UKRI
Research Node on
TAS Governance & Regulation

King's College London
Heriot-Watt University
The University of Edinburgh
University of Glasgow
University of Nottingham
UK | CHINA | MALAYSIA
University of Sussex

INTRODUCING THE TAS NODE

Our Vision

To devise a framework and tools for the regulation of autonomous systems – to help them function safely and respond appropriately to the complexities of the world they function in

Principal Investigator's Vision

Principal Investigator's Vision

How can we trust autonomous computer-based systems? By autonomous we mean "independent and having the power to make your own decisions". We tackle the issue of trusting autonomous systems (AS) by building: experience of regulatory structure and practice, notions of cause, responsibility and liability, and tools to create evidence of trustworthiness into modern development practice.

Modern development practice includes continuous integration and continuous delivery. These practices allow continuous gathering of operational experience, its amplification through the use of simulators, and the folding of that experience into development decisions. This, combined with notions of anticipatory regulation and incremental trust building form the basis for new practice in the development of autonomous systems where regulation, systems, and evidence of dependable behaviour co-evolve incrementally to support our trust in systems.

Our Research Node brings together a diverse multi-disciplinary team from Edinburgh, Heriot-Watt, Glasgow, KCL, Nottingham and Sussex, involving computer science and AI specialists, legal scholars, AI ethicists, as well as experts in science and technology studies and design ethnography. Together, we present a novel software engineering and governance methodology that includes:

1) New frameworks that help bridge gaps between legal and ethical principles (including emerging questions around privacy, fairness, accountability and transparency) and an autonomous systems design process that entails rapid iterations driven by emerging technologies (including, e.g. machine learning in-the-loop decision making systems)

2) New tools for an ecosystem of regulators, developers and trusted third parties to address not only functionality or correctness (the focus of many other Nodes) but also questions of how systems fail, and how one can manage evidence associated with this to facilitate better governance.

3) Evidence base from full-cycle case studies of taking AS through regulatory processes, as experienced by our partners, to facilitate policy discussion regarding reflexive regulation practices.

Our work is grounded in concrete problems arising within our wide network of partner organisations, who are working at the forefront of mobility, health and social care, and a variety of other application domains that impact on our daily lives – raising the need for a multi-faceted understanding of trustworthiness. Therefore our methodology is iterative, interleaving basic methods development within the node with engagement in the form of case studies to guide the Node's activities towards questions that are not only scientifically interesting but also societally relevant.

Beginning such an initiative in the middle of a global pandemic, with all of the uncertainties implied by public health measures, has made reaching outwards particularly hard. Not only has this restricted engagement with public, but also the pressure on frontline staff such as in health and social care has set limits on engagement with the professional community.

At the same time, the various conversations around formal regulation of AI and mechanisms for informal governance within organisations have now become even more timely. Governments representing all major economies in the world today are deliberating mechanisms for governance and regulation of technology, some of which speak directly to autonomous systems and the specific issues of relevance to the TAS programme. We are excited to contribute to this conversation, developing a methodology for better governance that allows society to benefit from the potential envisioned by the technological optimists, without compromising the broader social good.

Prof. Subramanian Ramamoorthy
Principal Investigator
UKRI Research Node on TAS Governance & Regulation

Pillar 1: Legal and Social Aspects of TAS

The Aim of Pillar 1 is to analyse how formal laws, standards, ethics and other forms of normative systems shape and impact on design practices in the TAS development cycle develop understanding of the best approaches for TAS governance, with particular attention to the TAS development process and the role of a new type of regulator: system designers.

We need to situate and unpack what their legal and ethical responsibilities are and how they can respond to these. We explore how legal norms, rules and human values can be embodied in TAS, to reflect on how legal problems can be anticipated and even addressed before the system is deployed in the world

This task can be broken up in a number of sub-questions:

1. What is the current regulatory landscape, how do different modes (e.g. formal laws vs ISO standards vs ethics guidelines) and levels (subnational, national, international) of regulation interact?

2. What aspects of TAS are not captured, or not captured well, by these existing forms of control?

3. What is the track record of these regulatory approaches, what has worked well in the past, what didn't, and what can this tell us about the future of TAS regulation?

4. What is the role of law in creating trust? How does regulation and design together create trust?

Answering these questions will allow us to help **understand** the existing regulatory regime and evaluating its strength and weaknesses. It will **assist developers** as new "subjects and agents of regulation in law-compliant design choices, and to **influence** policy makers to affect better regulation. Finally, it will inform better **communication** to the wider public, enabling more realistic assessments of the dangers and potential of AI and lead ot more prudent interactions.

Representative examples of our work:
Both the EU and the US have since the inception of the project initiated ambitious legislative projects to build comprehensive regulatory regimes for AI. For post-Brexit UK, this poses a significant regulatory challenge: which, if any, of the two models to align with, and how the UKs emerging governance framework for AI would interact with these if there is substantial divergence. We organised an expert roundtable on the EU Act, and on the basis of this identified a number of concerns and opportunities stemming from the proposal Our results are already feeding into the expert group of Police Scotland on the use of advanced technologies in policing. We explore also how these ideas interact with existing regulatory regimes, including the indirect effect that fields of law such as intellectual property have on the success of Ai specific regulation or the lack thereof

The Covid crisis brought AI enabled medical technologies and their regulation into the public limelight. It also brought the issue of "trust" into sharp relief, and asks the question how a crisis can shape our expectations of trustworthy technology. Within the context of medical devises as one of our case studies, we explore in particular how laws respond, or do not respond, to crisis situations, how that has effects on the trust we place in technologies
le domains and past experiences, alongside legal analysis.

