UKRI TAS NODE IN RESILIENCE

ANNUAL REPORT 2021

CONTENTS

Foreword

UK and international strategy documents promise a future where autonomous systems unobtrusively maintain the infrastructure of our cities, empathetically support the elderly in living fulfilled independent lives, and dependably manage our road traffic to minimise environmental impact. Fully observant of social, legal, ethical, empathic and cultural (SLEEC) norms, these autonomous systems operate resiliently: they avoid, withstand, recover from, adapt, and evolve to handle unforeseen uncertainty and disruptions.

Current autonomous systems cannot deliver the resilience levels needed to achieve this promise. Their resilience is severely limited by an inability to consider essential socio-technical concerns in their decision-making. Which resilience-enhancing actions are safe and compliant with the SLEEC norms that underpin even routine human decisions? How to quantify and reduce socio-technical uncertainty, and how to predict and detect disruptions? How to ensure that resilience is not achieved at the expense of other human values?

The TAS Node in Resilience project addresses these fundamental questions. Funded by a £3M EPSRC grant, our team of computer scientists, engineers, lawyers, mathematicians, philosophers and psychologists from five UK universities embarked on a 42-month research programme aimed at delivering a step change in the socio-technical resilience of autonomous systems.

Despite the challenges brought by the COVID-19 pandemic in 2021, the first year of the project has been highly successful. Together with our 30+ project partners, we co-defined assistive care, multimodal transportation, and emergency response case studies for the project. We have exploited the synergies between our disciplines to lay the basis for the development of a comprehensive toolbox of principles, methods, and systematic approaches for the engineering of resilient autonomous systems. We started new collaborations with researchers from the TAS Hub, the other TAS Nodes, and many other initiatives. We look forward to continuing to deliver our ambitious objectives, and to expanding these collaborations in 2022.

Prof. Radu Calinescu
TAS Node in Resilience
Principal Investigator

Creativity and Multidisciplinarity

Undertaking recognised world-leading fundamental research, with benefits to real-world applications and adoption of autonomous systems

Publications
20+ journal, conference and workshop papers

Use cases
Co-creation of use cases in robotic assistive dressing, multimodal transportation and emergency response with 30+ project partners

New multidisciplinary autonomous system grants

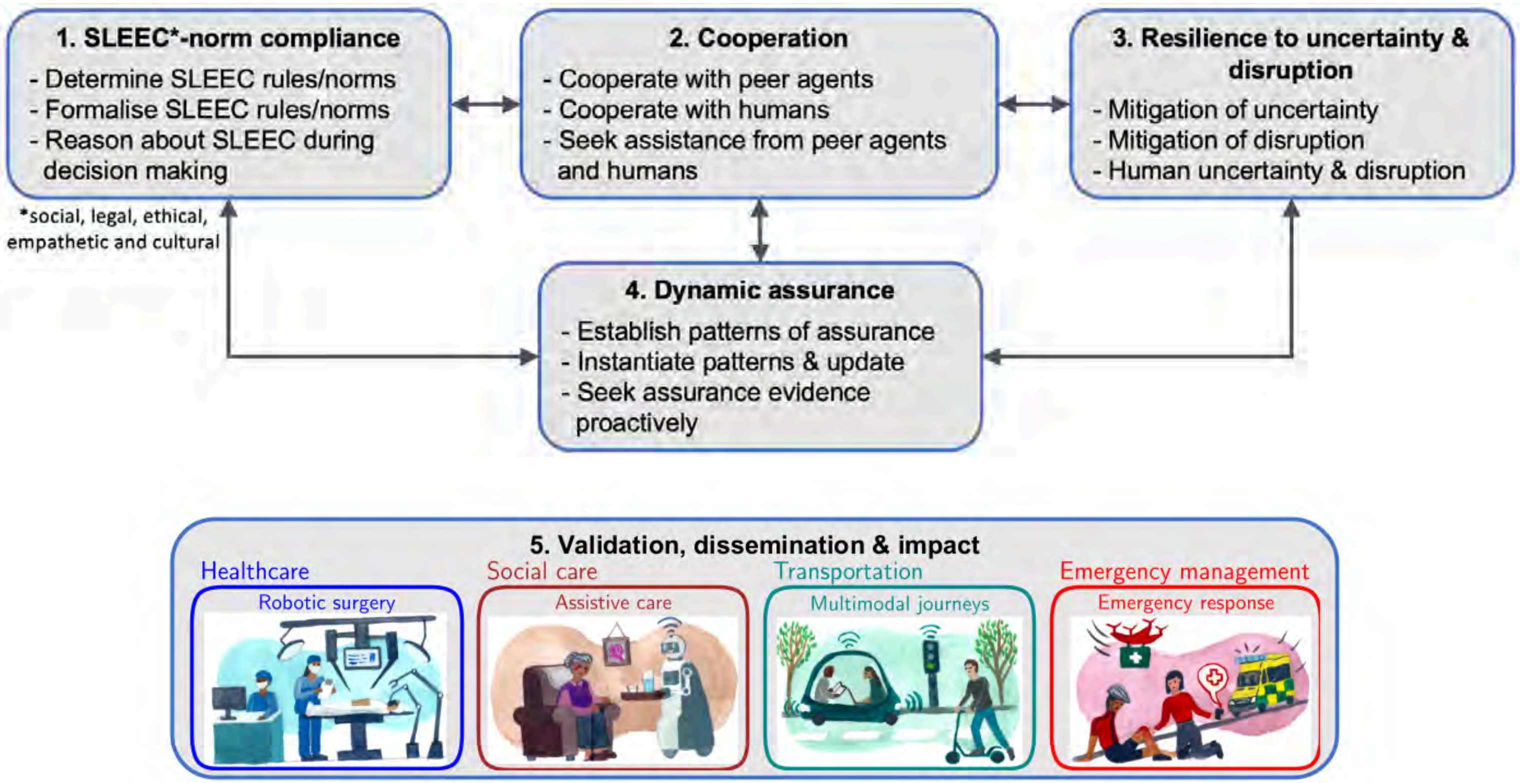
Project team members involved in three of 12 Round 1 TAS Hub pump-priming projects (£320,000)

