ENGINEERING TRUSTABLE AI

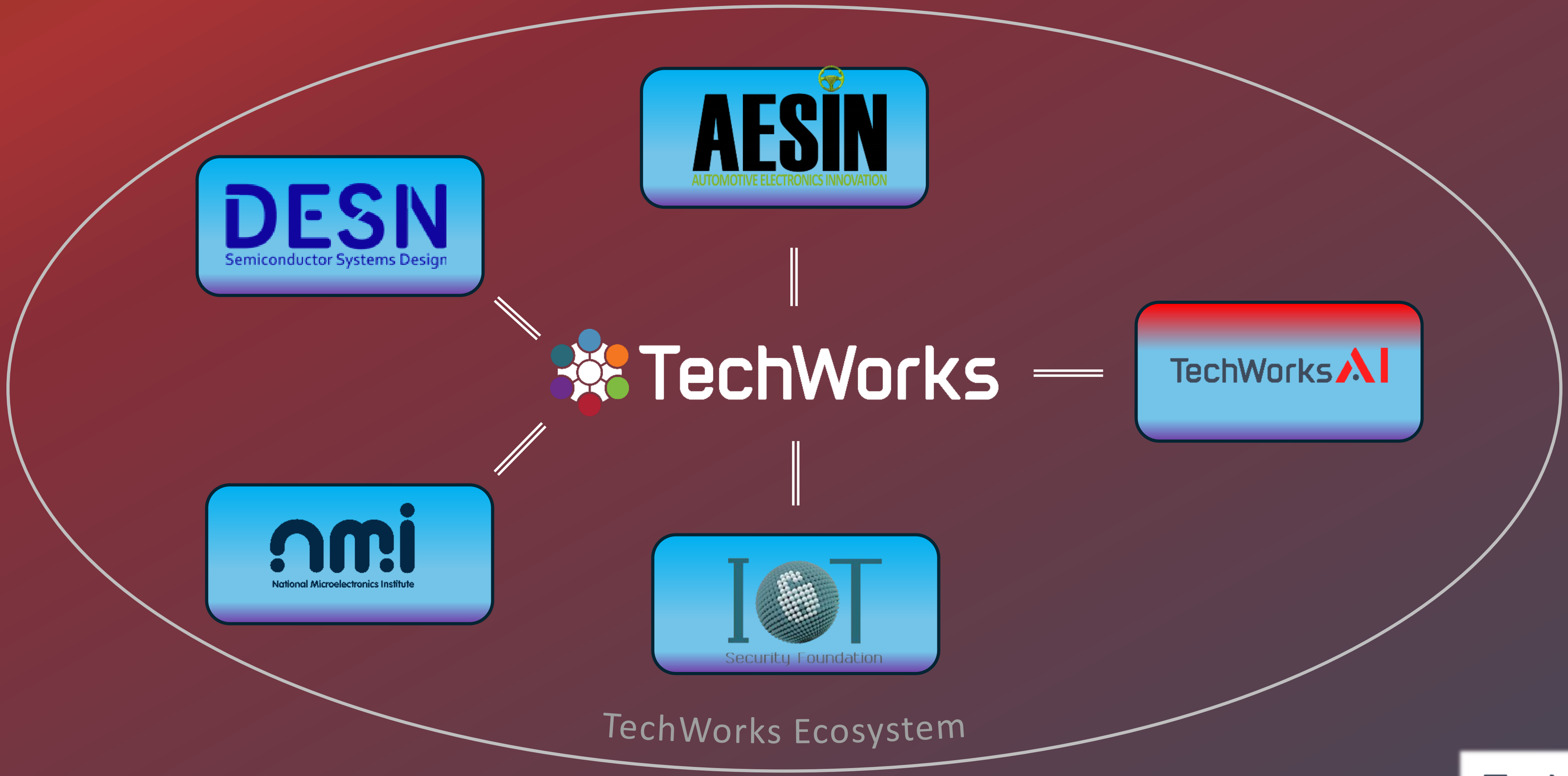
Verification Futures

Gareth Richards
AI Network Manager

TechWorks AI

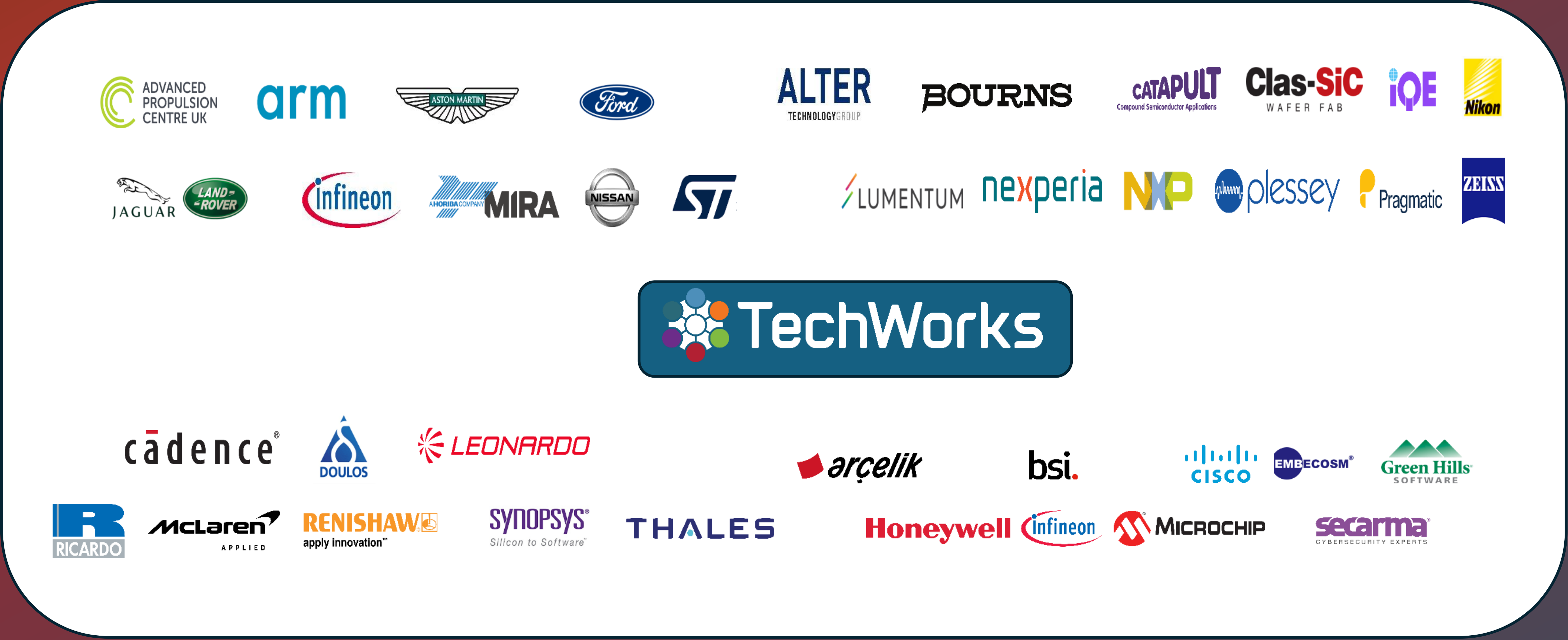
TechWorks AI

The TechWorks Ecosystem


AESIN
AUTOMOTIVE ELECTRONICS INNOVATION
DESN
Semiconductor Systems Design
TechWorks
TechWorks AI
nmi
National Microelectronics Institute
IoT
Security Foundation
TechWorks Ecosystem

The TechWorks Ecosystem

>300 members



What we do

Thought Leadership & Cross-sector Collaboration

Sector Analysis & Research

Networking & Events

