

Claims, Arguments, Evidence



Introduction to the CAE Framework

Dennis Vetter

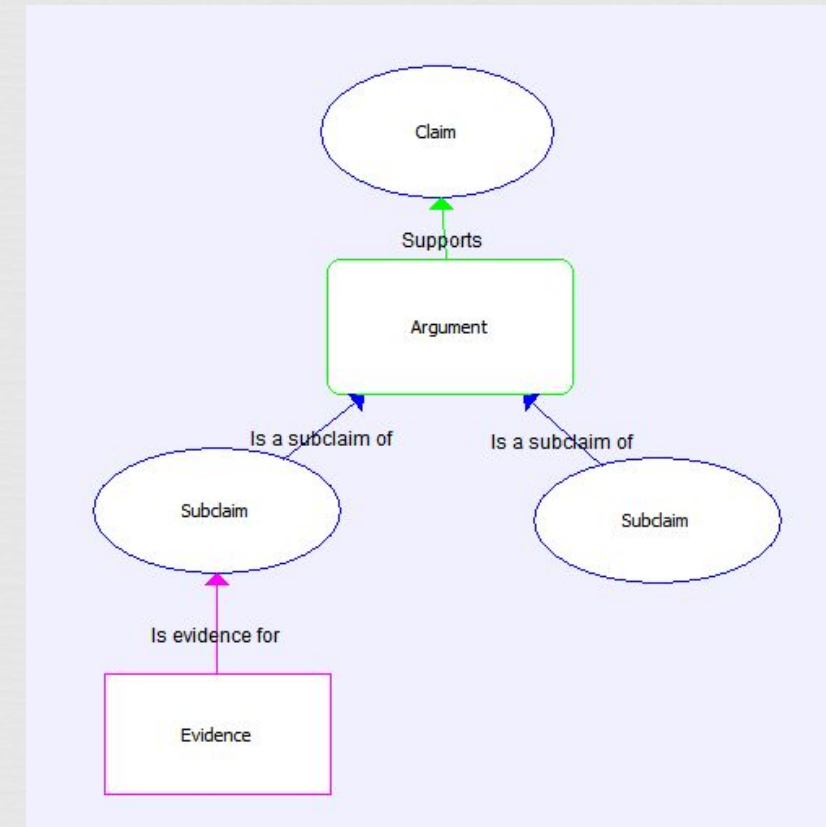
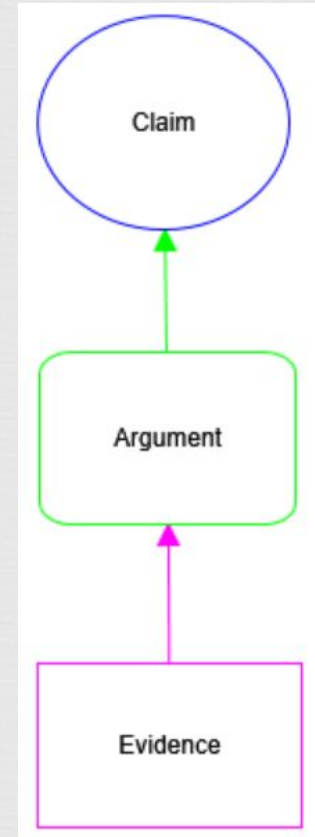
CAE Framework?



- CAE = Claims, Arguments, Evidence
 - Framework for structured reasoning in complex systems
 - Applications in multiple disciplines
 - Nuclear Industry
 - Medicine
 - Security
 - NASA
- For us: how can we connect vendor statements to available evidence?
 - mental guideline

Claims

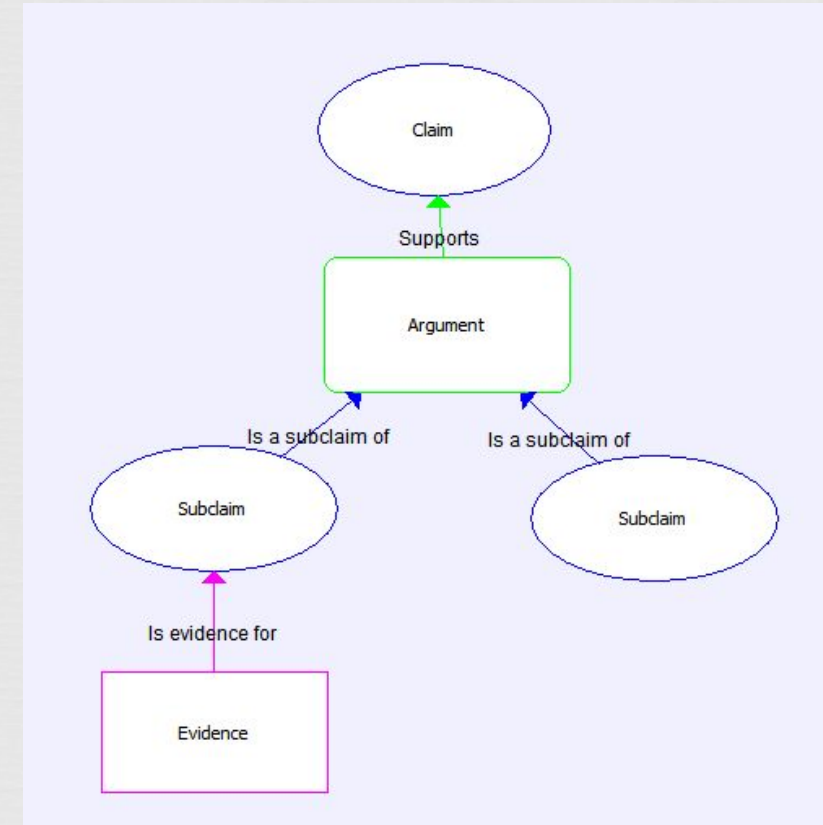
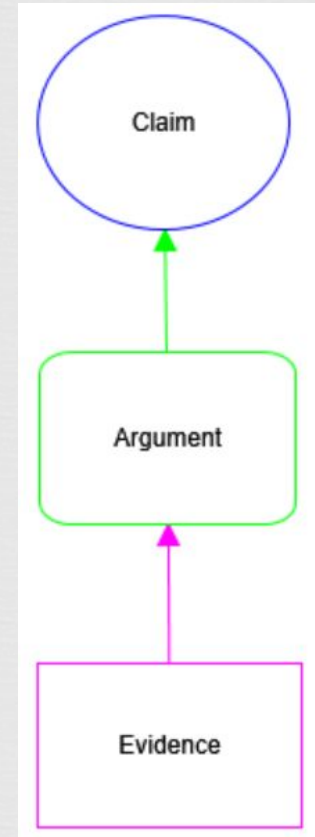
- Statement on one property of a system / object
 - Can be assessed **true** or **false**
 - May require definition of context
- Is supported by sub-claims, arguments or evidence
- Example:
“This AI system is trustworthy”



