SCC PANEL

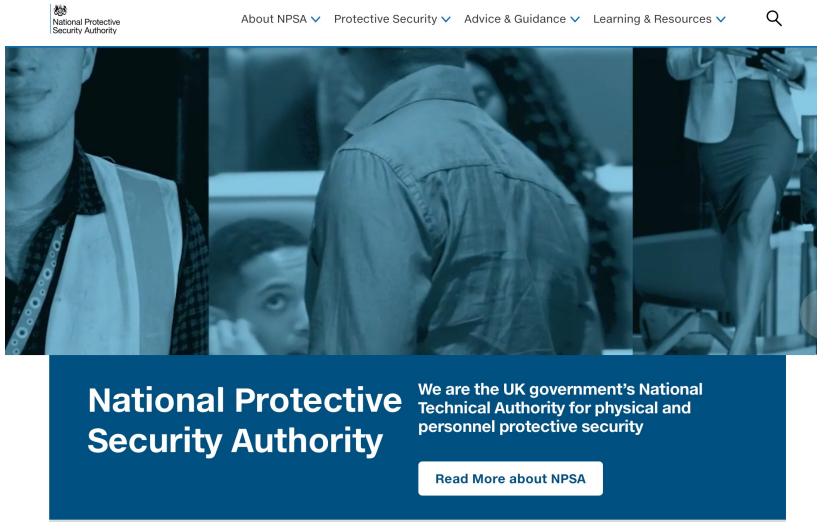
USING ASSURANCE CASES FOR A GO/NO GO DECISION: THEORY AND PRACTICE

Prof Robin Bloomfield FREng
City, University of London and Adelard
May 2024

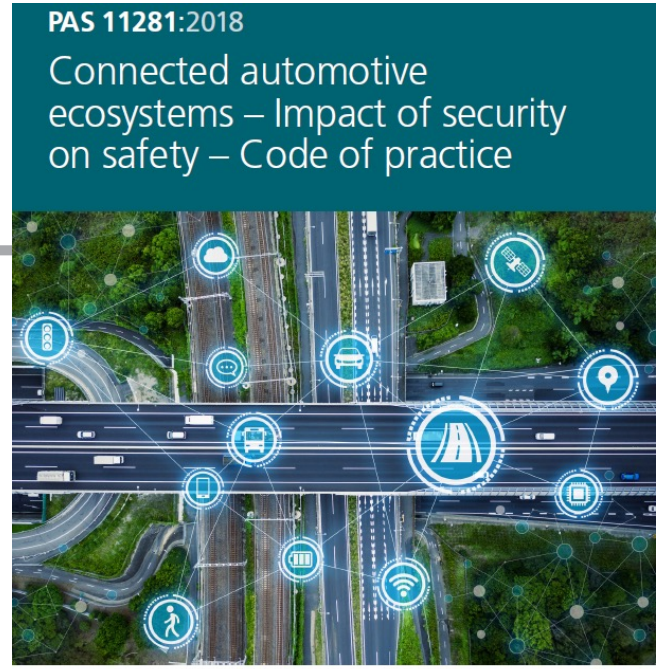
robin.bloomfield@nccgroup.com
r.e.bloomfield@city.ac.uk

Adelard, 24 Waterside, 44-48 Wharf Road, London N1 7UX
T +44 20 7832 5850 E office@adelard.com W www.adelard.com

“IF IT’S NOT SECURE, IT’S NOT SAFE”.