Government Relations

Industry Strategy & Advocacy

Collaborative Projects

Best Practice & Open Standards

TAIBOM

UK Research and Innovation

Trustable AI Bill Of Materials - Project partners

- BAE Systems
- BSI
- Copper Horse
- NquiringMinds
- TechWorks-AI
- University of Oxford

- Versioning, Attestation, Integrity

TechWorks AI

TAIBOM

UK Research and Innovation

Trustable AI Bill Of Materials - Project partners

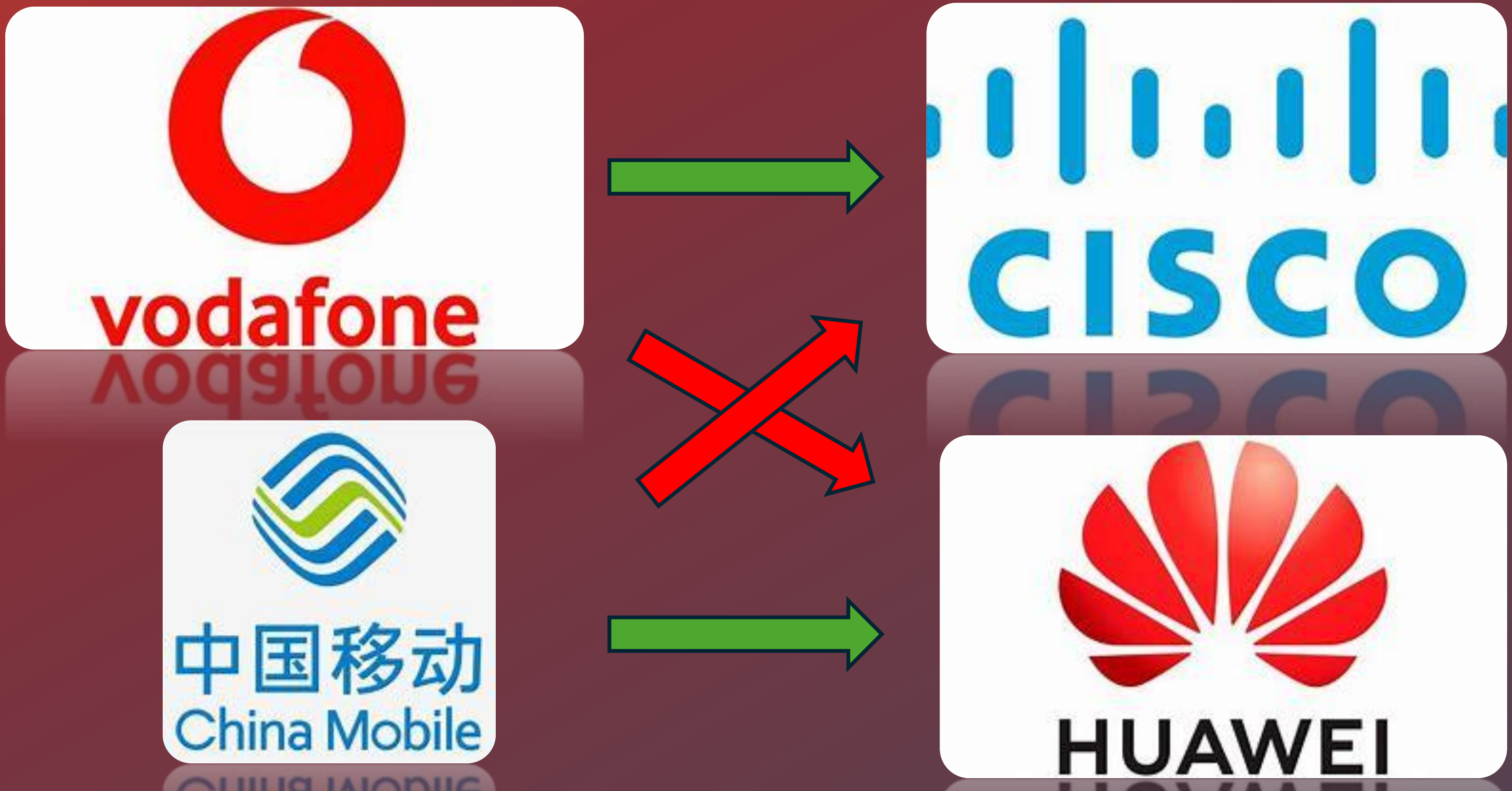
- BAE Systems
- BSI
- Copper Horse
- NquiringMinds
- TechWorks-AI
- University of Oxford

