# The Alan Turing Institute

# A Relational Approach to Human-AI Trust Building

Dr Rupert Barrett-Taylor

# Summary

– **Trust is a Relational Problem,**

– Trust is a label often overloaded and lacking specificity,

– Artificial Intelligence (AI) is a label for technology also lacking specificity in research,

– Survey of literature suggests trust related to technology is treated instrumentally,

– The trustor, often meaning the user, lacks agency in theories of trust relations,

– Trust, like AI, is a question of epistemology as well as technics.

# Constructing Trust

– Epistemology

– Human-Machine Trust

– A Framework for Analysis


TRUST NO ONE

## **Trust**

– Makes us vulnerable

– Prospect of betrayal

– Is inherently risky

– Generated between equals

# Who Trusts What?

– Trustor: 'Users'

– Trustee: 'Machine'



On the one hand we tend to overestimate it and imagine it as having superpowers,
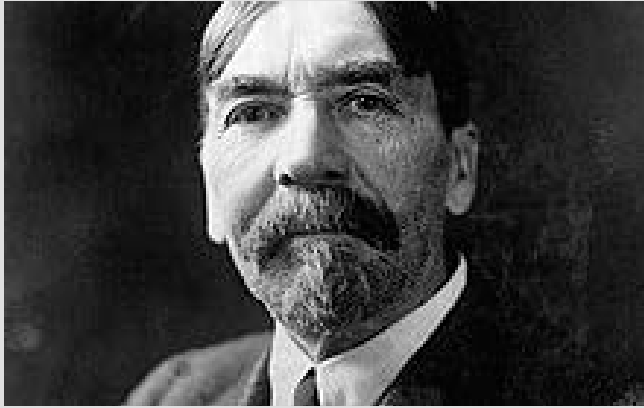
# Human-Machine Trust

– **Reliance**



– **Trust**

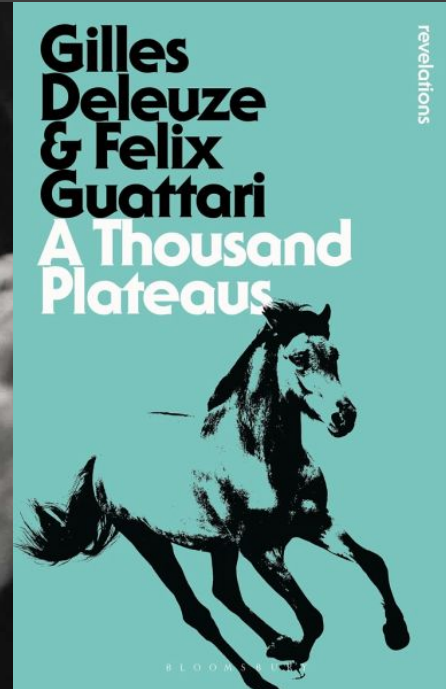# Epistemology: Frameworks for Understanding Trust in Technology

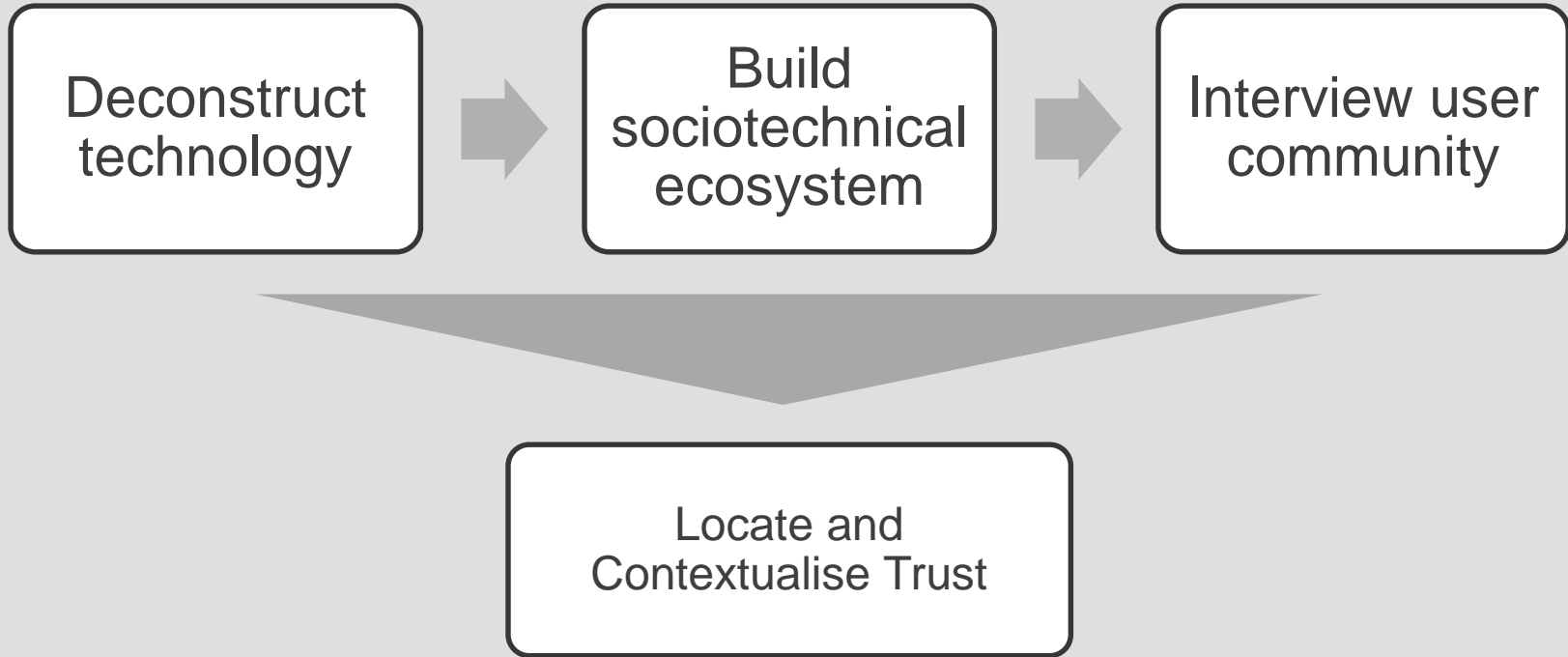– **Determinism**



– **Positivism**

# Relational Approaches

– Actor Network Theory

– Assemblage Theory

# Empirical Analysis

Deconstruct technology → Build sociotechnical ecosystem → Interview user community

⬇

Locate and Contextualise Trust

## Outcomes

– Guidance Notes

– Archive User Experiences of Trust & Reliance

– Illuminate Distrust & Scepticism

– Generalise Experiences

– Organisational Change

# Questions


I HAVE BEEN A GOOD CHATBOT!