1) Situational Awareness and Trust during Shift between Autonomy Levels in Automated Vehicles
(https://www.tas.ac.uk/current-research-projects/sa2ve/)
2) Imagining Robotic Care: Identifying Conflict and Confluence in Stakeholder Imaginaries of Autonomous Care Systems
(https://www.tas.ac.uk/current-research-projects/imagining-robotic-care/)
3) An Open Laboratories Programme for Trustworthy Autonomous Systems
(https://www.tas.ac.uk/current-research-projects/open-tas/)

Two Assuring Autonomy International Programme grants (£420,000)

1) Confident Safety Integration for Cobots – Phase 2
(https://cutt.ly/CSI-Cobot-2)
2) Ambient Assisted Living for Long-term Monitoring and Interaction
(https://cutt.ly/ALMI)

Fundamental research
The TAS Node in Resilience research strands and work being undertaken by researchers can be found on the following pages.



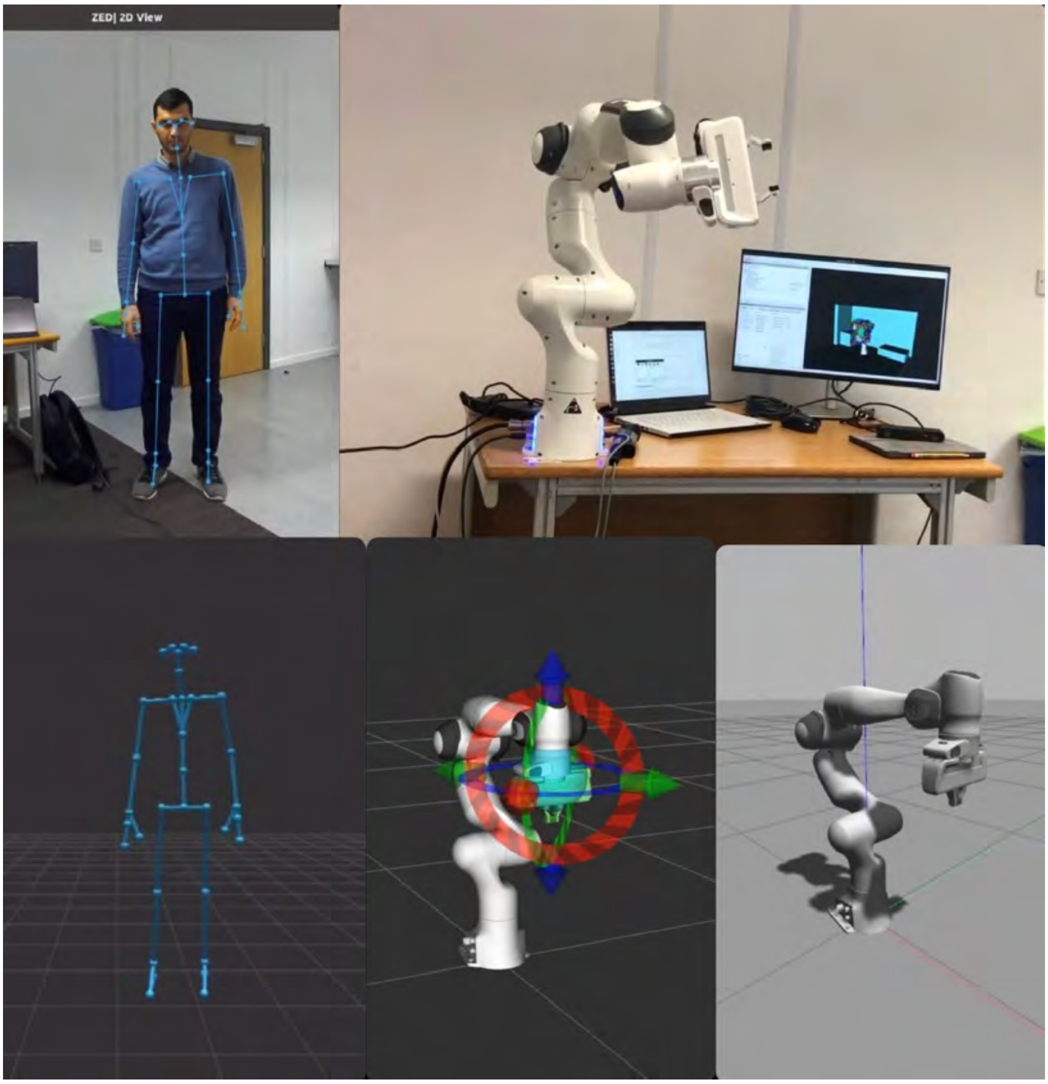
TAS Node in Resilience Research Strands

Dr. Mohammed Al-Nuaimi / The University of Sheffield

Summary of Research

Work has focused on an assistive dressing robot, represented by a robot arm (Franka Emika). The same robot arm in simulation has been developed using the Robot Operating System (ROS) and Gazebo simulator, and a control method has been designed using the MoveIt control system. An interface to control the real robot arm and its Gripper (in real-time) have also been developed.

Current work is aimed at developing the connection between an Omega 6 Force Dimension haptic device and the robot arm, so that the arm can be controlled remotely using the haptic device through ROS. Work is also taking place on a motion tracking system (skeleton tracking) using the latest generation ZED 2i stereo camera with Artificial Intelligence (AI) software.