- Versioning, Attestation, Integrity → Trustability

TechWorks AI

TRUST

What do we mean by trust?



The willingness of a party to be **vulnerable** to the actions of another party based on the **expectation** that the other will perform a particular action important to the trustor, **irrespective of the ability to monitor or control that other party***

*The Academy of Management Review, Vol.20, No. 3 (Jul., 1995), pp.709-734 (26 pages)
https://doi.org/10.2307/258792

Foundations of Trust

Trusting soup...

Integrity

FILTERED WATER, ORGANIC DICED TOMATOES, ORGANIC ONIONS, ORGANIC CARROTS, ORGANIC KIDNEY BEANS, ORGANIC POTATOES, ORGANIC CELERY, ORGANIC GREEN BEANS, ORGANIC PASTA (ORGANIC DURUM WHEAT SEMOLINA FLOUR, WATER), ORGANIC PEAS, ORGANIC LEEKS, SEA SALT, ORGANIC HIGH OLEIC SAFFLOWER AND/OR SUNFLOWER OIL, ORGANIC SPICES, ORGANIC BASIL, ORGANIC GARLIC, ORGANIC BLACK PEPPER. **CONTAINS WHEAT.**

Attestation

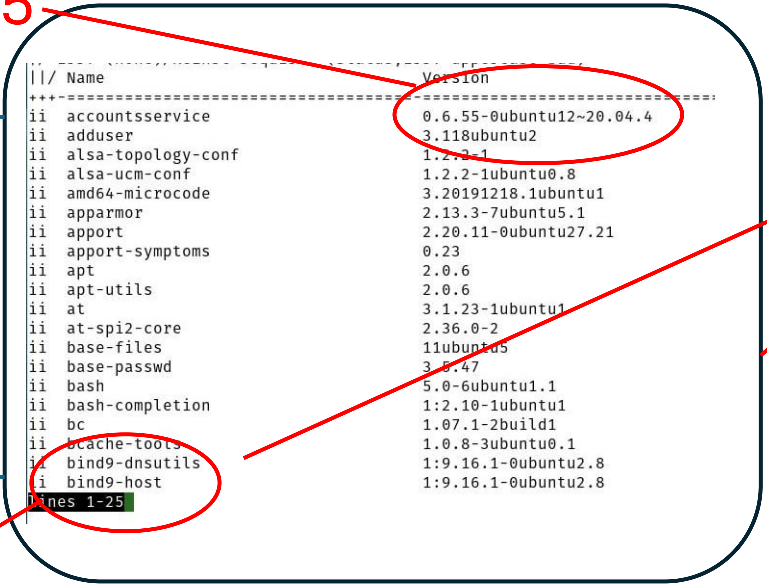
Provenance

Foundations of Trust

Trusting software…

Integrity — # = 5F3D75

XYZ Corp



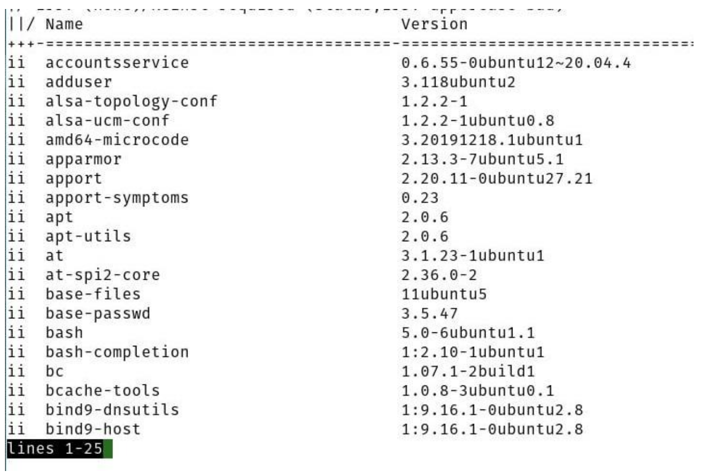
```
||/ Name                                         Version
+++-=============================================-=================================
ii  accountsservice                              0.6.55-0ubuntu12~20.04.4
ii  adduser                                      3.118ubuntu2
ii  alsa-topology-conf                           1.2.2-1
ii  alsa-ucm-conf                                1.2.2-1ubuntu0.8
ii  amd64-microcode                              3.20191218.1ubuntu1
ii  apparmor                                     2.13.3-7ubuntu5.1
ii  apport                                       2.20.11-0ubuntu27.21
ii  apport-symptoms                              0.23
ii  apt                                          2.0.6
ii  apt-utils                                    2.0.6
ii  at                                           3.1.23-1ubuntu1
ii  at-spi2-core                                 2.36.0-2
ii  base-files                                   11ubuntu5
ii  base-passwd                                  3.5.47
ii  bash                                         5.0-6ubuntu1.1
ii  bash-completion                              1:2.10-1ubuntu1
ii  bc                                           1.07.1-2build1
ii  bcache-tools                                 1.0.8-3ubuntu0.1
ii  bind9-dnsutils                               1:9.16.1-0ubuntu2.8
ii  bind9-host                                   1:9.16.1-0ubuntu2.8
lines 1-25
```