Security Informed Safety video



CPNI. Centre for the Protection of National Infrastructure

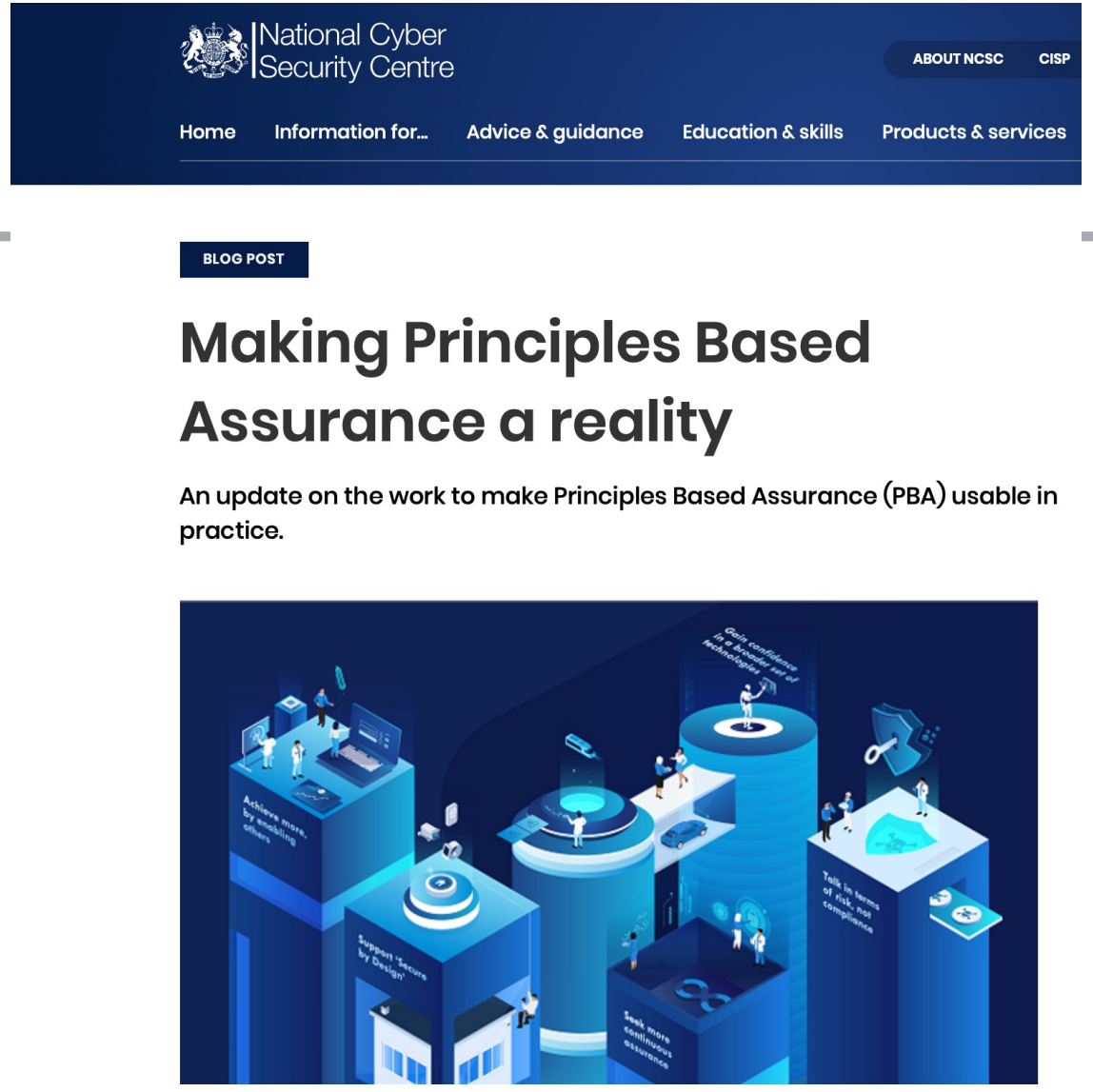
bsi.

https://www.npsa.gov.uk/security-informed-safety

ADELARD part of nccgroup

SECURITY – UK NCSC

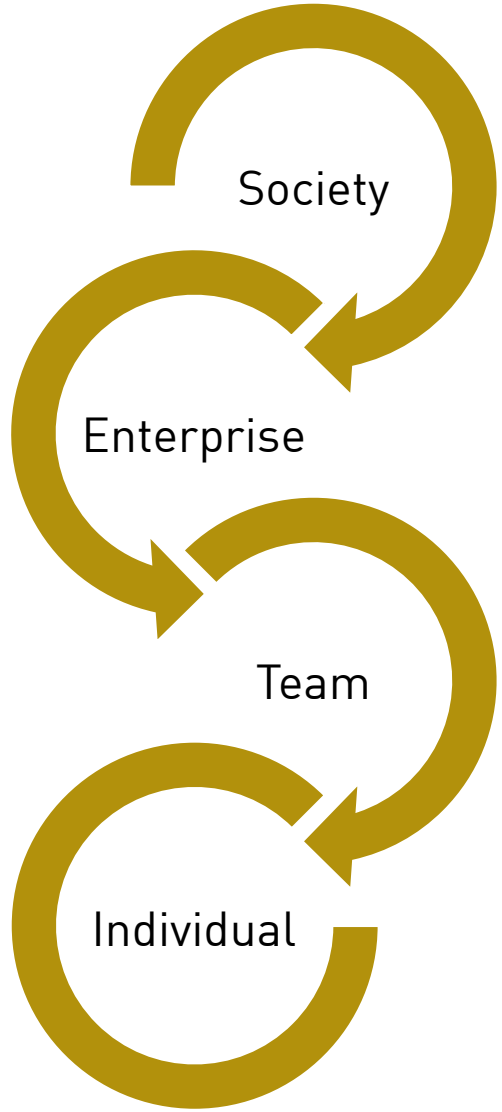
- https://www.ncsc.gov.uk/blog-post/making-principles-based-assurance-a-reality



TWO QUESTIONS

- How confident am I in the claim being made?
- What is the impact on the decision?