Pillar 1: Team

Burkhard Schafer

Co-Investigator – Pillar Lead

- AIpology: the role of apologies for trustworthy AI. This project looks at the way in which trust in AI can be restored after a problem procured, and investigates in particular side based and legal solutions that create a space for AIs to apologise.
- Enacting code: One way to bridge the gap between legal language and computational compliance is to change the way laws and regulation are enacted. Can we think of new forms of legislative drafting that incorporate code directly into the legislative text and create “authoritative computational translations” of laws
- Counting (on) AI: one problem for the legal regulation of trustworthy systems is to provide appropriate identity criteria. Physical products are considerably easier to regulate in this respect, as they remain stable over time. This part of the project looks at the way in which legislative proposals such as the EU AI Act individuate AIs

Chris Marsden

Co-Investigator

- My research focuses on the application of Artificial Intelligence to democratic processes, in particular election cybersecurity and political disinformation, with major recent reports for the European Parliament (2019) and Commonwealth (2020)
- Within TAS G&D Pillar 1, my main focus is on the future of Artificial Intelligence regulation, focussed on those issue areas above, and the proposed EU Artificial Intelligence Act and its enforcement, where the rhetoric very often is completely at odds with the reality and poorly understood by observers outside the European institutions.

Lachlan Urquhart

Co-Investigator

- Providing IT Law and Human Computer Interaction expertise in cross cutting and legal pillars.
- Leading work packages on developing 'trustworthy by design' approaches to translate between legal compliance requirements and TAS development practices.
- Mapping key principles of UK/EU governance frameworks around TAS, particularly within a case study focusing on UAVs.
- Developing 'legal provocations for design', particularly drawing on the proposed EU Artificial Intelligence Act.

Shannon Vallor

Co-Investigator

- I contribute to this pillar on legal and social aspects, as well as to the cross-cutting themes working group by helping to integrate the ethical dimensions of governance across all the research pillars.

Phoebe Li

Co-Investigator

- Risk regulation and intellectual property aspects of AI systems, especially on medical device

Robin Williams

Co-Investigator

- Information and Communication Technologies (ICT);
- Life Science Innovation
- Environmental Innovation and Transition to a sustainable society

Daria Onitou

Research Associate

- Regulation of accountability and responsibility of AI systems currently focusing on AI in the medical domain:
- Scrutinising gaps in regulation regarding AI systems (such as, the new EU AI Act).
- Integrating knowledge on computational methods & studies to develop future legal and ethical governance of AI systems.

TAS Modelling: Accountability, Explainability and Responsibility

This pillar provides a novel modelling framework, and tools, to help regulators and developers consider chains of causal factors leading to unexpected/undesired outcomes in Autonomous Systems. Regulation requires terms such as “desirable” or “unexpected” in human-centred systems to be elaborated. For instance, we want decisions to be `fair’, but can its meaning be made auditably clear? Recent AI advances, e.g., methods for local explanations also address some aspects of such questions. However, not only are these methods insufficient for describing issues arising due to complex interactions between human, institutional and technical factors, but also good TAS governance must juggle disparate forms of responsibility: moral, legal, professional, and causal.

Pillar 2 research directions:

To achieve the above goal, Pillar 2 has the following objectives,

1. Develop a responsibility framework (RF) that will focus on elucidating the causal chains of accountability and responsibility to help the regulator/developer specify anticipated harms and audit system behaviour with respect to these. To do this, we will first understand elicit different notions of responsibility by involving clinicians, system developers, regulators, auditors, and customers. We will then design models to assess the extent to which the distinct types of responsibilities have been realised.

2. Develop techniques to help explain and interpret the outputs of autonomous system. The explanation and interpretation techniques will be used to assess the importance of each input variable for the outcome. These approaches will be integrated with analysis relying on causal models to rank inputs and assign quantitative responsibilities of the inputs in the outcome.

3. Integrate quantitative and qualitative responsibilities into the regulatory framework. The qualitative and quantitative responsibilities will complement each other and provide the necessary data to determine whether the regulations have been met.

4. Analyse the process in cases of errors in outcomes. Provide means to determine whether the error occurred in the autonomous system component or in the interaction between the system and a human user. Feedback the conclusions to improve the interpretability of the output of autonomous systems.

5. Consider different scenarios where interpretability is needed and characterize different types of output to support the interpretability, for instance: explanations of recommendations, apologies, liability, apportionment of blame, etc.

Pillar 2: Team

Hana Chockler

Co-Investigator – Pillar Lead

- Investigating reasons, causes and explanations of software engineering and machine learning procedures
- Ongoing research project on explanations of reinforcement learning policies
- Leading Pillar 2 research activities and acting as Liaison for the TAS Hub.

Vaishak Belle

Co-Investigator

- Integrating reasoning and learning,
- explainability and ethics in AI,
- Automated reasoning, planning and programming

Phoebe Li

Co-Investigator

- Attribution and demarcation of liabilities in the complex chain of stakeholders

Shannon Vallor

Co-Investigator

- My core research interests are the ethics of data, AI and robotics
- My present research focus is the impact of AI and autonomous systems on human capabilities and virtues
- In Pillar 2 I supervise PhD research on a qualitative framework that maps and integrates the different types of responsibility that impact governance of autonomous systems (moral, legal, causal, professional, organisational/corporate, and regulatory responsibility).
- I also contribute to the Cross-Cutting Themes working group by helping to integrate the ethical dimensions of governance across all the research pillars.

Xin Du

Research Associate

- Model evaluation;
- Causal Learning;
- Robustness of Machine Learning Models;

Ajitha Rajan

Co-Investigator

- Automated software testing techniques
- Artificial intelligence for cancer immunotherapy.