Attestation

Provenance

Foundations of Trust

Trusting software…SBOM



```
||/ Name                                           Version
+++-==============================================-==============================
ii  accountsservice                                0.6.55-0ubuntu12~20.04.4
ii  adduser                                        3.118ubuntu2
ii  alsa-topology-conf                             1.2.2-1
ii  alsa-ucm-conf                                  1.2.2-1ubuntu0.8
ii  amd64-microcode                                3.20191218.1ubuntu1
ii  apparmor                                       2.13.3-7ubuntu5.1
ii  apport                                         2.20.11-0ubuntu27.21
ii  apport-symptoms                                0.23
ii  apt                                            2.0.6
ii  apt-utils                                      2.0.6
ii  at                                             3.1.23-1ubuntu1
ii  at-spi2-core                                   2.36.0-2
ii  base-files                                     11ubuntu5
ii  base-passwd                                    3.5.47
ii  bash                                           5.0-6ubuntu1.1
ii  bash-completion                                1:2.10-1ubuntu1
ii  bc                                             1.07.1-2build1
ii  bcache-tools                                   1.0.8-3ubuntu0.1
ii  bind9-dnsutils                                 1:9.16.1-0ubuntu2.8
ii  bind9-host                                     1:9.16.1-0ubuntu2.8
lines 1-25
```

- But wait - we DID write all our own software!
- What about the compiler, what about the OS etc?

- Did we write **all** of the software (probably not)?
- What is the **provenance** of library elements, object code drivers etc?
- Do we know they are all **vulnerability free**? Can we prove it?

- EU Cyber Resilience Act (CRA) 2026/7 – software must have an SBOM and **have no known exploitable vulnerabilities**

- US EO 14028 – now - SBOM must be provided for federal projects

Foundations of Trust

"The reality is we know more about what is in our sausages than our software"

Ollie Whitehouse, CTO, NCSC

Foundations of Trust

Regulated industries have been worrying about this for some time. The CRA, in particular, now imposes similar requirements on most commercial products

- Payments – PCI & UK Finance
  - Mandatory third-party (accredited lab) security evaluation of all software components. Certification to Common Criteria EAL 4+
- Medical Devices
  - IEC 62304 – Software Lifecycle Process – required (or equivalent) for US FDA and EU MDR compliance
- Etc

SOUP
Software Of Unknown Provenance

Foundations of Trust

EU Product Liability Directive – implementation by December 2026

*”For the first time, **software** - both embedded and standalone - will be treated as a product, subject to the <u>same rigorous liability rules as physical goods</u>.*

Manufacturers face liability for up to 10 years for defective products, extended to 25 years for latent personal injury claims.”

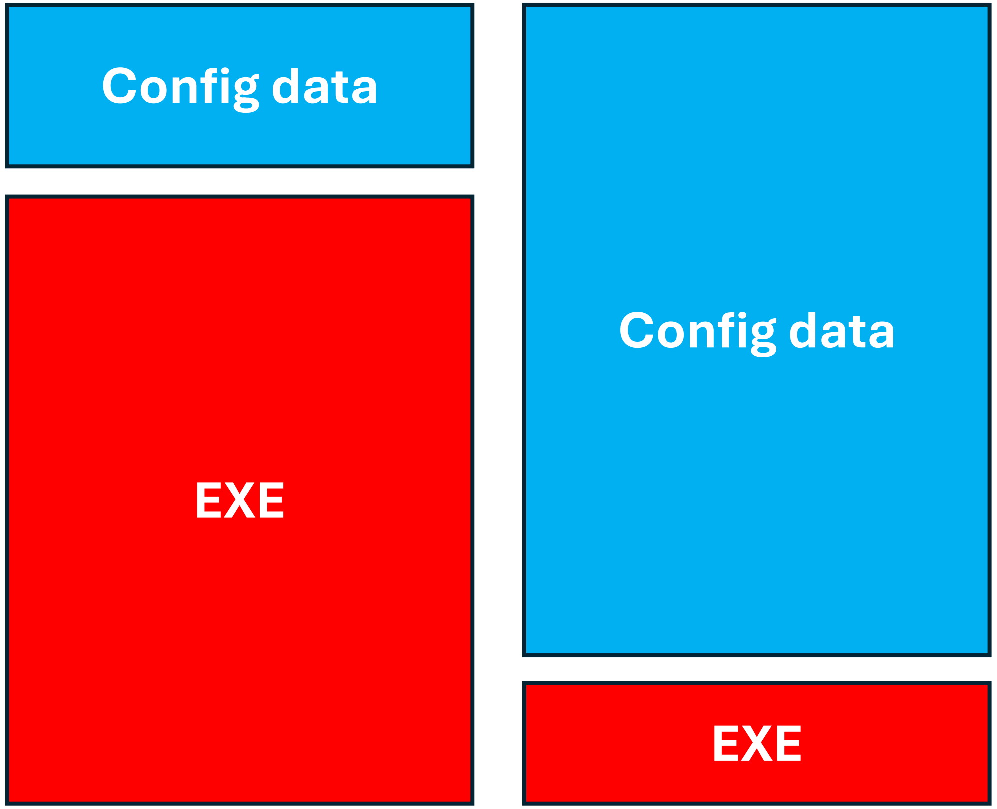
www.fladgate.com

Foundations of Trust

Traditional software

Locked functional (compiled) specification- updated regularly with small configuration file.