Upcoming Work

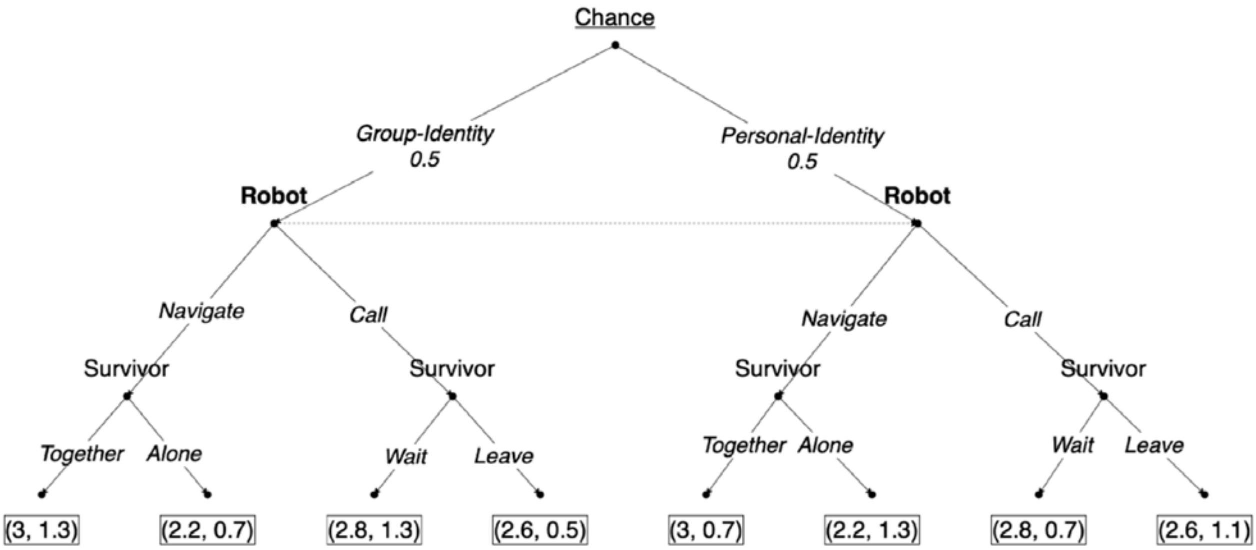
In the short term, physiotherapists from North Yorkshire council will join the team for a laboratory demonstration to collect necessary information for the dressing operation including the Force and Torque data and how to dress and support elderly people or people with upper limb impairments. A dressing handle with a Force/Torque sensor attached has been designed for this purpose and a set of IMUs will be used for tracking motion and to record the necessary data. This will contribute to a conference paper for IROS 2022. In the longer term, the plan is to: develop an AI verbal communication system for the robot to make simple conversation with the end users; use a second robot arm within the system for the dressing operation; and integrate the AI software into more functions of the dressing operations and controlling the system.

Dr. Carlos Gavidia-Calderon / The Open University

Summary of Research

Autonomous system resilience to human behaviour is being explored with the aim to engineer autonomous systems such as delivery drones, self-driving cars, and customer support robots that can infer human preferences and act accordingly, thereby increasing trust and utilization.

An evacuation scenario to test candidate solutions has been modelled with an evacuation robot traversing a disaster zone looking for survivors. Two possibilities exist: 1) call a first-responder or 2) guide the survivor to safety (preferable to alleviate the workload of first-responders). To make an optimal decision, the robot needs knowledge of the preferences of the surrounding survivors, and their willingness to help evacuate victims. The model uses game theory in conjunction with the social identity approach where agent types are mapped to social identities. When the robot finds a survivor, it: 1) uses sensors to gather identity markers; 2) infers probabilities over social identities from them; 3) feeds these values into a game-theoretic model; and 4) uses the model to obtain its optimal behaviour for the person it is helping. A preliminary evaluation using synthetic data compared the project's architecture against a proself-oriented robot and a prosocial-oriented one. The proself-oriented robot always calls for first-responder support when facing a victim, assuming survivors will not help with the evacuation, whereas the prosocial-oriented robot always offers guidance, relying on the help of other survivors in the area. The adaptive architecture outperforms the other two robots on mean number of successful evacuations.



.

Upcoming Work

A key component in the architecture is the type estimator which is responsible for translating sensor data into the probability distribution over types needed for game-theoretic modelling. Moving forward, the aim is to train this component using real-world data to enhance type estimation. The preliminary evaluation was made over the game-theoretic model but the aim is to evaluate the robot using a high-fidelity simulation (targeting Gazebo as a simulation platform).

Research Outputs

Gavidia-Calderon, C., Kordoni, A., Bennaceur, A., Nuseibeh, B., Levine, M. (2022). What do you want from me? Adapting Systems to the Uncertainty of Human Preferences. Submitted to ICSE NIER'22.

