Trust Me?… I'm an Autonomous Machine
Identifying Expert Narratives Around Trustworthy Autonomous Systems

Lancaster University | The University of Nottingham

Workshop Organisers: **David Green, Joseph Lindley, Glenn McGarry**
Artwork and Sketch Noting: **Miriam Sturdee**
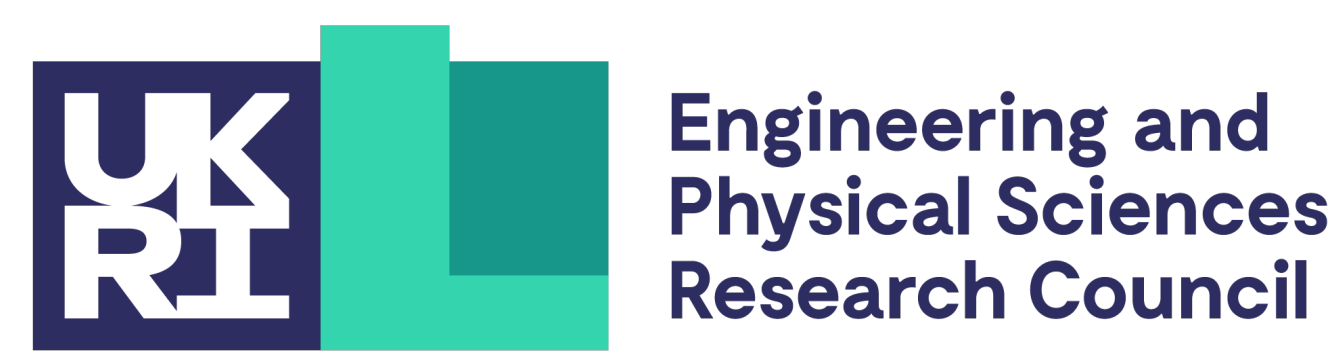