Benedicte Legastelois

Research Associate

- Formal methods for AI and ethical AI
- Combine them and focus on practical problems like trustworthiness and understanding of Mlgorithms
- main role is to study explanation technics on algorithm for medical diagnosis and especially models using cancer images

Subramanian Ramamoorthy

Principal Investigator

- Machine learning for decison making and safe autonomy
- Human-robot interaction
- Simulation frameworks for performance evaluation
- Uncertainty and robustness quantification

Bhargavi Ganesh

PhD Student

- Application of data science/AI tools to public policy research
- Conceptual and empirical work on risk mitigation/governance of AI given its broader impact on society

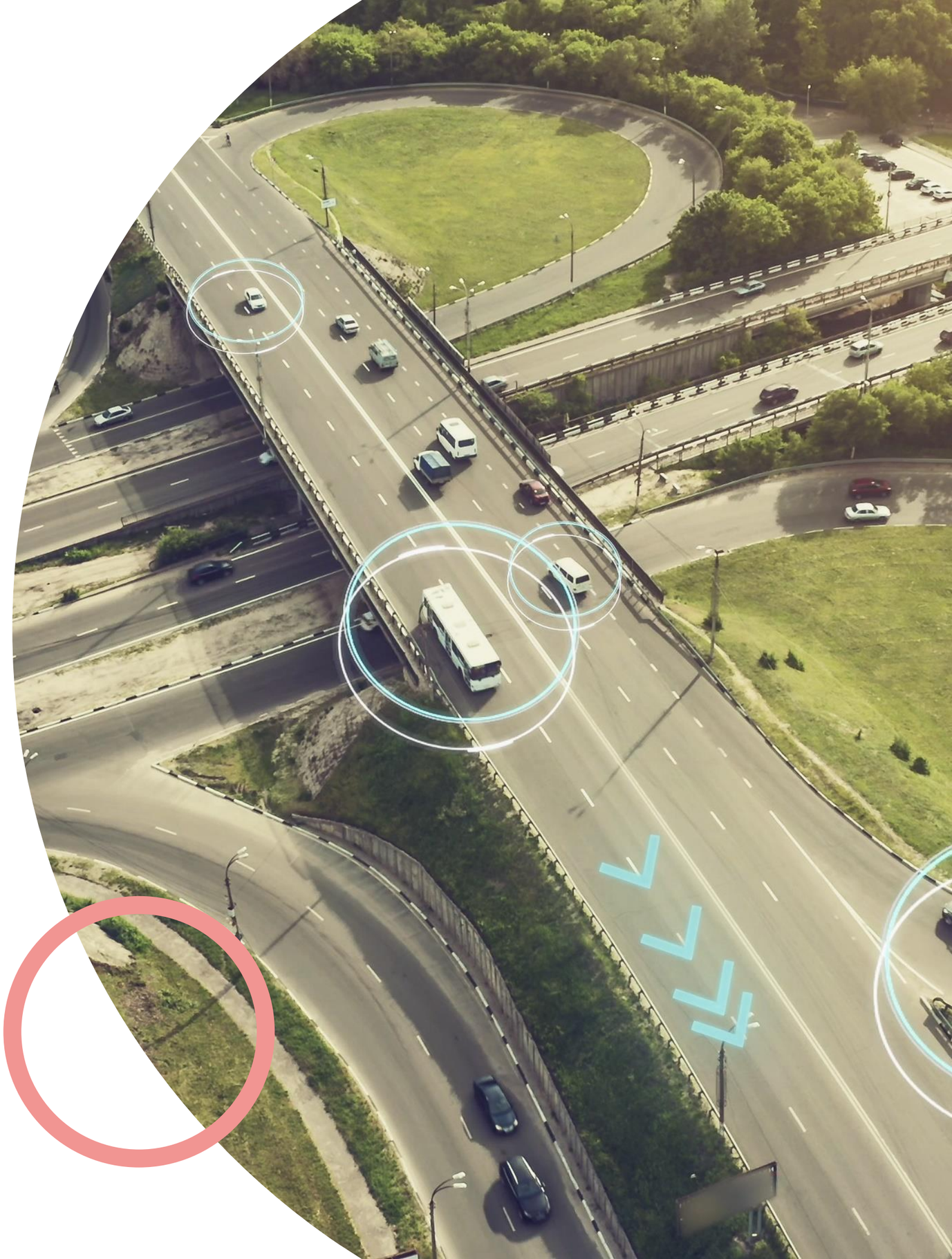
Pillar 3:
Computational Tools

In terms of tool support, we are focusing on bridging the gaps that traditionally exist during the early stages of system development. For instance, the gaps between regulatory regimes, systems engineers and software engineers. These gaps can result in the loss or distortion of knowledge that can have catastrophic consequences if undetected until system deployment. We are also focused on bridging the gaps between high-level design specifications and the variability and unpredictability of operational environments, such as commercial air space or the public road network. Deploying Natural Language Processing technology is critical to our approach. We will be using it in three ways:

A To enable users to control the simulations by articulating environment descriptions and task specifications in natural language;

B To provide novel technology for linking natural language utterances to formal symbolic representations that express the intended information about the environment and the task (as required to achieve (a));

C By extracting structured information from incident reports to make it more amenable to analysis and search. Bridging these gaps will strengthen the integrity of autonomous systems, and ultimately, we believe, will increase the trust that the public have in autonomous systems.