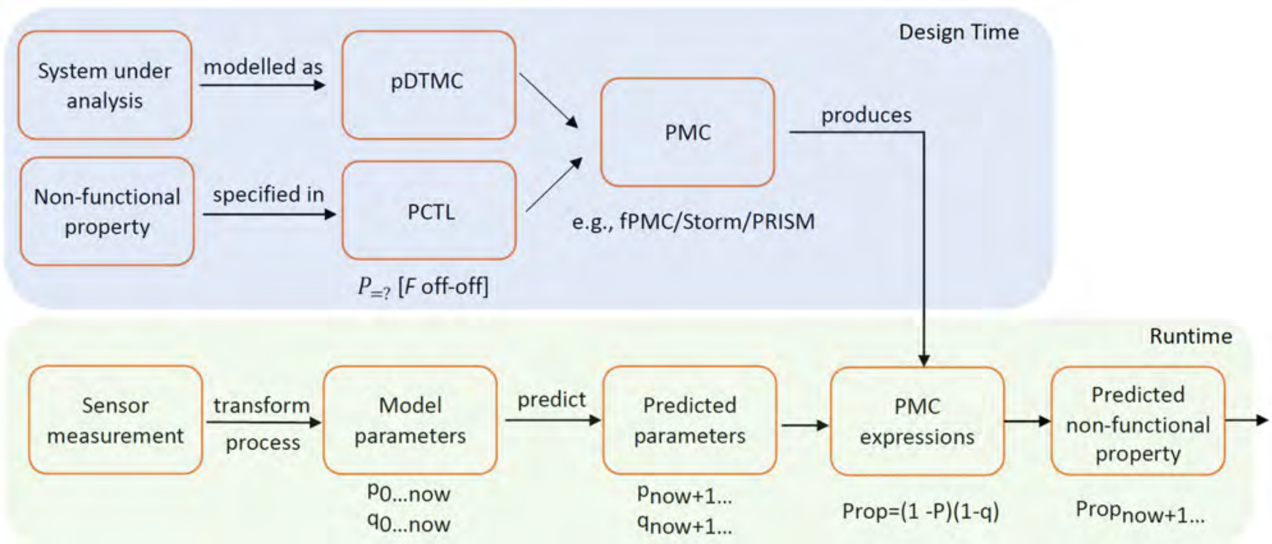
Gavidia-Calderon, C. (2021). Game-Theoretic View on Resilient Cyber-Physical Systems. TAS All Hands Meeting '21 Research Talks.

Dr. Xinwei Fang - University of York

Summary of Research

Predicting autonomous system disruptions through quantitative verification to enable mitigation strategy planning and resilience. Discrete-time Markov chains (DTMCs) were used to model the behaviour of the autonomous system under analysis, where every variable of the system and its environment is associated with a parameter in the DTMC (e.g. transition probability or reward), and the non-functional requirements of the system were specified in probabilistic computation tree logic (PCTL). With the parametric DTMC models and their associated properties, model checkers (e.g. fPMC, PRISM or Storm) were used to obtain the relationship between the system variables (i.e. the parameters of the DTMC) and the non-functional properties of interest. This relationship was then used to map predicted changes in the system and its environment to predicted changes in the system properties.

A model has been built for a robotic assistive-care application, and preliminary results show that the proposed method is able to predict system disruptions with accuracy.



Upcoming Work

In the short term, the evaluation of the current work will be finalised and published. In the long term, a framework that can be used to mitigate a wide range of disruptions (e.g. sudden changes, trends, noise) will be investigated.

Research Outputs

Gerasimou, S., Camara Moreno, J., Calinescu, R., Alasmari, N., Alhwikem, F., and Fang, X. (2021) "Evolutionary-guided synthesis of verified Pareto-optimal MDP policies." 36th IEEE/ACM International Conference on Automated Software Engineering.

Fang, X., Calinescu, R., Gerasimou, S., & Alhwikem, F. (2021). Fast Parametric Model Checking through Model Fragmentation. 43rd IEEE/ACM International Conference on Software Engineering.

Fang X. (2021). Fast parametric model checking for verifying autonomous system compliance with non-functional requirements TAS All Hands Meeting '21 Research Talks.

Dr. Anastasia Kordoni / Lancaster University

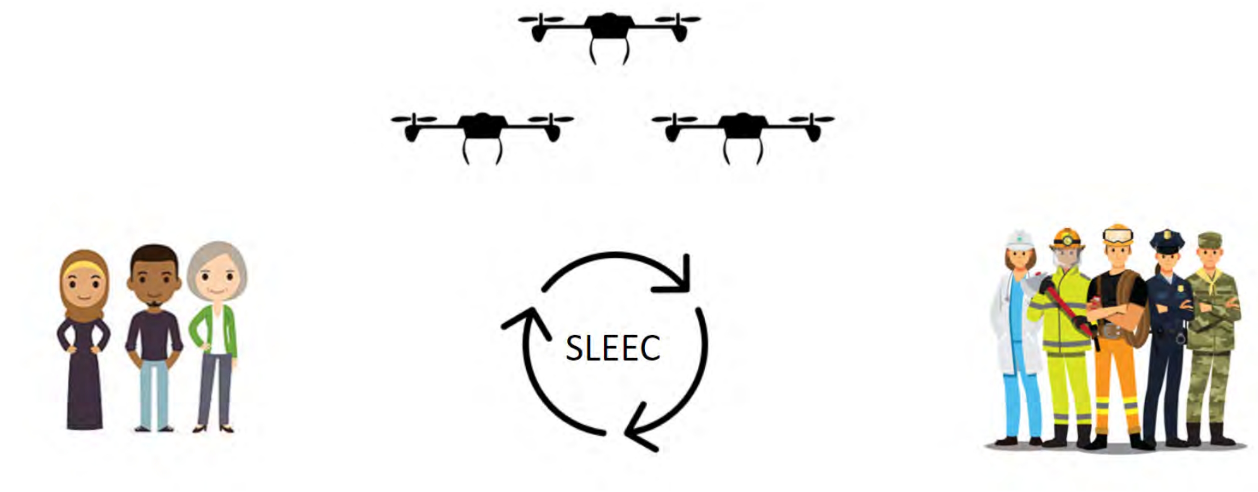
Summary of Research

The research focused on how social psychological theory, and specifically the social identity approach, can be used to identify and inform SLEEC rules and norms in the context of search and rescue operations (SaRs). SLEEC rules have been examined in three ways:

1) Accounting for multi-agency via interviews with emergency services personnel exploring human-UAV dynamics using the Critical Decision-Making method as well as interviews from the perspective of refugees who had experience with UAVs. This work is in collaboration with colleagues from the International Organization of Migration and NGOs.