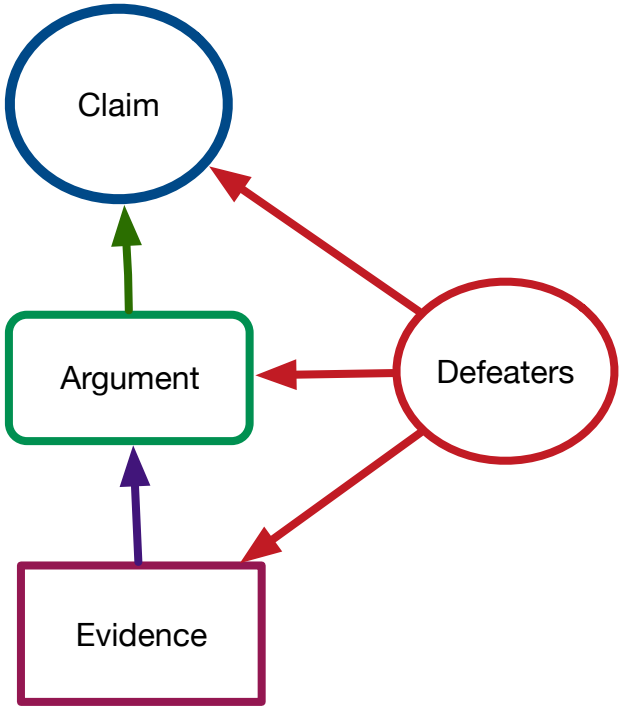
Reasoning and communication



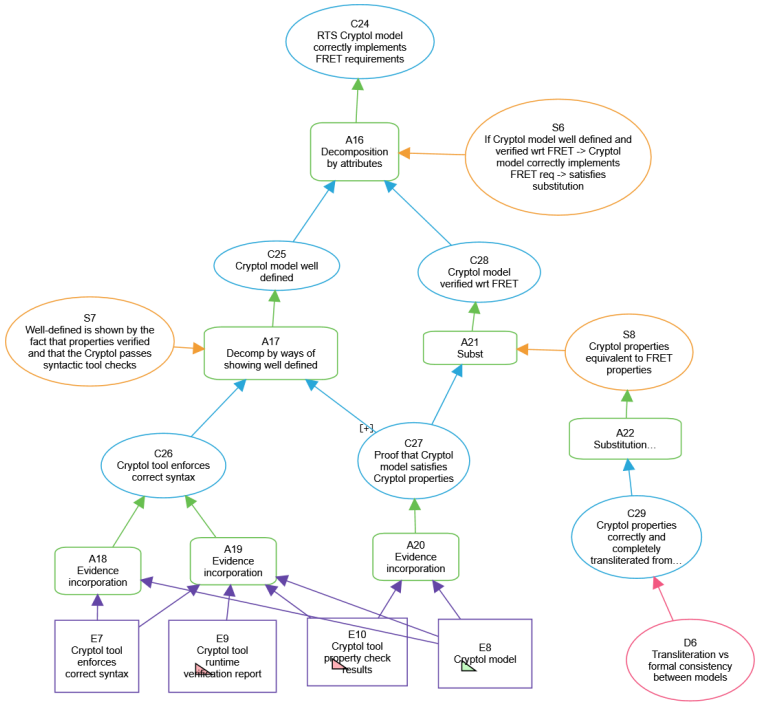
ASSURANCE 2.0

R Bloomfield and J Rushby, Assurance 2.0 Manifesto
https://arxiv.org/abs/2004.10474