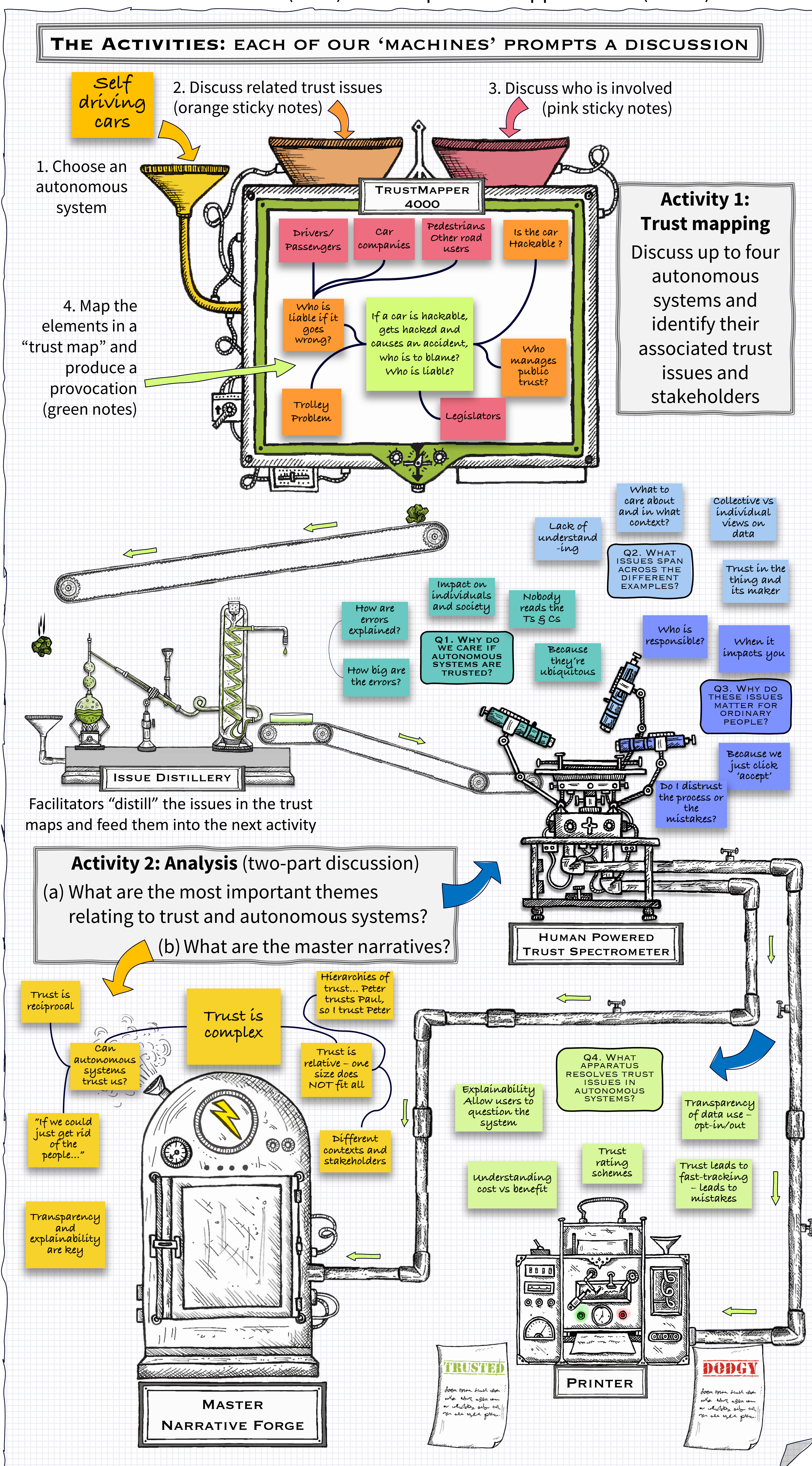Facilitation Support: **Zach Mason**
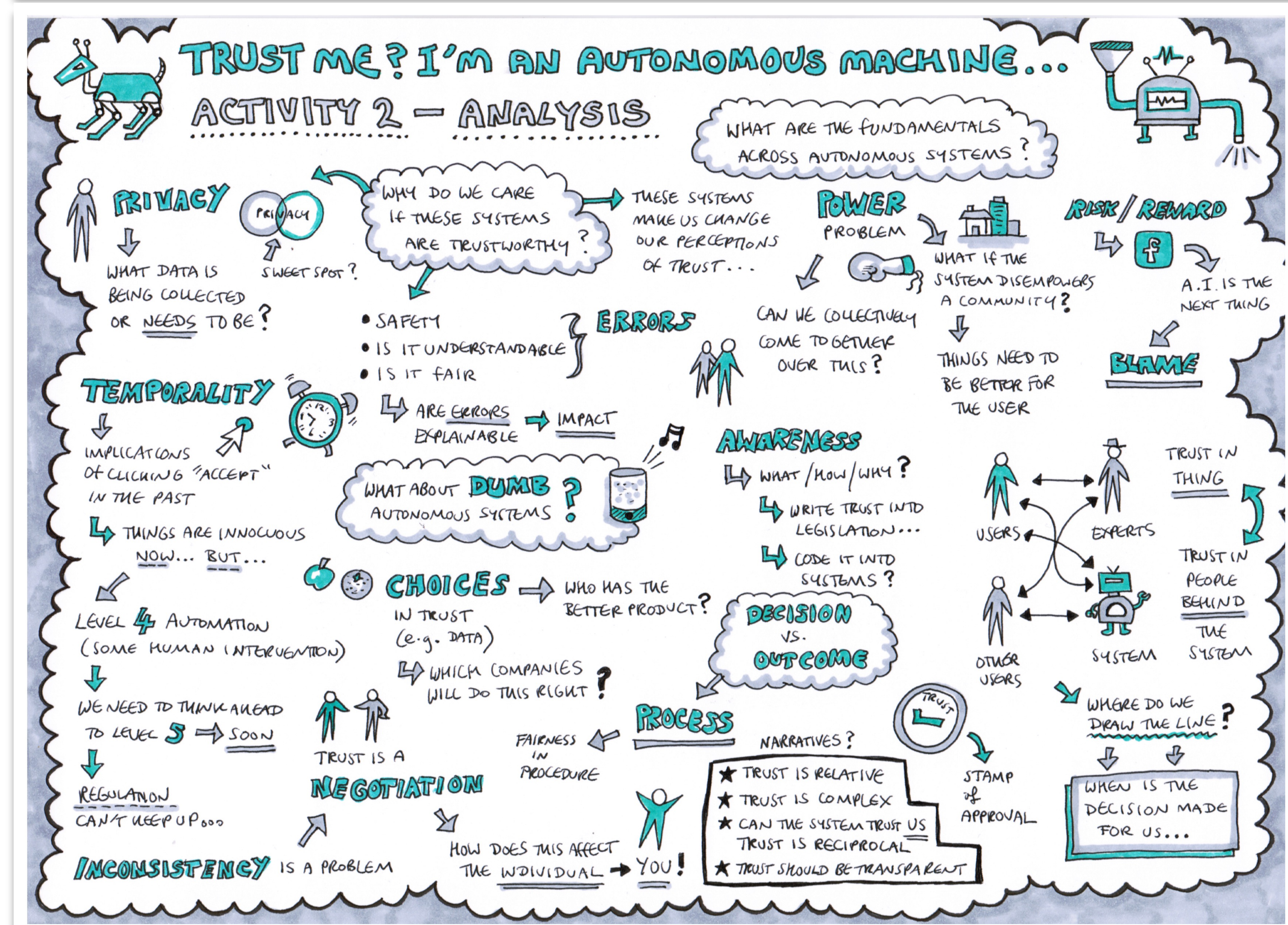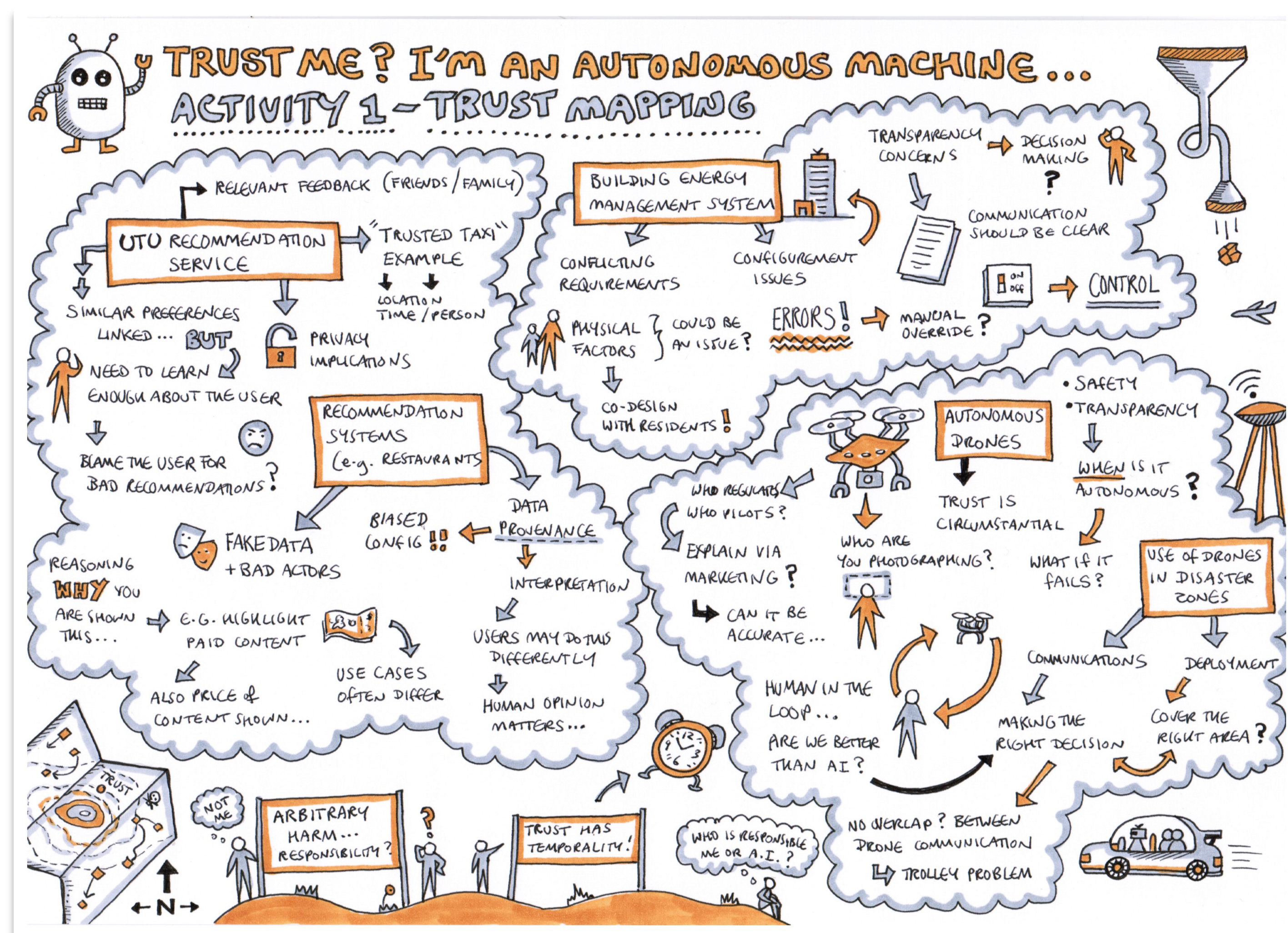
The Workshop