2) Accounting for the context and stakeholders via interviews with fire-fighters who have used UAVs for non-UK emergency operations. Work with experts in International Humanitarian Law (IHL) to investigate the complexity of human laws regarding asylum seekers to inform changes in SLEEC rules and technical capacity. A Serious Lego Play workshop (SLP) on SLEEC rules and norms will be held with project stakeholders.

3) Accounting for psychological states and human-UAV cooperation via a simulation study examining how social identities create an autonomous system that adapts to human preferences to promote resilience.



Upcoming Work

Work will focus on completing interviews with emergency services and stakeholders followed by papers on SLEEC and the embedded dynamics. A report will be written for developers on existing and needed drone capacities across the emergency services, Another report will be prepared for emergency personnel on similarities, differences and recommendations for cooperation (with colleagues and UAVs) across the sector. A series of experiments will be designed to validate/formalise SLEEC rules and norms driven by the social identity approach and an identity for human-drone cooperation will be specified.

Research Outputs

Gavidia-Calderon, C., Kordoni, A., Bennaceur, A., Nuseibeh, B., Levine, M. (2022). What do you want from me? Adapting Systems to the Uncertainty of Human Preferences. Submitted to ICSE NIER'22.

Kordoni, A. (2021). Understanding Social Behaviour to Shape Surveillance. Presentation at Surveillance Conversations, Security Lancaster.

Kordoni, A. (2021). Humanitarian Emergency: SLEEC considerations / Use of UAV in Humanitarian Response. TAS All Hands Meeting '21 Research Talks.

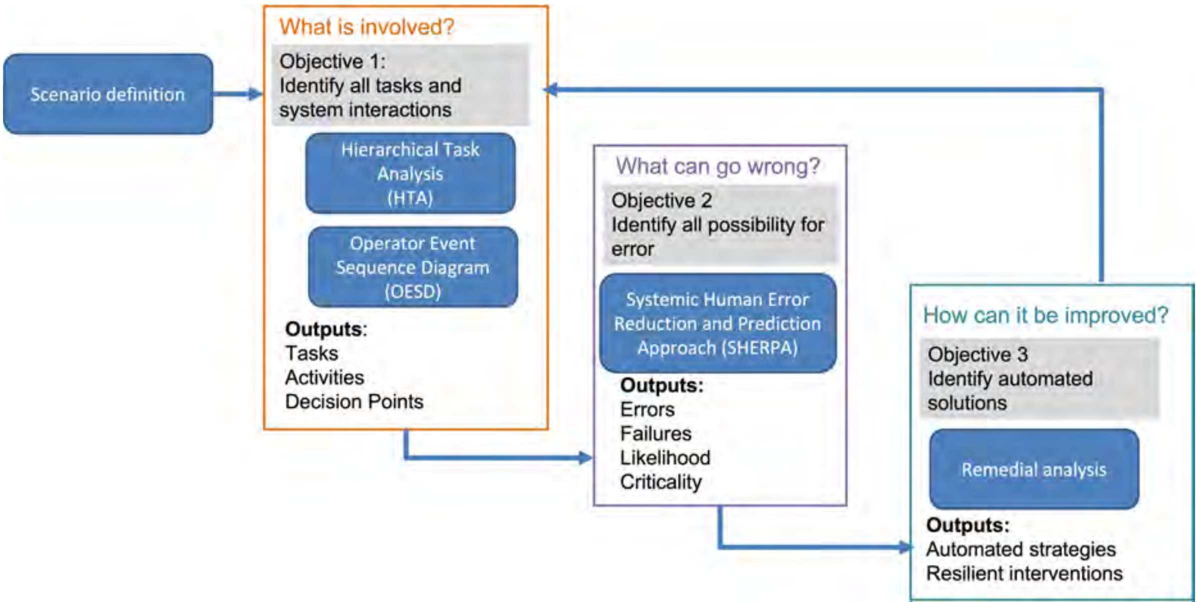
Dr. Katie Parnell / University of Southampton

Summary of Research

Literature has been reviewed on human factors engineering and transport scenarios generating a plan to capture the interaction between different road users in future road transport systems, wherein both autonomous vehicles and active travel modes (cyclists) will be prevalent, to enhance traffic efficiency and sustainability in a safe and resilient manner. Initial work on this scenario aims to provide a comparison between a 'non-automated' vehicle overtake scenario (Level 0 automated, SAE, 2021) and a Level 3 automated vehicle overtake (pending issues), which will still require the driver to have a supervisory role but not be fully in control of driving the vehicle.

A new approach to modelling system resilience using Operator Event Sequence Diagrams (OESDs) in combination with a predictive error analysis approach aims to map out all system interactions and the possibility for failure, while providing recommendations for the integration of automation to overcome current failures and enhance system resilience. Inclusion of Hollnagel's four pillars of resilience (monitor, anticipate, respond, learn) have also been applied to the analysis, with the proposition to include a 'detect' pillar. Detection plays a critical role in the safe interaction between vehicles and cyclists, especially in the automated scenario. This work has been written up into a first draft of a journal paper.

Research into the human factors methods that can be utilised to map out the SLEEC concerns proposed within this project has identified the value of the Cognitive Work Analysis methodology. This work is in the early stages and will progress alongside others working on mapping the SLEEC norms.



Upcoming Work

In the short term, the finalisation of the OESDs in combination with the error analysis work will be completed to provide a first draft of the paper. Further work will be done on the identification of the SLEEC norms within the transport scenario - including the application of methods used in the caregiving scenario and cognitive work analysis. In the longer term, there are plans for a user study that will build on the current work.

