognition of images by hand need time and effort. Even though many image classification tools have been created to solve this issue, academics are wary of delegating the classification task to software tools alone - out of concern that these tools may misclassify images containing valuable data.



To address this issue, we propose CamTrap.AI, a tool that uses ensemble learning to classify the presence of animals in an image and determine their type. The crucial aspect of our solution is that we combine the results of more than a dozen well-known image classification models to create a single conclusion, improving the system's overall performance and boosting the confidence of human decision-makers. Our tool enables researchers to override or affirm the ensemble model's decisions through a semi-autonomous decision-making approach. We have examined each model's performance separately using various ensemble learning techniques to determine which outperforms and makes solutions based. In addition, we re-executed the full evaluation on a Raspberry Pi to evaluate if our strategy could produce better results when decisions were to be made at the edge. We investigated the proposed method with two distinct massive datasets (LILA BC Dataset and our DGFC dataset).



Outcomes

- **[Demo]** Akin Kaki, Naeima Hamed, Pablo Orozco Ter Wengel, Benoit Goossens, Omer Rana, Charith Perera, **CamTrap.AI: Semi-Automatic Wildlife Image Classification Using Ensemble Learning**, Work-in-Progress RESOURCES CODE CODE

Search Images from Similar Locations in a Camera Trap Dataset

MSc

Researcher: Akin Kaki (MSc Student-2022)

Camera trap images were widely used in the study of wildlife, but most of the time, the location information of the images (Camera traps) was kept unknown. This can be due to using low-cost cameras that don't have a GPS sensor or to keep the location of rare species hidden. This information is lost once the researcher finishes their work. However, this causes an issue when new research is done using the same data or when a researcher uses an older dataset due to a lack of access to real-time data directly from the forest. To solve this problem, we propose a technique that uses a few images from a particular location to generate a background image and compare it with the rest of the data to find images from the exact location and develop a proof of concept to implement it. Some manual work is required initially. While it can also be solved using machine learning algorithms, these models generally create tags for images that might give a false positive when searching for a location, as the same animal can be present in two different locations. The models can tag the animals instead of the location. Initial data required to train the model is enormous and need to be manually identified, defeating the purpose of automation.

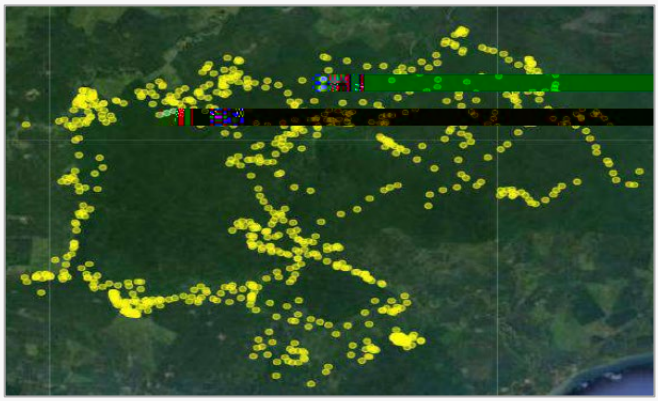
We tested this method with some custom data and got an accuracy of 88.48%, while using the Machine learning model only got an accuracy of 44.68%. Once the whole dataset is processed using this approach, a single image can be used to find to which location it belongs (search by image) and achieve an overall accuracy of 95.6%. Camera-trap datasets are not open sourced sometimes, in that case manual processing of the whole dataset (which might contain 100000 to millions of images) or once the researcher identifies images from a location, a manual process is required to determine the presence of animals in the images. Machine learning models cannot be used sometimes as they might only have low power devices and/or no internet access in a remote location, or the dataset is a closed source. So, we used the same process of comparing the background image to identify animal presence in the images. We tested on a subset of the custom dataset we used in the above process and achieved an accuracy of 71.48%.

Predict Animal Movements using Collar Data

BSc

Researcher: Jacob Harkins (BSc Student-2022)

The project explores the dataset to help patrollers better understand elephant movements and thus locate them easily. With the prevalence of consistent elephant movement trends in the usage of natural forest corridors, this project aims to develop and train a machine learning regression framework to predict the future GPS locations and movement trends of a Bornean pygmy elephant located in Sabah, Malaysia, in the hopes that it can be used as a secondary tool to base patrol routes around. CODE