Pillar 3: Team

Alan Bundy

Co-Investigator – Pillar Lead

- The automated diagnosis and repair of faulty representations of the environment.
- Drawing inferences from information stored on the Internet, using diverse methods of inference, including statistical and deductive methods.
- My contribution to Pillar 3 of the Governance and Regulation TAS Node will be to use representational repair techniques to correct and enrich both the specifications of autonomous systems and the representation of the environment they operate in.

Alex Lascarides

Co-Investigator

- Multimodal Natural Language Processing
- Interactive Task Learning
- Dialogue Systems
- Learning optimal planning strategies

Andrew Ireland

Co-Investigator

- Formal methods, software engineering, and automated reasoning.
- Bridging the gaps between regulatory regimes, specifications and design prototypes will strengthen the integrity of autonomous systems. We believe that this will also lead to an increase in the trust that the public have in autonomous systems.
- The key to “bridging the gaps” lies in understanding the synergies that exist between formal modelling and simulation. Therefore, to support the development of high integrity autonomous systems, we will exploit these synergies to enhance the capabilities of existing formal modelling and simulation environments.

Yuhui Lin

Research Associate

- Automatic reasoning
- Formal verification
- Formal specification
- Software engineering

Subramanian Ramamoorthy

Principal Investigator

- Machine learning for decison making and safe autonomy
- Human-robot interaction
- Simulation frameworks for performance evaluation
- Uncertainty and robustness quantification

Craig Innes

Research Associate

- Automated Testing and Verification of Data-Driven Systems
- Combining formal logic with robotic control
- Probabilistic specification languages and learning specifications from examples
- Post-doctoral research associate. Leveraging UAV and AV simulators to test whether current automated systems conform to established traffic rules. Extracting formalized rules for driving and flight from human incident reports, road data and flight data.

Alice Miller

Co-Investigator

- Formal verification, specifically model checking and in particular the use of model checking for planning
- combinatorial search, and the use of symmetry for state space reduction
- probabilistic reasoning
- co-supervisor of Yuhui
- knowledge of probabilistic model checking (prism)

Paul Jackson

Co-Investigator

- How to describe high-level requirements from autonomous system stakeholders in formal languages
- How to check conformance to such formalised requirements by simulations or live tests
- How to use formal verification techniques to analyse the consistency of such formalised requirements

Ajitha Rajan

Co-Investigator

- Automated software testing techniques
- Artificial intelligence for cancer immunotherapy.

Matthias Appelgren

Research Associate

- Natural language processing
- Interactive task learning

4.4 Cross-cutting Theme: HCI; Design Ethnography

The cross-cutting theme focuses on 1) translating between social and technical perspectives on TAS governance, 2) enabling responsible innovation, and 3) informing design through ethnographic studies. The theme thus seeks to sensitise technological perspectives on TAS to governance requirements, b) to shape the design of socio-technical solutions that enable responsible innovation, and c) to leverage ethnographic studies to investigate and elaborate the governance challenges that confront autonomous systems from the point of view of current practice.

These distinct foci are explored through three dedicated work packages (WP) to inform interdisciplinary development of TAS:

Translating governance

This work package focuses on assessing the alignment of proposed technical developments with existing and emerging regulation. The purpose is to bring the regulatory provocations that emerge from Pillar 1 to the forefront, understand the extent to which technological proposals address them, and identify significant gaps that require attention.

Enabling responsible innovation

This work package foregrounds the importance of responsible research and innovation (RI) for TAS development. The purpose of is to understand how RI can inform developer, regulator and stakeholder thinking and practice. It engages with upstream developments, shaping processes for anticipating new harms and informing design choices, and downstream decisions about system deployment and post-implementation monitoring.

Informing design

This work package reaches beyond the pillars to engage with ethnographic studies relevant to understanding TAS governance challenges. The purpose is to understand governance challenges from the perspective of current practice by investigating cases of autonomous systems development and exploring the governance challenges implicated in their realization

Progress to date has involved mapping governance concepts, including the EU's proposal for the regulation of AI, to understand key themes governing TAS development: results highlight the strong equation of trust with risk management. We have also conducted ethnographic work with our stakeholders, which explore the engineering of trust and the governance challenges of managing margins of error in autonomous systems. An additional thread of work leverages pump-priming funding from the TAS hub to explore citizens reactions to autonomous systems.

Crosscutting Theme Team

Andy Crabtree

Co-Investigator – Theme Lead

- Design ethnographer, understanding governance challenges in current practice
- Leading work package (CCT WP3) on ‘informing design’
- Coordinating with CCT WP1 on translating governance
- Cross-cutting theme lead

Robin Williams

Co-Investigator

- Leading work package on Enabling responsible innovation which foregrounds the importance of responsible research and innovation (RI) for TAS development.

Shannon Vallor

Co-Investigator

- I contribute to the Cross-Cutting Themes working group by helping to integrate the ethical dimensions of governance across all the research pillars.

Lachlan Urquhart

Co-Investigator

- Leading work package on Translating governance – focusing on assessing the alignment of proposed technical developments with existing and emerging regulation.

Alex Lascarides

Co-Investigator

Glenn McGarry

Research Associate

- General research interests: field studies for interdisciplinary design
- Specific interests in audio production and the creative industries
- Contributing to the governance node through field studies of autonomous systems in practice.

Project Partners

Adelard

- Adelard LLP is an independent product and services company, founded in 1987, that supports its clients in the areas of safety, dependability, security and risk management. Our applied research programme enables us to bring cutting-edge techniques and methods to bear to solve our clients' difficult problems.