Research Outputs

Parnell, K. J., Banks, V. A., Allison, C. K., Plant, K. L., Beecroft, P., & Stanton, N. A. (2021). Designing flight deck applications: combining insight from end-users and ergonomists. Cognition, Technology & Work, 23(2), 353-365 (https://doi.org/10.1007/s10111-020-00629-w)

Parnell K.J. (2021). Trustworthy and Resilient Transport Systems. TAS All Hands Meeting '21 Research Talks.

Dr. Colin Paterson / University of York

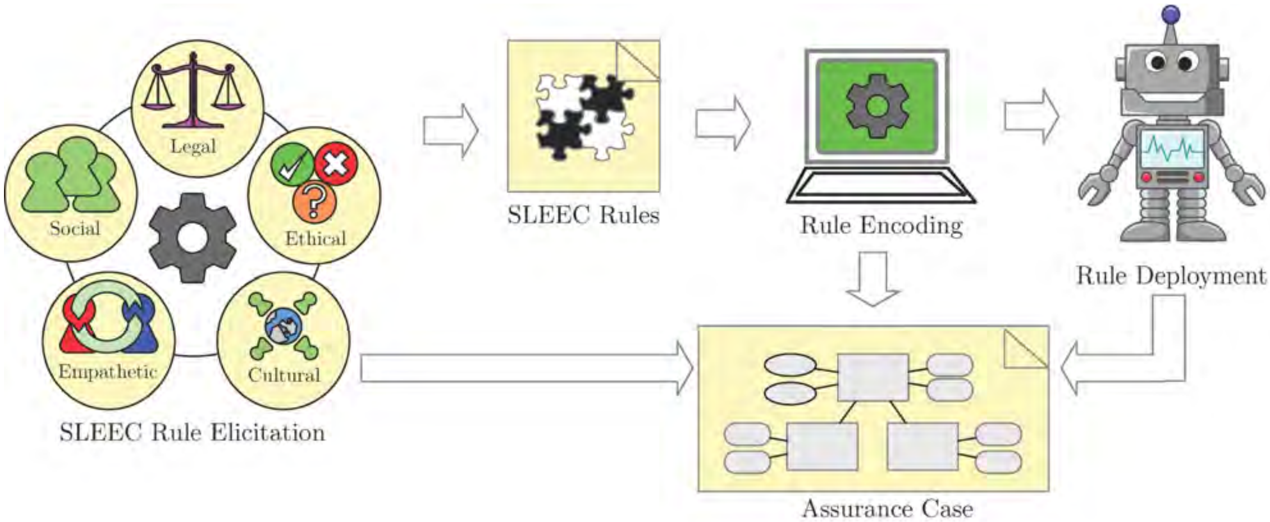
Summary of Research

Work has been focused on three tasks: 1) The derivation of a systematic process for the elicitation of SLEEC rules; 2) A formal language to allow for such rules to be encoded for verification; and 3) The assurance of SLEEC rules for autonomous systems.

Capabilities and measures which are critical to the implementation of rules are elicited from stakeholders. When a rule is not applicable under certain conditions (defeated), the potential impacts on SLEEC principles within the operating context should be recorded. These impacts will be critical in deciding trade-offs associated with rules as well as providing evidence used in the assurance of resilience SLEEC processes. Work has begun on encoding defined SLEEC rules in a structured language which allows for their verification. The language is intended to be interpretable by stakeholders and an example of how the final specification may look is given below:

```
WHEN UserFallen THEN SupportCalled SHOULD OCCUR WITHIN 60s
UNLESS UserCancelledSupportCall
UNLESS BloodPressure < MinBP
```

The assurance work is commencing to confirm that proposed patterns specified in the goal structuring notation (GSN) can be applied in a variety of contexts and for a range of SLEEC concerns.



Upcoming Work

The next stage of language formalisation is to link the language to modelling forms and verification techniques. Work on producing a first paper on the assurance of SLEEC concerns in autonomous systems is also underway.

Research Outputs

Colin Paterson, Haoze Wu, John Grese, Radu Calinescu, Corina S. Pasareanu and Clark Barrett (2021). DeepCert: Verification of Contextually Relevant Robustness for Neural Network Image Classifiers. 40th International Conference on Computer Safety, Reliability and Security, SAFECOMP.

Colin Paterson, Radu Calinescu and Chiara Picardi (2021). Detection and Mitigation of Rare Subclasses in Neural Network Classifiers. 3rd IEEE International Conference on Artificial Intelligence Testing.

Paterson, C. (2021). DeepCert: Verification of Contextually Relevant Robustness for Neural Network Image Classifiers. TAS All Hands Meeting '21 Research Talks.

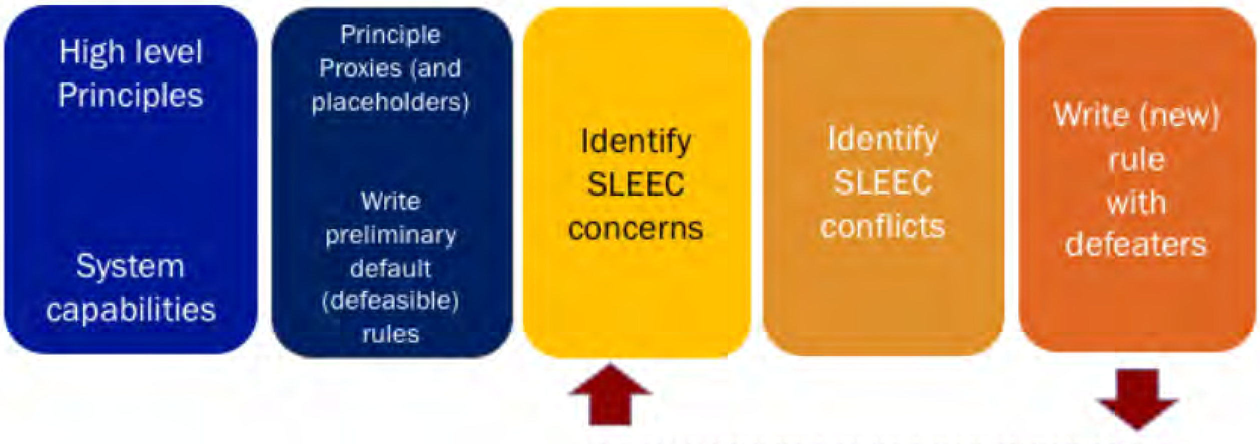
Dr. Beverley Townsend / University of York