Functionality - defined behaviour



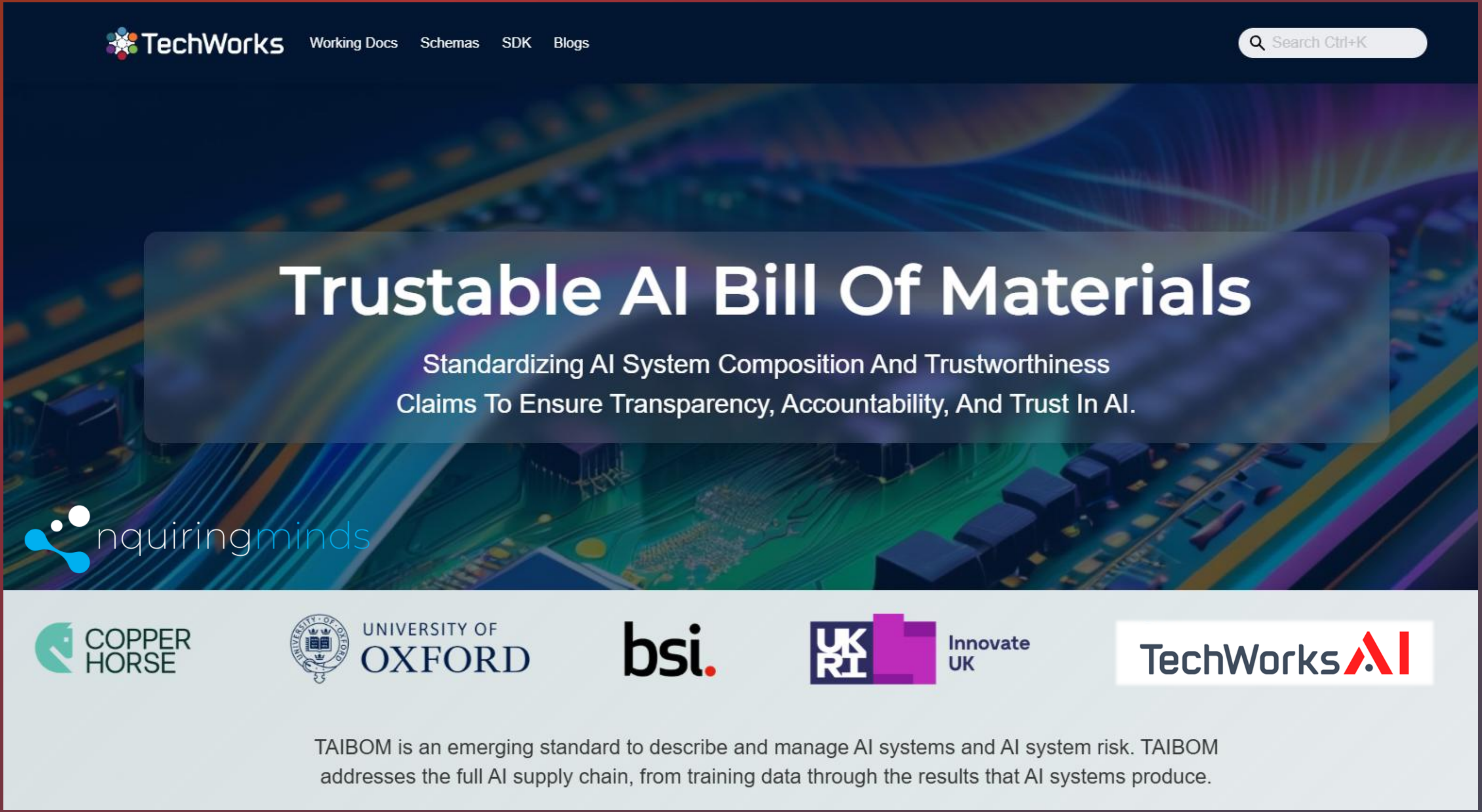
Can we hash and sign this quickly enough to be useful?

AI

Small functional code, infrequently updated with **huge configuration file** (weights)