BAE Systems

- At BAE Systems we serve, supply and protect those who serve and protect us, in a corporate culture that is performance driven and values led. We employ a skilled workforce of 89,600 people in more than 40 countries. We help our customers to stay a step ahead when protecting people and national security, critical infrastructure and vital information. We also work closely with local partners to support economic development through the transfer of knowledge, skills and technology.
- The highly dynamic operational environments associated with the Defence domain pose unique challenges for the Safe, Legal and Ethical development and deployment of autonomous systems. Appropriate Governance and Regulation will be key to ensuring that well-formed Legal and Ethical principles guide the evolution and use of such systems. We hope to draw on the combined cross-domain knowledge and real-world experience associated with the TAS Governance and Regulation Node to assist with addressing these challenges; identifying specific Defence use-cases to guide research, and leveraging non-Defence application read-across where possible.

CAA

As the UK's aviation regulator we work so that:

- The aviation industry meets the highest safety standards,
- consumers have choice, value for money, are protected and treated fairly when they fly,
- through efficient use of airspace, the environmental impact of aviation on local communities is effectively managed and CO2 emissions are reduced,
- The aviation industry manages security risks effectively.

We are a public corporation, established by Parliament in 1972 as an independent specialist aviation regulator. The UK Government requires that our costs are met entirely from charges to those we provide a service to or regulate.

Most aviation regulation and policy is harmonised across the world to ensure consistent levels of safety and consumer protection. Worldwide safety regulations are set by the International Civil Aviation Organisation and within Europe by the European Aviation Safety Agency.

Craft Prospect

- Craft Prospect is a space engineering practice that delivers mission-enabling products and develops novel mission applications in order to realize SMART SECURE SPACE. We aim to be the partner of choice on small space missions and provide dedicated services in an agile manner, adding value to the overall service offering of our clients. We architect and deliver smarter satellites making full use of the limited resources available, targeting their products and services at the global small satellite market.
- We leverage state-of-the-art advances in embedded computing, machine learning and quantum cryptographies to create solutions that make space systems smarter and more secure and vastly improve the return for end users.

- At Craft Prospect, we are developing autonomous mission operations and data management capabilities for small satellites. These activities, currently performed on the ground either by a human or with human oversight, have a critical role in the success of Earth satellite missions, particularly in the domain of Earth observation. As the activities are automated and moved on-board the satellites, many benefits are realised, but at the cost of oversight and transparency, diminishing the trust in the activities. We are developing assurance methods and a suite of artefacts to instil trust in our customers and other end users that autonomous on-board operations and data management will be improve the return of their mission and not compromise its security, safety or value. In tandem to this work is the exploration of how autonomous space systems can be governed and regulated. Current regulations are based on the expectation of human-operated satellites and do not consider the value or importance of the data acquired by the satellites. We hope to better understand how we can engage with UK government and others to ensure autonomous space system development and its benefits aren't hindered by out of date regulations and licencing approaches.

Project Partners

DHI

- The Digital Health & Care Innovation Centre (DHI) is part of the Scottish Funding Council's Innovation Centre Programme, which is designed to support transformational collaboration between universities and businesses.
- We transform great ideas into real solutions. Our vision is "innovation in digital health and care will help the people of Scotland live longer, healthier lives and provide sustainable and inclusive growth for our economies."
- We play a pivotal role in supporting collaboration to co-design person-centred digital health and care solutions across service, technical and business innovation. We are shifting the balance of care from a traditional treatment focused model to one that focuses on prevention, detection, post event care and self-management.

Digital Health & Care Innovation Centre

D-RisQ

- D-RisQ is a leading high-tech UK company specialising in producing software and systems verification technologies based in Malvern, Worcestershire. Since 2012, we've built an expert team of software consultants, analysts and developers to help "change the way the world does software". By bringing advanced automated software verification tools to safety-critical, security-critical and business-critical system developers, we enable substantial savings in the time and cost of software development. These tools also provide organisations with the proof of compliance to internationally recognised standards in various sectors. The legacy technology we use is based on over thirty years of development in the UK defence sector and has been licensed to us for further development and commercialisation for the benefit of all business sectors. Over the years, we've harnessed hundreds of man-years of development effort and applied it through a team of highly qualified mathematicians and computer scientists. Following eight years of R&D, we have taken the core technology and created a suite of world-leading software assurance tools.
- D-RisQ has developed assured "safety cages" for autonomous air, surface and sub-surface sea vehicles using our assurance techniques. The interest is in aligning assurance evidence with Governance for autonomous systems to be trusted at an affordable price.

D-RisQ SOFTWARE SYSTEMS

DSTL

The Defence Science and Technology Laboratory (Dstl) is the science inside UK defence and security.
Dstl is one of the principal government organisations dedicated to science and technology in the defence and security field.

Dstl supplies specialist services to MOD and wider government, working collaboratively with external partners in industry and academia worldwide, providing expert research, specialist advice and invaluable operational support.

We are innovative, collaborative and impactful.

[dstl] The Science Inside

Ethical Intelligence

- Ethical Intelligence (EI) makes ethics accessible and affordable for all by pioneering EaaS (Ethics as a Service) using its worldwide network of renowned interdisciplinary and verified ethics professionals.
- EI is particularly invested in the development of the verification tools and validation of professionals and systems. Through EI's experience in operationalising ethics in technology, the trends of private companies and governments favoring such tools and certifications has been prevalent, so having the ability to work with researchers and PhD students in these areas makes this an attractive proposition. Additionally, EI is keen to be involved in collaborating with the Centre's researchers, helping define the scope of the case studies, and facilitating access to data and use cases as this is seen to be greatly beneficial to not only the node but for the wider application of ethics and AI governance in general.