CLAIMS, ARGUMENTS, EVIDENCE, DEFEATERS



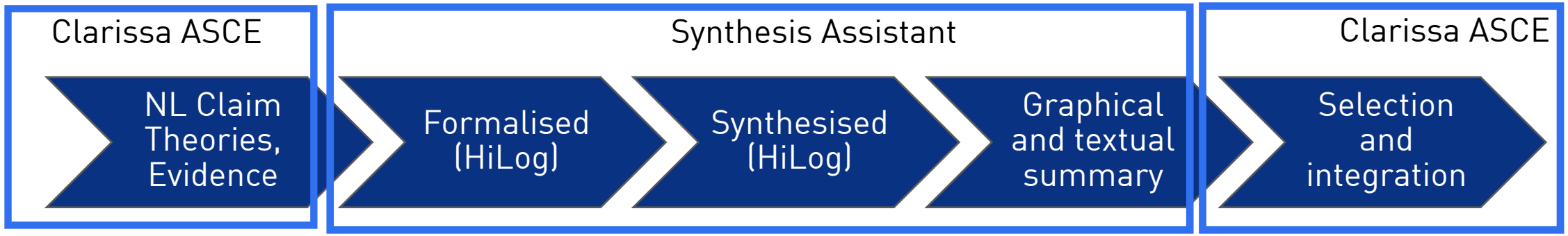
- ***Claims*** **-** assertions put forward for general acceptance
- ***Arguments*** **-** link the evidence to the claim
- ***Evidence*** **-** the basis of the justification of the claim
- ***Defeater*** **–** reasons for doubting



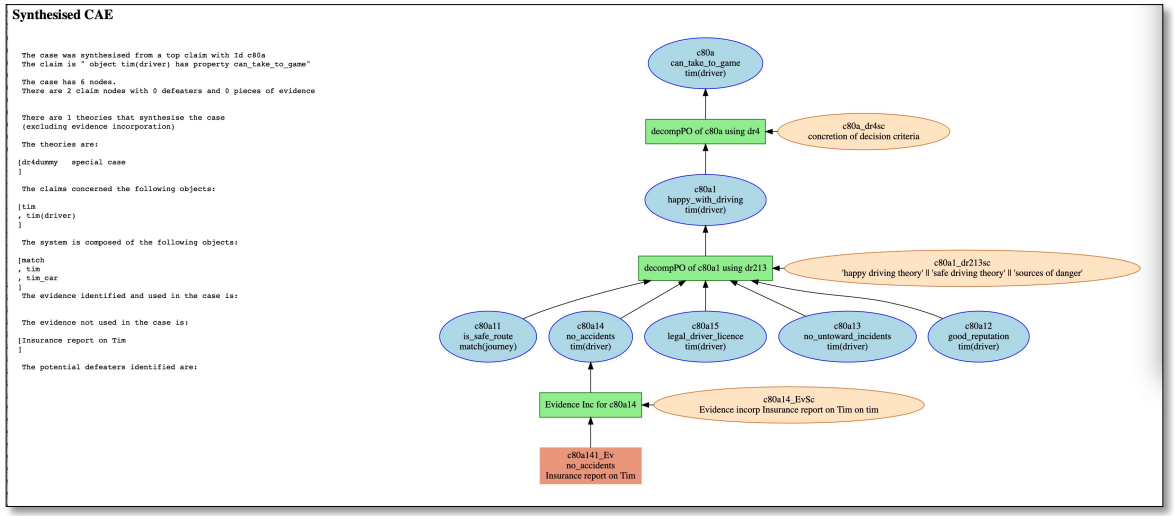
ADELARD part of nccgroup

ASSURANCE CASE SYNTHESIS

Synthesis Assistant is a research tool designed to synthesize claims, arguments and evidence structures from a root or top-level claim.



- Given:
  - Top-level claim (defined in ErgoAI or node imported from an ASCE file)
  - Definition of the system structure
  - Possible defeaters
  - Theories used to develop the case
  - Evidences for the case
  - LLM support



DEVELOPMENT AND ASSESSMENT OF ASSURANCE CASES
Positive, negative, residual doubts

- **Positive:** logical soundness of argument plus scientific assessment of theories
  - Soundness is logical validity (checkable) plus credibility of evidence and reasoning
  - Credibility of evidence is "weighed" by confirmation measures
    - *Forces contemplation of defeaters at evidence level*
  - And ensured for reasoning steps by (checkable) side-conditions (for deductiveness)
- **Negative:** active search for and resolution of defeaters
  - Defeaters are retained to assist evaluators
  - Value their coverage, significance, and diversity more than quantity
- **Residual Doubts:** what about the gaps?
  - Localized for analysis as potentially valid defeaters, inductive steps
  - Need to assess risk: consequences and likelihood
  - We propagate probabilistic belief in several ways to assist different stakeholders
    - Internalized explicitly within claims and associated models/theories
    - Conservative sum of doubts
  - Purpose is to explore assessments and tradeoffs, not deliver verdict
- Overall evaluation yields degree of belief in top claim
  - Sentencing statement or Assurance Case report supports overall verdict