Summary of Research
Deriving rules from SLEEC norms considers the importance of understanding and addressing SLEEC norms and normative concerns in the development and deployment of trustworthy resilient autonomous systems. 'Concerns' are those points of impact within a system that directly affect SLEEC norms. The aim of this research is to show how one might move from a high-level set of norms for an autonomous system to explicit, encoded specification rules.

The three central intentions of the research are: to uncover the high-level normative principles that underpin trustworthy systems; to defend a rule elicitation process that may be conducive to their translation, application, and integration; and lastly, to pragmatically demonstrate how this might be achieved using a robotic assistive-care dressing system as an example.

The process framework has been established and the documentation is being completed. Rules, set up as defeasible default rules with exceptions (or hedges), are being developed as examples of how the process may be practically implemented. This process in the context of a transport case scenario is also under consideration.

How normative 'conflicts' or tensions arising from the various norms (and concerns) may be resolved - either through weighing and balancing of the norms or through a method of Hortian default reasoning, prioritisation, and hedging is also being investigated.



Upcoming Work
Work is underway to produce a research paper on the SLEEC rule elicitation process (Deriving rules from SLEEC norms); to contribute to a SLEEC assurance case paper; and to publish a paper on facilitating normative decision-making in autonomous systems.

Research Outputs
Townsend BA (2021). Decoding the proposed EU AI Act. American Society of International Law. - (https://www.asil.org/insights/volume/25/issue/20)

Townsend BA (2021) SLEEC rule elicitation process. TAS All Hands Meeting '21 Research Talks.

Coordination and Collaboration

Building a connected and multidisciplinary UK research community tackling the challenges of Trustworthy Autonomous Systems

January 2021

The project welcomed over 25 collaborators from 19 partner organisations including autonomous system developers, operators, end users, researchers, regulators, policy makers, at our first partner workshop. Partner presentations and break-out discussion groups provided a unique opportunity to learn about the views and activities of autonomous systems stakeholders, to identify obstacles to the socio-technical resilience of autonomous systems, and to obtain essential partner input into the early activities of our project.

February 2021

Joint workshop with the TAS Node in Verifiability and ongoing collaborative work on TAS uncertainty and ethical concerns

Discussions with TAS Hub and Assuring Autonomy International Programme (AAIP) to collaborate on joint calls, events, projects, TAS contributions to AAIP Body of Knowledge, etc.

February / March 2021

Discussions with TAS Programme and UK-RAS Network and arrangements for a TAS White Paper to be published in UK-RAS series.

March 2021

Joint 'probabilistic verification' workshop with the TAS Node in Functionality and ongoing collaborative work on modelling and verification of robotic swarms

April 2021

European Robotics Forum 'Needs and Challenges for Assuring the Safety of AI' workshop – joint presentation with AAIP.

June/July 2021

Four presentations given by or involving our researchers at the 'Trusting Machines? Cross-sector Lessons from Healthcare and Security' TAS/RUSI Conference, which brought together academic experts, policy leaders, industry professionals and the public to discuss a future where autonomous machines integrate into two of our vital sectors.

September 2021

At the TAS Hub All-hands meeting (AHM), our project researchers contributed to the 'TAS Leads Fireside Chat', presented 7 of the 12 pre-recorded TAS AHM research talks (https://vimeo.com/showcase/8775391), organised a 'Handling Uncertainty in Autonomous Systems' panel and participated in additional AHM sessions.

September 2021

Engagement and agreed collaboration plan with leading AI researchers from the USA at Carnegie Mellon University/NASA (Dr Corina Pasareanu), the University of Maryland (Prof John Horty) and Rowan University (Prof Nidhal Carla Bouaynaya)

September / October 2021

Collaboration with the 'Imagining robotic care' TAS pump-priming project to plan and test co-design workshops for project stakeholders using Serious LEGO Play (SLP)

November 2021

Meeting with group of TAS Hub researchers led by Prof. Gopal Ramchurn to plan further collaborations and a 2022 'Autonomous Systems for Health and Social Care' workshop.

Advocacy and Engagement

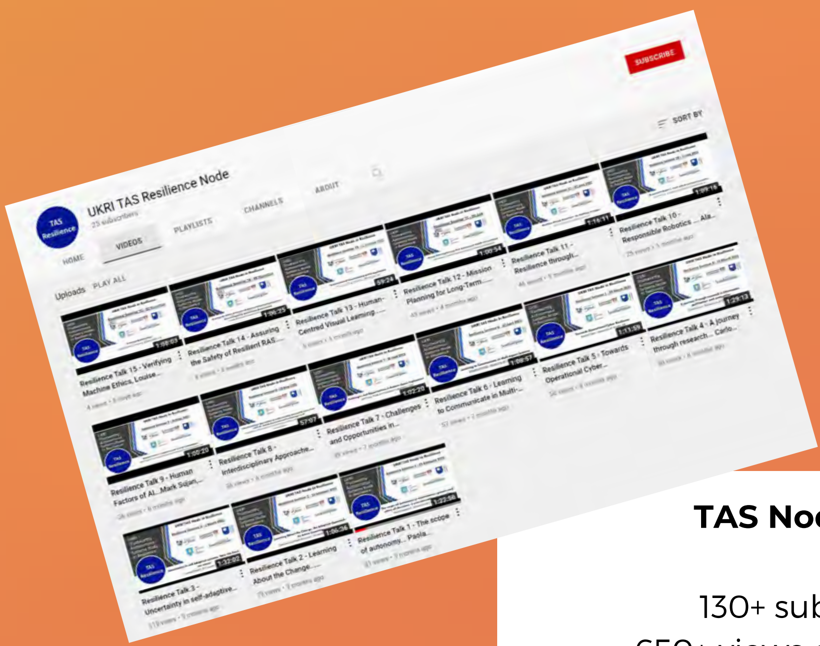
Putting in place a clear single point of contact for TAS expertise and engagement with key takeholders