EI

Imandra

- Since its founding, Imandra has pioneered advances in AI for algorithm safety and compliance and their application to new industries. Their formal verification technology was originally designed to automatically analyse financial algorithms for glitches and unfair behaviour. Today, Imandra's "Reasoning as a Service ®" platform makes deep advances in symbolic AI for making algorithms safe, explainable and fair widely accessible to those without specialised background in these fields.

IMANDRA

Project Partners

Legal & General

Our purpose is to improve the lives of our customers, build a better society for the long term and create value for our shareholders. We use our long-term assets in an economically and socially useful way to benefit everyone in our communities.

Microsoft

- Microsoft was founded in 1975 with a mission to enable people and businesses throughout the world to realize their full potential by creating technology that transforms the way people work, play and communicate. We develop and market software, services, and hardware that deliver new opportunities, greater convenience and enhanced value to people's lives. We do business worldwide and have offices in more than 100 countries.
- Our products include operating systems for personal computers ("PCs"), servers, phones, and other intelligent devices; server applications for distributed computing environments; productivity applications; business solution applications; desktop and server management tools; software development tools; video games; and online advertising. We also design and sell hardware including the Xbox 360 gaming and entertainment console, Kinect for Xbox 360, Xbox 360 accessories, and Microsoft PC hardware products.

NASA Ames

NASA's Ames Research Center, one of ten NASA field centers, is located in the heart of California's Silicon Valley. Since 1939, Ames has led NASA in conducting world-class research and development in aeronautics, exploration technology and science aligned with the center's core capabilities.

National Physical Library

NPL is the UK's National Metrology Institute, developing and maintaining the national primary measurement standards. It is a Public Corporation owned by the Department of Business, Energy and Industrial Strategy (BEIS). It has a partnering agreement with BEIS and the University of Strathclyde and the University of Surrey. NPL is part of the National Measurement System (NMS) which provides the UK with a national measurement infrastructure and delivers the UK Measurement Strategy on behalf of BEIS.

We undertake excellent science and engineering to deliver extraordinary impact for the UK and provide the measurement capability that underpins the UK's prosperity and quality of life. From accelerating new antibiotics and more effective cancer treatments to developing unhackable quantum communications and superfast 5G, our expertise is crucial in researching, developing and testing new products and processes.

NVIDIA

We pioneered a supercharged form of computing loved by the most demanding computer users in the world -- scientists, designers, artists, and gamers. For them, we've built the equivalent of a time machine.

Fueled by the insatiable demand for better 3D graphics, and the massive scale of the gaming market, NVIDIA has evolved the GPU into a computer brain at the exciting intersection of virtual reality, high performance computing, and artificial intelligence.

Capgemini

Capgemini is a global leader in consulting, digital transformation, technology and engineering services. The Group is at the forefront of innovation to address the entire breadth of clients' opportunities in the evolving world of cloud, digital and platforms.

Project Partners

Optos

- Optos is a a division of Nikon Co. Ltd, Japan since 2015. Its products produce high resolution **opto**map images of 82% or 200◦ of the retina, something no other imaging device is capable of in a single comfortable capture. The most recent innovation from Optos integrates ultra-widefield retinal imaging and image-guided Optical Coherence Tomography (OCT). This combined device facilitates the early detection, management and effective treatment of disorders and diseases evidenced in the retina such as retinal detachments and tears, glaucoma, diabetic retinopathy, and age-related macular degeneration. More than 1,000 published and ongoing clinical trials as well as thousands of case studies and testimonials show the long-term value of **opto**map imaging and OCT in diagnosis, treatment planning, and patient engagement.
- Our interests lie in gaining a broader understanding of trust in the context of autonomous systems and in developing mechanisms to deliver trustworthy autonomous systems across a range of healthcare providers. Key to achieving these will be understanding the relationships of the stakeholders involved, which include first and foremost the patient, several types of practitioners and many regulatory bodies. Furthermore the combination of a specific healthcare context with a particular design of an autonomous system will have a significant impact on how trust can be instilled.

SICSA

The Scottish Informatics and Computer Science Alliance (SICSA) is a collaboration of 14 Scottish Universities. SICSA promotes international excellence in Unversity-led research, education and knowledge exchange for Scottish Informatics and Computer Science.

The SICSA Research Themes represent the key clusters of expertise within the SICSA Member Universities and facilitate interactions and collaborations between researchers across Scotland; as well as the exchange of ideas between Universities, businesses and the public sector.

Thales

- Thales is a global technology leader with more than 80,000 employees on five continents. The Group is investing in digital and “deep tech” innovations – Big Data, artificial intelligence, connectivity, cybersecurity and quantum technology – to build a future we can all trust. Trust is essential for societies to flourish, with humans playing a central role in every critical decision.
- Thales's high-tech solutions, services and products help companies, organisations and governments to achieve their goals and ambitions. And in each of our five vertical markets — digital identity and security, defence and security, aerospace, space, and transport — our customers play a vital role in society.

An overview of the TAS Hub

The UKRI Trustworthy Autonomous Systems Hub (TAS Hub) is led by the University of Southampton, with partners from the University of Nottingham and King's College London. The TAS Hub sits at the heart of the programme and assembles a diverse and multi-disciplinary team of researchers and industry partners, and government representatives.

The TAS Hub is a programme that adopts Responsible Research and Innovation (RRI) and Equality, Diversity, and Inclusivity (EDI) principles to deliver the best practices that will enable the development of autonomous systems that are both trustworthy in principle and trusted in practice.