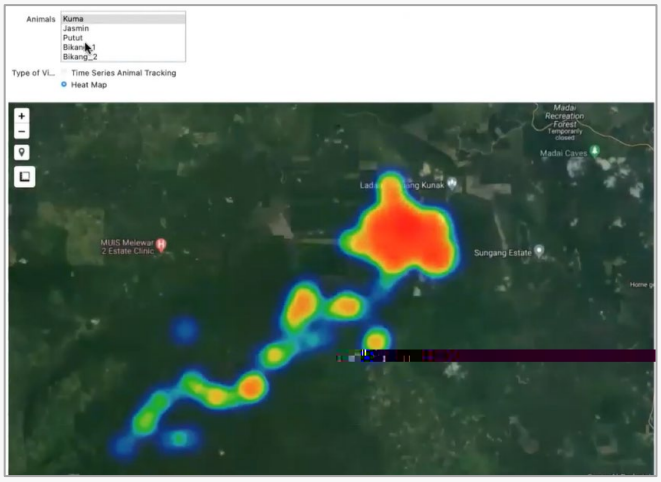


Interactive Data Science for Forest Observatory

BSc

Researcher: Sam Winkworth (BSc Student-2022)

Observing forest environments gives researchers vital information for effective planning and appropriate responses to these events. It can also help spotlight trends that otherwise may not have been noticed. This project explores various data science techniques to tell a story of how the Kinabatangan Forest and the surrounding area of Sabah have changed over time. VIDEO CODE

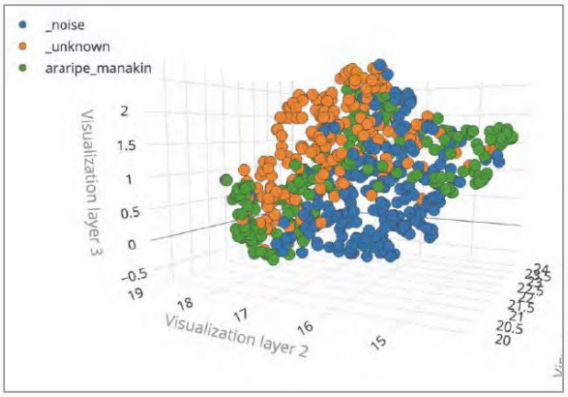


Classify Species using Sounds at the Edge

BSc

Researcher: Ruslan Levond (BSc Student-2022)

Automatic sound recognition systems have proven to be effective during conservation activities. Several devices out on the market can be installed in the wild and record animal sounds. However, they are inaccessible due to being expensive, and they only record sounds and require additional proprietary software to classify sounds elsewhere. The project intends to create an alternative low-cost device that can be installed in the wild and be able to record and classify animal sounds right on edge. This involved developing a machine learning model from scratch that can classify bird sounds. The project also created edge frameworks for two architectures, Raspberry Pi and Arduino, to which



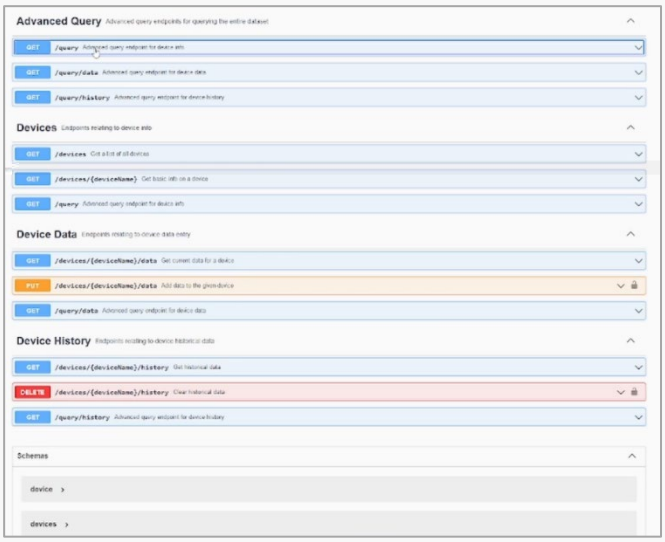
the machine learning model is deployed to. The project then created a gateway device and a framework for it which is used to store results transmitted by edge devices. The project has also investigated the performance of the model running on both architectures and compared architectures to understand which one is more suitable to use when. VIDEO CODE

Visualise IoT Data in Abacws Building

BSc

Researcher: George Grainger (BSc Student-2022)

This project aims to develop a web application to visualise data gathered from IoT sensors in the Abacws building. The purpose of this application is to make it easier to access and visualise IoT data within the Abacws building. Displaying the data on a 3D model like this is useful as it lets the user know visually where the sensor/device is located within the building. This is useful within a university setting as some studies may, for example, require access to data such as the temperature in a particular room at a particular time. Our application empowers people to easily collect and use data from these sensors with little technical understanding.



The application's target audience is professors and students at Cardiff University who want or need to access the data out of curiosity, for information, or as part of a wider study. This application presents a navigable 3D model of the Abacws building with devices clearly shown inside it. Upon selecting a device, the information and data related to that device should be fetched from a data store and shown to the user. The user can then view historical data for that device to easily identify trends. For more advanced use cases, the application will expose several endpoints to make it easy to query the dataset of all devices. This provides a powerful aggregation of IoT data in a single source. VIDEO CODE RESOURCES

gitlab.com/IOTGarage

bit.ly/2JMoSd4

@IOTGarageNews

INTERNET OF THINGS
GARAGE
iotgarage.net