Aims
- To identify key themes or 'master narratives' in TAS
- To make sure we are clear about the shape of these narratives, as a first step towards bridging these understandings with the general public.

Method
- Our workshop coordinated and captured a focused discussion about TAS with domain experts (17 in academia + 5 in industry).
- Activities and discussion were designed and facilitated through online collaborative whiteboard (Miro) and telepresence applications (Zoom).

THE ACTIVITIES: EACH OF OUR 'MACHINES' PROMPTS A DISCUSSION

Self driving cars

2. Discuss related trust issues (orange sticky notes)

3. Discuss who is involved (pink sticky notes)

1. Choose an autonomous system

4. Map the elements in a "trust map" and produce a provocation (green notes)

TrustMapper 4000

Drivers/Passengers — Car companies — Pedestrians Other road users — Is the car Hackable?

Who is liable if it goes wrong? — If a car is hackable, gets hacked and causes an accident, who is to blame? Who is liable. — Who manages public trust?

Trolley Problem — Legislators

Activity 1: Trust mapping
Discuss up to four autonomous systems and identify their associated trust issues and stakeholders

ISSUE DISTILLERY

Facilitators "distil" the issues in the trust maps and feed them into the next activity

What to care about and in what context? — Collective vs individual views on data — Lack of understanding — Q2. WHAT ISSUES SPAN ACROSS THE DIFFERENT EXAMPLES? — Trust in the thing and its maker — Impact on individuals and society — Nobody reads the Ts & Cs — How are errors explained? — Q1. WHY DO WE CARE IF AUTONOMOUS SYSTEMS ARE TRUSTED? — Because they're ubiquitous — Who is responsible? — When it impacts you — How big are the errors? — Q3. WHY DO THESE ISSUES MATTER FOR ORDINARY PEOPLE? — Because we just click 'accept' — Do I distrust the process or the mistakes?

HUMAN POWERED TRUST SPECTROMETER

Activity 2: Analysis (two-part discussion)
(a) What are the most important themes relating to trust and autonomous systems?
(b) What are the master narratives?

Trust is reciprocal — Trust is complex — Hierarchies of trust... Peter trusts Paul, so I trust Peter — Can autonomous systems trust us? — Trust is relative – one size does NOT fit all — "If we could just get rid of the people..." — Different contexts and stakeholders — Transparency and explainability are key — Explainability Allow users to question the system — Q4. WHAT APPARATUS RESOLVES TRUST ISSUES IN AUTONOMOUS SYSTEMS? — Transparency of data use – opt-in/out — Understanding cost vs benefit — Trust rating schemes — Trust leads to fast-tracking – leads to mistakes

TRUSTED — PRINTER — DODGY

MASTER NARRATIVE FORGE

Capturing the Narratives

Trust Me? I'm an Autonomous Machine… Activity 1 – Trust Mapping

Trust Me? I'm an Autonomous Machine… Activity 2 – Analysis

What next?

Post Workshop Analysis
- Theme development
- Share report with TAS members

Develop Speculative Designs and Consult with Publics
- Develop speculations which distil, embody, and communicate key positions, ideas and issues
- Deliver these to public groups in collaboration with Ruskin and Design Museums

Reflections and Reporting
- Construct 'master narratives' which negotiate between expert and lay positions
- Share and report to TAS members

Preliminary findings

- Multiple hierarchies, contexts, stakeholders and technologies always interact in any given circumstance
- Trust changes over time
- Just because Alice trusts Bob, and Peter trusts Bob, doesn't mean that Peter should or would trust Alice

Trust is a Distributed Concern