Data driven behaviour

Foundations of Trust

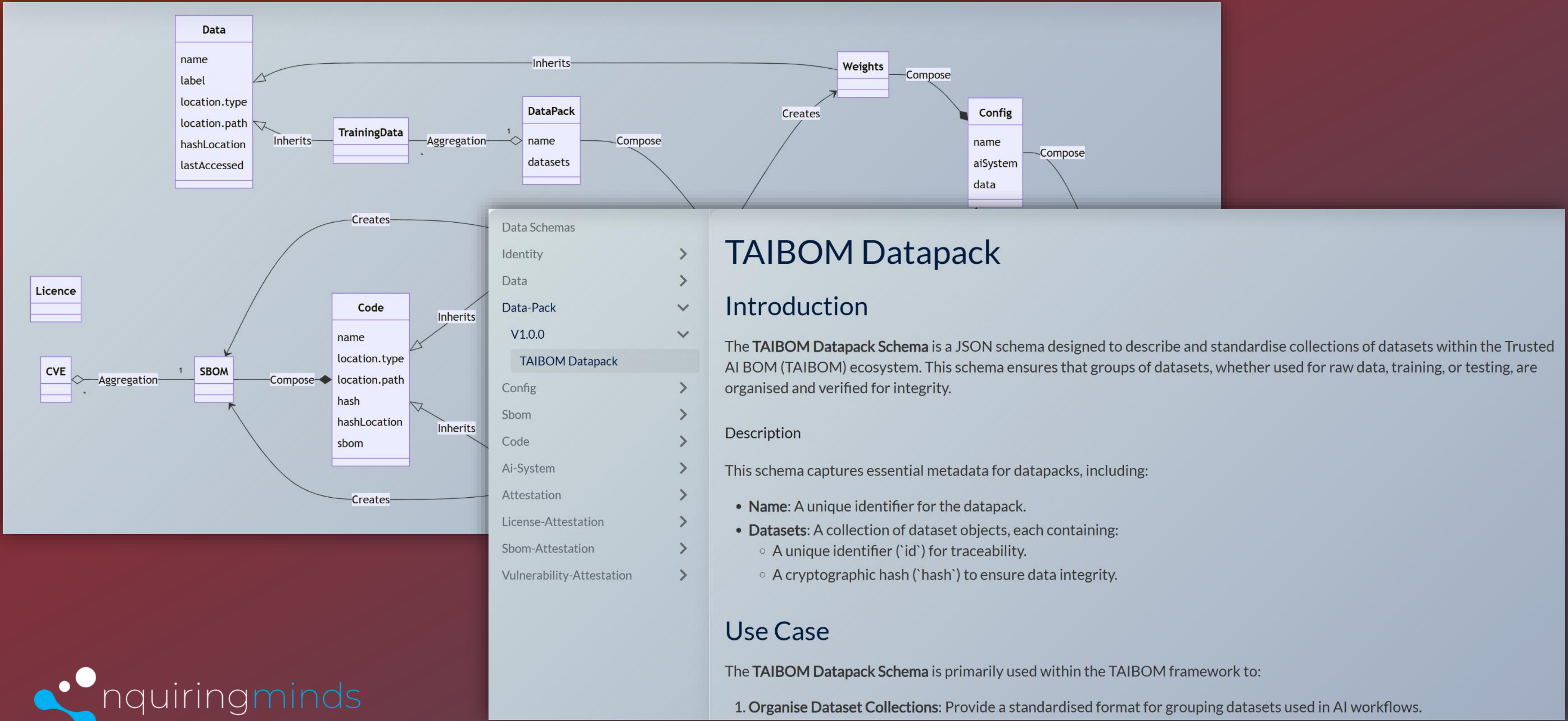


Trustable AI Bill Of Materials

Standardizing AI System Composition And Trustworthiness Claims To Ensure Transparency, Accountability, And Trust In AI.

TAIBOM is an emerging standard to describe and manage AI systems and AI system risk. TAIBOM addresses the full AI supply chain, from training data through the results that AI systems produce.

Foundations of Trust



TAIBOM Datapack

Introduction

The **TAIBOM Datapack Schema** is a JSON schema designed to describe and standardise collections of datasets within the Trusted AI BOM (TAIBOM) ecosystem. This schema ensures that groups of datasets, whether used for raw data, training, or testing, are organised and verified for integrity.

Description

This schema captures essential metadata for datapacks, including:

- **Name**: A unique identifier for the datapack.
- **Datasets**: A collection of dataset objects, each containing:
  - A unique identifier (`id`) for traceability.
  - A cryptographic hash (`hash`) to ensure data integrity.

Use Case

The **TAIBOM Datapack Schema** is primarily used within the TAIBOM framework to:

1. **Organise Dataset Collections**: Provide a standardised format for grouping datasets used in AI workflows.

nquiringminds

Foundations of Trust

What Claims / Attestations can be made?

XYZ made this
Authorship claim

It was made this way
A statement of conformance or best practice

It works well
A statement of performance – test results

It didn't break
A positive pen test

Software licenses
Code can be used this way

Data licenses
Data was intended for this purpose

Export Conditions
Restrictions or FOCI elements

Anything else
You decide what else you want….

Foundations of Trust

Who can/could make the claims?

