SLEEC & Resilience: Doing Lego Serious Play with TAS Stakeholders

Lancaster University| Open University| University of Sheffield
a.kordoni@lancaster.ac.uk | mark.levine@lancaster.ac.uk

Anastasia Kordoni, Mark Levine, Stevienna de Saille,
Amel Bennaceur, Carlos Gavidia-Calderon, Bashar Nuseibeh

Introduction

While "*resilience*" can refer to an autonomous system, individual capabilities, or a human-robot team as a whole (Matthews et al., 2016), there are often differences in the *meaning* and *content* of this term.

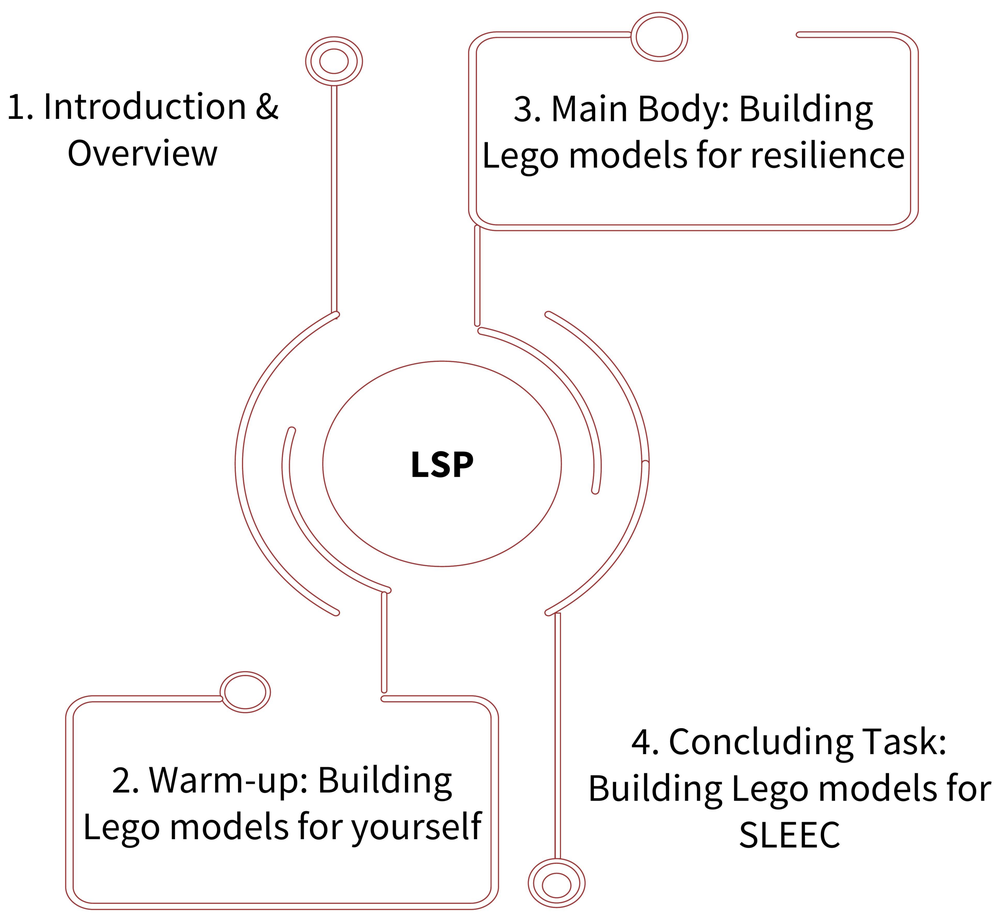
RQ1: How do different stakeholders define "*resilience*" based on their expertise?

RQ2: What Social, Legal, Ethical, Empathic, Cultural (SLEEC) values may be relevant to designing resilient autonomous systems?

Methodology

Participants:
2 Groups of TAS stakeholders
Group 1: *N* = 5 (industry, government)
Group 2: *N* = 7 (industry, government, third party)

Procedure:
Online **Lego Serious Play** workshop (LSP , Frick et al., 2013)



RQ1. Resilience is…

- **A process of recovering** from failure. Building a simpler but more robust model in case the complex model fails.
- The **interaction** between human-robot and the **process** of getting them into a good state.
- A **bridge of trust** that the system does what it says in a reliable way and what it wants us to do.
- **Diversity of perception** & **flexibility.**
- **Self-adaption** to unforeseen circumstances.
- **Having a vision** and maintaining it throughout the designing process.

RQ2. SLEEC Values

Social
- Respect equality
- Account for social biases
- Account for public readiness
- Build a shared understanding of reality

Legal
- Explainability
- Consistency in performance
- Conformance to standards
- Use multi-disciplinary research to inform policy/legislation

Ethical
- Fairness & safety
- Understand conflicts & dilemmas
- Do no harm
- Design for human
- Don't systematise inequalities
- Be transparent
- Promote social responsibility

Empathic
- Flexible explanations according to human situation
- Relatable communication
- Subdivide to separate goals

Cultural
- Protect the environment
- Inclusivity
- Cultural appropriateness of system operation
- Cultural determination of what can/cannot be accepted

Conclusions & Future Directions

Conclusions:

- Resilience in autonomous systems is associated with reasoning about i) what we have **implicitly taken out** of the design, ii) what needs to be built in to deal with **novel situations**, ii) **diversifying** routes to recovery.
- Respect for social equality, minimising the creation of new inequalities and the systematisation of existing ones, benefit humanity, and reflect on the human interpretation of trust to build relationships with machines are among the SLEEC values that were prioritised for this resilience.

Future Directions:

- Where does resilience lie as an end responsibility?
- How do we deal with a lack of distinction between how we approach a problem and how it's perceived, understood or interacted with by the technology?