The TAS Hub has so far initiated 18 community-driven projects worth around £3m, situated in four application areas: Autonomous Vehicles, Health and Care, Defence and Security, AI Ethics and Governance. It has also initiated its Creative Engagement programme through four artist-in-residence projects and a co-developed gallery at the National Gallery X. These projects were created by a diverse community of researchers from different disciplines (Computer Science, Engineering, Social Sciences, Law etc..) The Hub also brings together over 80 industrial partners, many of which are directly contributing to these projects and look to exploit their outputs over the next 3 years. We aim to build upon this momentum to address the most pressing challenges our society faces, while also looking to develop the UK's economic edge globally. To this end, our current call for pump priming projects highlights three key application areas: (i) The transition to net zero (ii) Post-pandemic recovery (iii) Inclusion. By so doing we look to ensure that autonomous systems are built for the purpose of improving quality of life and sustainability for all and not only those who can afford the technology.

We also endeavour to ensure that the public is aware of the new issues autonomous systems will create; how autonomous systems will affect their privacy and autonomy among others. Through our policy, events such as 'Trusting Machines? Cross-sector Lessons from Healthcare and Security' and public engagement initiatives, we are constantly contributing to government policy and directly talking to diverse stakeholders to internalise their concerns about future autonomous systems into our research. We will also be growing the impact of the Hub through our Skills programme that aims to translate research outputs into training and educational material that will shape the workforce across different disciplines and industries.

We look forward to continuing to work with the Governance and Regulation Node to drive the TAS programme forward.

Professor Gopal Ramchurn FIET
Director UKRI TAS Hub

Representative
Related Projects

Lachlan Urquhart

- Emotional AI in Smart Cities: Cross Cultural Lessons from UK and Japan on Designing for An Ethical Life. https://www.law.ed.ac.uk/research/research-projects/emotional-ai
- Horizon Digital Economy Institute: Trusted Data Driven Products https://www.law.ed.ac.uk/research/research-projects/horizon (which is also part of the TAS Hub) including hoRRIzon https://www.horizon.ac.uk/project/horrizon/

Andy Crabtree

- Trust Me? I'm an Autonomous Machine, TAS Pump Priming Project https://www.tas.ac.uk/current-research-projects/trust-me/
- Horizon Digital Economy Institute: Trusted Data Driven Products https://www.law.ed.ac.uk/research/research-projects/horizon

Ajitha Rajan

Royal Society Industrial Fellowship
The Fellowship will start in March 2022 and will allow Ajitha to pursue her research project on assessing the safety of artificial intelligence perception tasks within autonomous vehicles such as those that are responsible for detection of vehicles, pedestrians, lanes, traffic lights and obstacles. Ajitha will work closely with Codeplay developers to enhance and integrate the automated testing tool developed by her research group into a Codeplay product that will help assess safety of perception tasks.

Robin Williams

Optometry Pilot Study (OPS). Data Driven Innovation 'Building Back Better' Open Call Awards 2021.
Investigator: Robin Williams.
https://ddi.ac.uk/ddi-building-back-better-open-call-awards

Seminar Series

9 June 2021 – Lorenzo Strigini
How safe is this autonomous vehicle? Difficulties of probabilistic predictions and some ideas for improving trust in them.

23 June 2021 – Ekaterina Komendantskaya, Heriot Watt University
Analyse, search, reason, interpret, verify: the "Maslow Pyramid" of AI modelling.

14 July 2021 – Joel Fischer, University of Nottingham
What we talk about when we talk about Trustworthy Autonomous Systems.

28 July 2021 – Jack Stilgoe, UCL
How can we know a self-driving car is safe?

11 August 2021 – Helen Hastie, Heriot Watt University
Trustworthy Autonomous Systems – the Human Perspective.

25 August 2021 – Daniel Angelov, Efemarai
Extending the Industrial MLOps pipeline with the needs for robustness and continual improvement

8 September 2021 – Robin Bloomfield, Adelard
Assurance 2.0 and impact of AI/ML on regulation

22 September 2021 –John Downer, University of Bristol (TAS Functionality Node)
Overview of TAS Node in Functionality

13 October 2021 – Carina Prunkl, Oxford
Human Autonomy and the Governance of Autonomous Systems

27 October 2021, Jane Fenn, BAE Systems
Evolution in Standards and Regulation in the Aerospace Domain

8 December 2021– Marina Jirotka, Oxford
Responsible Robotics: What could possibly go wrong?

Publications

F.J.C. Garcia, S.C. Smith, J. Lopes, S. Ramamoorthy, H. Hastie, **Self-explainable robots in remote environments**, Videos and Demos track. In Proc. *ACM/IEEE International Conference on Human-Robot Interaction, 2021.*

S.D.S. Marín, D. Gomez-Vargas, N. Céspedes, M. Múnera, F. Roberti, P. Barria, S. Ramamoorthy, M. Becker, R. Carelli, C. A. Cifuentes, **Expectations and perceptions of healthcare professionals for robot deployment in hospital environments during the COVID-19 pandemic**, Research topic: Robotics, Autonomous Systems and AI for Nonurgent/Nonemergent Healthcare Delivery During and After the COVID-19 Pandemic, *Frontiers in Robotics and AI - Biomedical Robotics, 2021.*

M. Appelgren; A. Lascarides. **Symbol Grounding and Task Learning from Imperfect Corrections.** Second International Combined Workshop on Spatial Language Understanding and Grounded Communication for Robotics. ACL-IJNLP, 2021: Association for Computational Linguistics (ACL), 2021.