The author
Creator of data

The developer
Creator of code

The publisher
Financial underwriter

The auditor
Someone who tested it

A third party
Any trusted (untrusted) entity

An agent
AI analysing AI

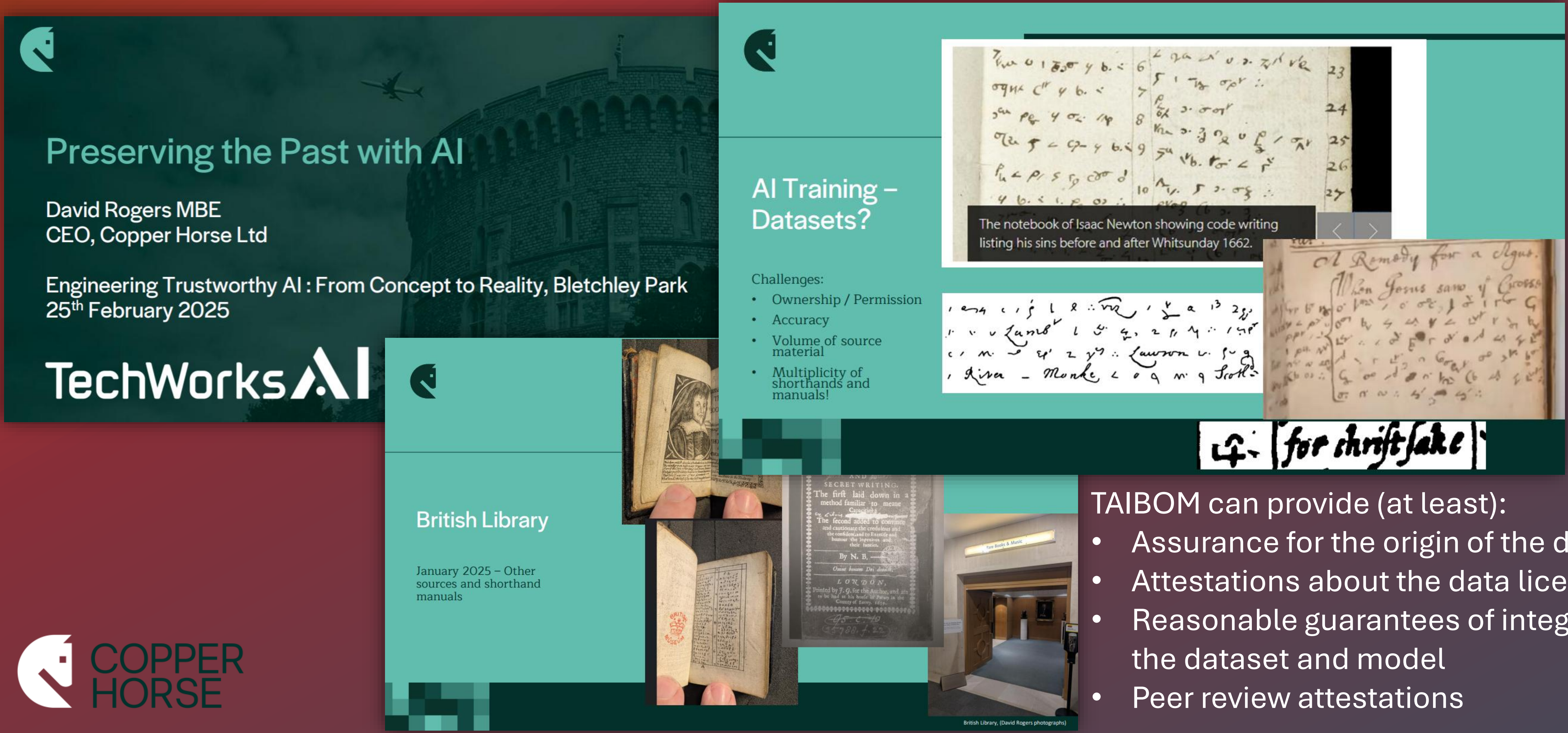
....

Foundations of Trust

Use cases

Foundations of Trust

Use case – 17th century shorthand



Preserving the Past with AI

David Rogers MBE
CEO, Copper Horse Ltd

Engineering Trustworthy AI : From Concept to Reality, Bletchley Park
25th February 2025

TechWorks AI

AI Training – Datasets?

Challenges:

- Ownership / Permission
- Accuracy
- Volume of source material
- Multiplicity of shorthands and manuals!

The notebook of Isaac Newton showing code writing listing his sins before and after Whitsunday 1662.

British Library

January 2025 – Other sources and shorthand manuals

British Library, (David Rogers photographs)

TAIBOM can provide (at least):

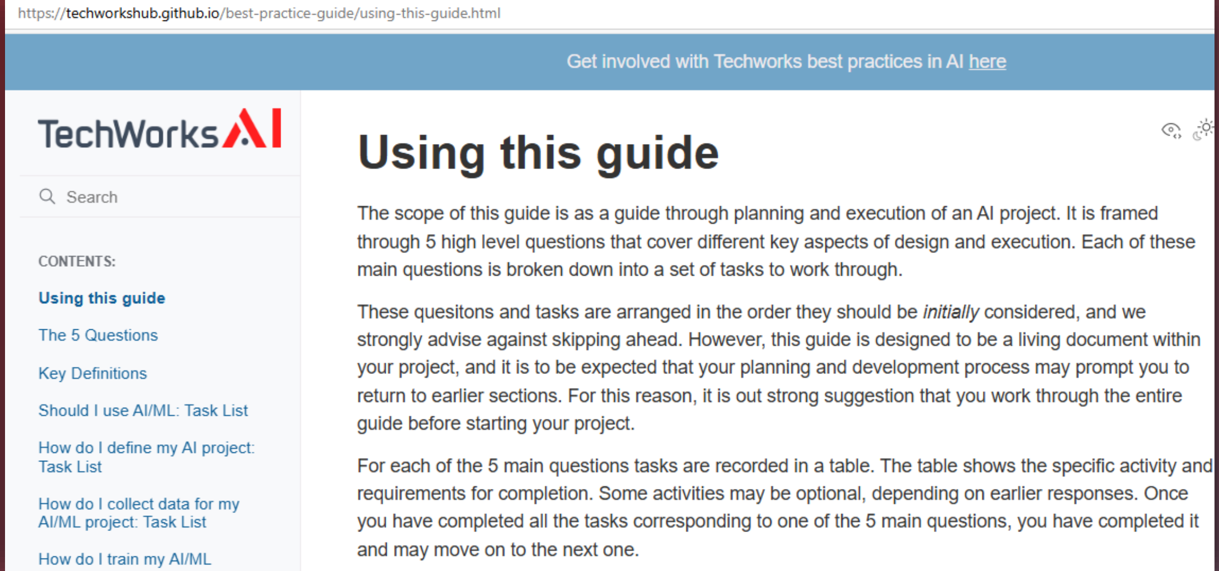
- Assurance for the origin of the dataset
- Attestations about the data license(s)
- Reasonable guarantees of integrity for the dataset and model
- Peer review attestations

COPPER HORSE

Foundations of Trust

And also.....