Confidence report coming on Arxiv

ADELARD part of nccgroup

SUMMARY REPORT

- The purpose of an assurance case is to support decision to deploy (or not) a system or service. The task of evaluators is captured in a summary report:

  " On the basis of this case and an examination of other relevant documentation, I judge the proposed system to be effective/adequately safe/unsafe/secure. . . "

  or, the case is insufficient to make a judgement

  - "I believe my judgement of this case is sound and valid because. . .
    - I understand the context and criticality of the decision. . .
    - I understand the system. . .
    - I find a clear thread of reasoning from evidence to claim. . .
    - Evidence provided is sufficient/insufficient for evidence-based decision making
    - I have actively explored doubts. .
    - I have also identified what evidence would be capable of disproving. . .
    - I have considered and addressed biases and fallacies. . . "

- For each of these can map Assurance 2.0 methodology to show where it provides support.

ADELARD part of nccgroup

DECISION MAKING - WHAT IS THE IMPACT ON THE DECISION?

- How wrong can I be?
- How much does it matter ?
- Models of "chain of confidence"

$$conf\,(PE \rightarrow property_{target}).$$

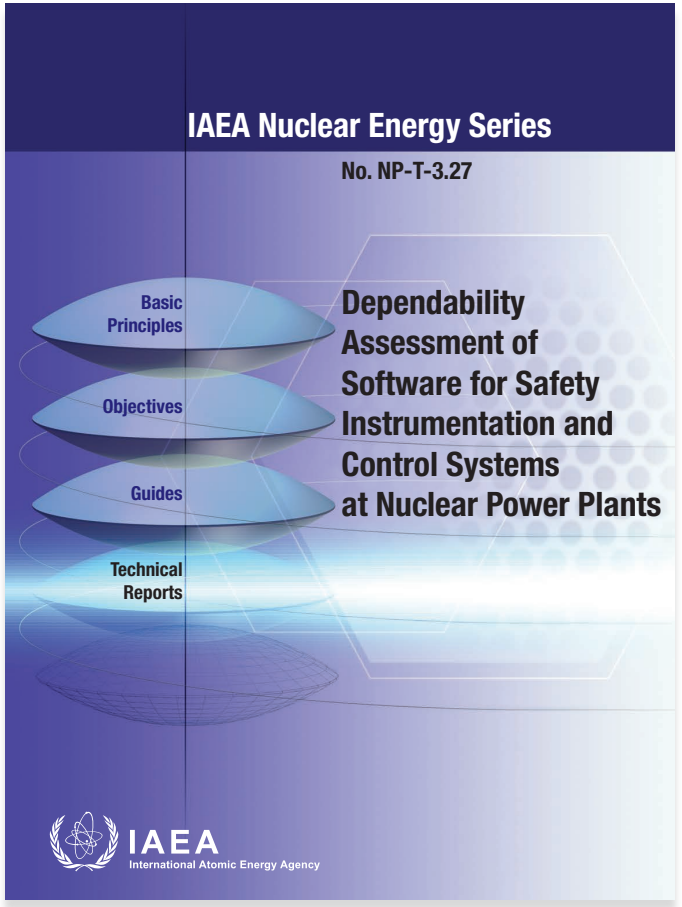
But for the case where the assumption does not hold, i.e.

$$\left(1 - conf\left(PE \rightarrow property_{target}\right)\right),$$

some alternative worst case bound property value is used, $property_{worst}$.

The expected value of the property includes both cases:

$$property_{expected} = conf\left(PE \rightarrow property_{target}\right) property_{target} + \left(1 - conf\left(PE \rightarrow property_{target}\right)\right) property_{worst}$$



MODELS OF CHAIN OF CONFIDENCE

- Example: Confidence that pfd requirement is met
- We can model this with a "chain of confidence" approach where :
  - expected pfd = pfd_{target}*$conf_{PE}$ + (1-$conf_{PE}$)* pfd_{max}
- Where
  - pfd_{target} is the required pfd
  - $conf_{PE}$ is our confidence in the claim of pfd_{target} (based on PE)
  - pfd_{max} is the upper bound on the pfd if our judgment of pfd_{target} is wrong

But ….avoid positional bargaining, 1 dimensionality

Prof Robin E Bloomfield FREng

Adelard (part of NCC Group)

and City, University of London

r.e.bloomfield@city.ac.uk

robin.bloomfield@nccgroup.com

ADELARD

Assurance 2.0 joint work with John Rushby, SRI