C. Marsden, I. Brown and M. Veale. **Responding to Disinformation: Ten Recommendations for Regulatory Action and Forebearance**, Chapter 11 in M. Moor and D. Tambini ed. Dealing with Digital Dominance, Oxford Press [2021]

A. McStay; L. Urquhart. **In cars (are we really safest of all?): Interior sensing and emotional opacity.**

D. Pagojus, A. Miller, B. Porr and I. Valkov. **Simulation and Model Checking for Close to Real-time Overtaking Planning.** In: M.Farrell & M. Luckcuck (Eds): Third Workshop on Formal Methods for Autonomous Systems (FMAS2021)

B. Schafer, C. Kutterer, E. Staudegger, E. Nerantzi, J. Slosser, J.J. Baker, M. Hildebrandt, R. Kennedy: Legal Services Industry in AIAI4People's 7 AI Global Frameworks, Atomium. 2021

Urquhart, L; Miranda, D. **Policing Faces: The Present and Future of Intelligent Facial Surveillance.**

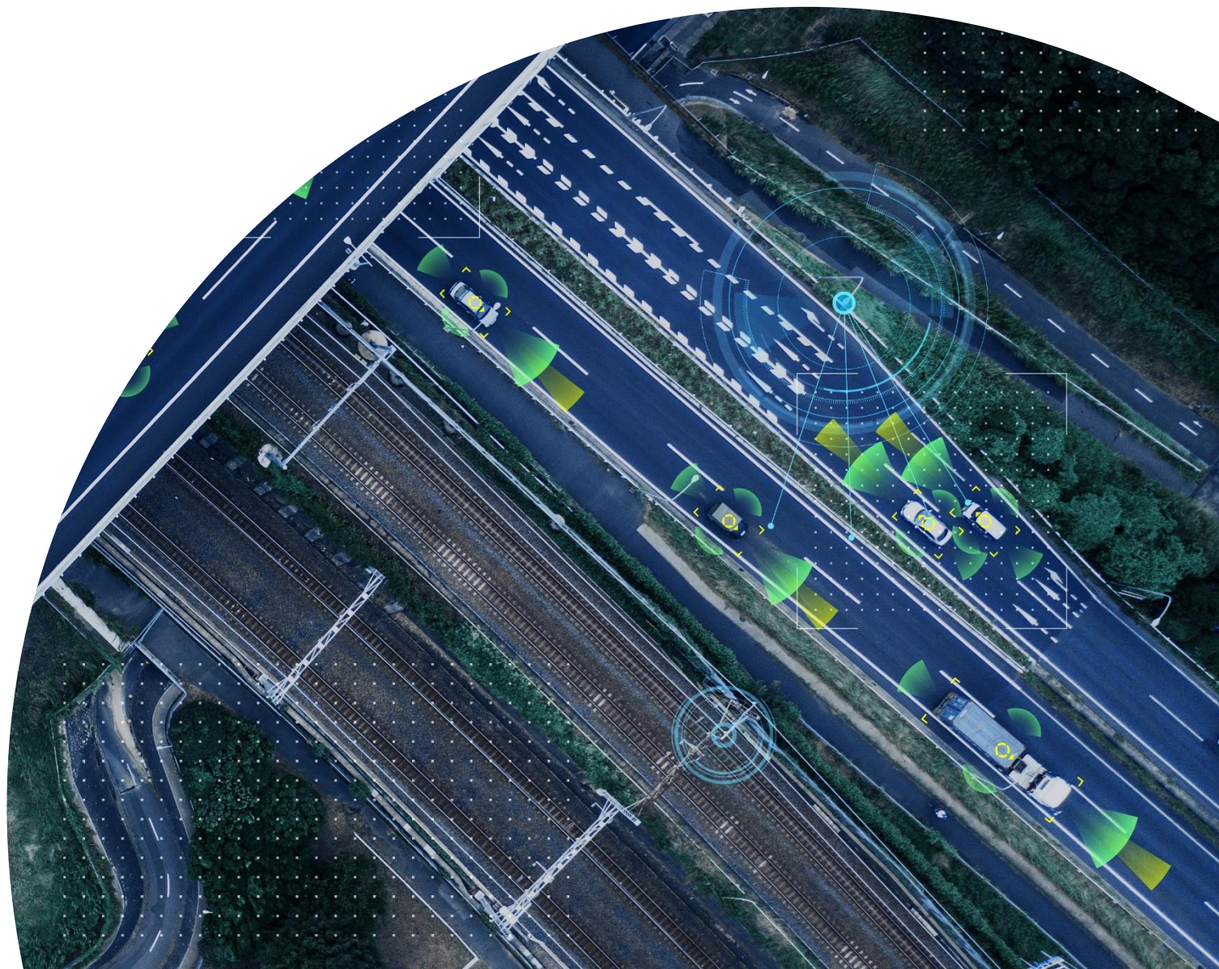
P. Inglesant; C. Ten Holter; M. Jirotka; R. Williams. **Asleep at the wheel? Responsible Innovation in quantum computing.**

M.L. Littman, I. Ajunwa, G. Berger, C. Boutilier, M. Currie, F. Doshi-Velez, G. Hadfield, M. C. Horowitz, C. Isbell, H. Kitano, K. Levy, T. Lyons, M. Mitchell, J. Shah, S. Sloman, S. Vallor, and T. **Walsh. Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report.** Stanford University, Stanford, CA, September 2021.

H. Chockler, D. Kroening, Y. Sun; **Explanations for Occluded Images.** Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021

H. Pouget, H. Chockler, Y. Sun, D. Kroening; **Ranking Policy Decisions** Proceedings of Advances in Neural Information Processing Systems (NeurIPS), 2021

V. Belle; I. Papantonis; **Principles and Practice of Explainable Machine Learning,** Frontiers in Big Data 2021.



https://governance.tas.ac.uk/

https://www.youtube.com/channel/UCTIkO9GeEJIPxIjhu9g7b7A

@tas_governance

tas.governance@ed.ac.uk