Contributions to Standards and Guidance

IEEE Standards Association
P2817 Guide for Verification of Autonomous Systems
https://standards.ieee.org/project/2817.html

BSI Standards Development
DS1 - Dependability
https://standardsdevelopment.bsigroup.com/committees/50001626

Assuring Autonomy International Programme
Guidance on the Assurance of Machine Learning in Autonomous Systems (AMLAS)
https://www.york.ac.uk/assuring-autonomy/guidance/amlas/



TAS Node in Resilience Seminar series

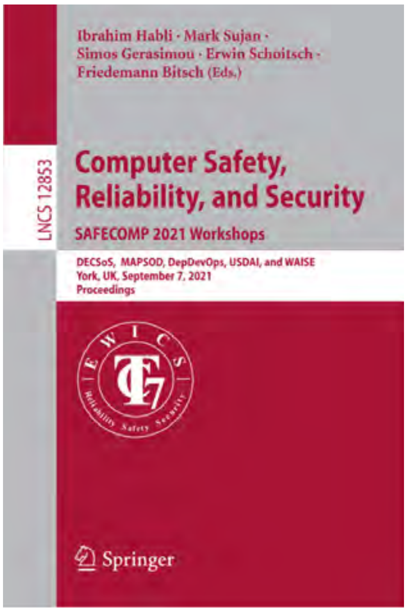
130+ subscribers from outside the project
650+ views of recorded seminars on YouTube
15 speakers from academia and industry
https://resilience.tas.ac.uk/events

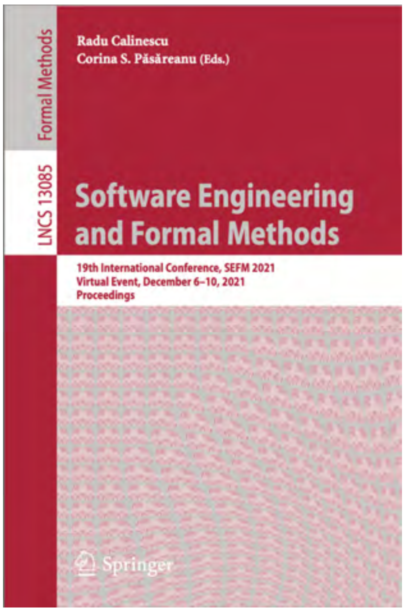
Influencing international TAS research agenda

Co-organisation of international conferences

40th International Conference on Computer Safety, Reliability and Security (SAFECOMP 2021) 7-10 September 2021.
Special theme: Safe human-robotic and autonomous system (RAS) interaction

19th International Conference on Software Engineering and Formal Methods (SEFM 2021) 6-10 December 2021.
Special topic: Software Engineering and Formal Methods for Resilient and Trustworthy Autonomous Systems





Guest editing journal special issues

Computing (Springer) Special Issue on Resilient Software and Software-controlled Systems, January 2021

Journal of Systems and Software (Elsevier) Special Issue on Software Engineering for Trustworthy Cyber-Physical Systems, April 2021

Looking Ahead

Global Research and Innovation Programme (GRIP)
January-March 2022
Collaboration activities and visits between TAS Resilience Node and US researchers, under the US/UK Statement of Intent (SoI) on Artificial Intelligence R&D.

- Rowan University
- University of Maryland
- Carnegie Mellon University

Co-creation Stakeholder Workshops - January/February 2022
Exploring Resilience via Serious LEGO Play.

- Examine how social, legal, ethical, empathic, cultural (SLEEC) values are built into autonomous systems for different groups (Industry/ Govt/ 3rd Party) and their impact on producing resilient, trustworthy autonomous systems
- In-depth analysis to map out SLEEC norms for the TAS Resilience Node, feedback to stakeholders, and produce paper for publication

UKRI TAS Resilience Thought Piece Workshop - January 2022
Workshop to contribute to producing a white paper or 'biscuit book' style publication.

- Key communications piece with policy makers and industry
- To be attended by the Node PI, Co-Is and other TAS researchers who can contribute to the discussion, and by representatives from the TAS Hub and from Thales to provide an industrial practitioner view.

NII Shonan Meeting - June / July 2022
No.168 Uncertainty in Self-Adaptive Systems (USAS)

- Focus on the central role of uncertainty in self-adaptive systems.
- Bring together leading researchers from multiple areas active in self-adaptation, foster collaboration, define a joint understanding of uncertainty and outline a joint research roadmap for future research on the topic.
- Edit a research volume based on the findings of the meeting.

LEARN MORE

Website:

https://resilience.tas.ac.uk/

Twitter:

@tas_resilience

YouTube Channel:

https://cutt.ly/resilience-channel

Email:

radu.calinescu@york.ac.uk

UKRI TAS NODE IN
RESILIENCE

This work is supported by the UK Research and Innovation (UKRI)
Trustworthy Autonomous Systems (TAS) programme
[EPSRC Ref: EP/V026747/1]