Best Practice Guide For AI Systems


https://techworkshub.github.io/best-practice-guide/using-this-guide.html
Get involved with Techworks best practices in AI here
TechWorks AI
Search
CONTENTS:
Using this guide
The 5 Questions
Key Definitions
Should I use AI/ML: Task List
How do I define my AI project: Task List
How do I collect data for my AI/ML project: Task List
How do I train my AI/ML

Using this guide

The scope of this guide is as a guide through planning and execution of an AI project. It is framed through 5 high level questions that cover different key aspects of design and execution. Each of these main questions is broken down into a set of tasks to work through.

These quesitons and tasks are arranged in the order they should be *initially* considered, and we strongly advise against skipping ahead. However, this guide is designed to be a living document within your project, and it is to be expected that your planning and development process may prompt you to return to earlier sections. For this reason, it is out strong suggestion that you work through the entire guide before starting your project.

For each of the 5 main questions tasks are recorded in a table. The table shows the specific activity and requirements for completion. Some activities may be optional, depending on earlier responses. Once you have completed all the tasks corresponding to one of the 5 main questions, you have completed it and may move on to the next one.

Model Cards Plus (MC+)

Department for Science, Innovation & Technology

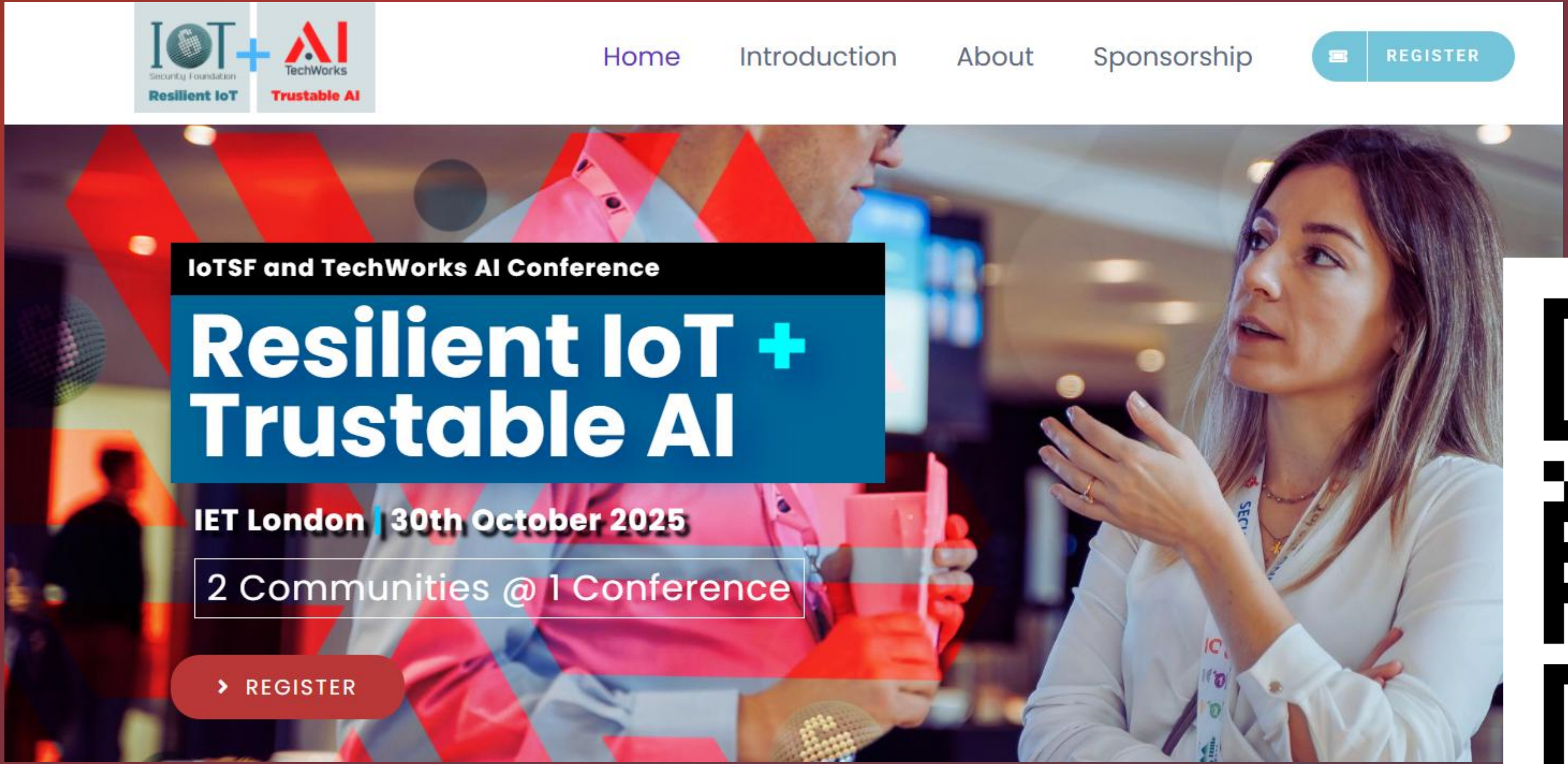
Foundations of Trust

Joining TechWorks-AI – www.techworksai.org

The TAIBOM project – www.taibom.org

We are very interested in **new use-cases** to apply TAIBOM , Best Practice Guide and MC+
Gareth.Richards@techworks.org.uk

Annual Conference – 30th Oct. IET London



IoT + AI

Security Foundation TechWorks

Resilient IoT Trustable AI

Home Introduction About Sponsorship REGISTER

IoTSF and TechWorks AI Conference

Resilient IoT + Trustable AI

IET London | 30th October 2025

2 Communities @ 1 Conference

› REGISTER

TechWorksAI