Arguments



- Explains the approach of supporting the parent claim
 - can be omitted if claim is trivial / understood by audience
 - can be a combination of sub-claims
- Example:
Prove the AI system is trustworthy by showing it has *high performance* and *fulfills the regulations*

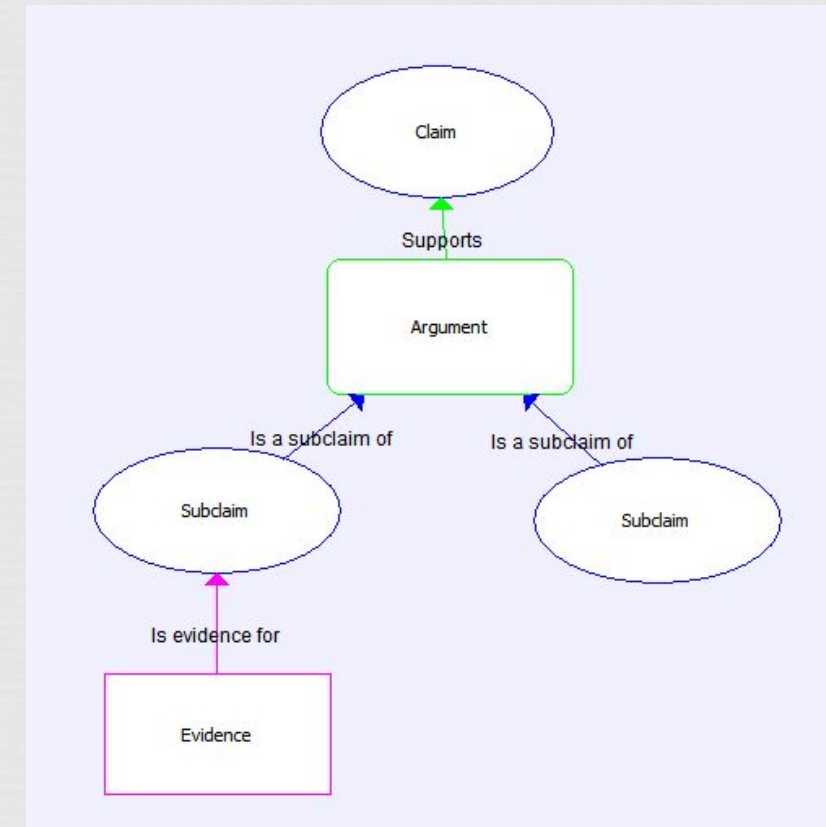
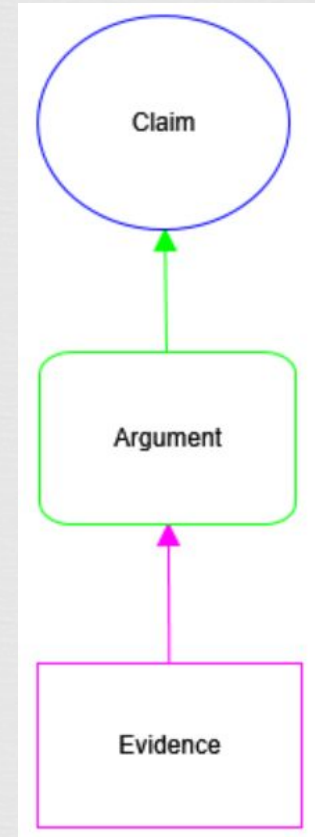


Evidence

- Establishes facts that support / oppose a claim
 - Can you trust the fact?
 - Who is the source?
 - How was it developed, is it verifiable, ...
 - (scientific) reports, clinical studies, official certifications, ...

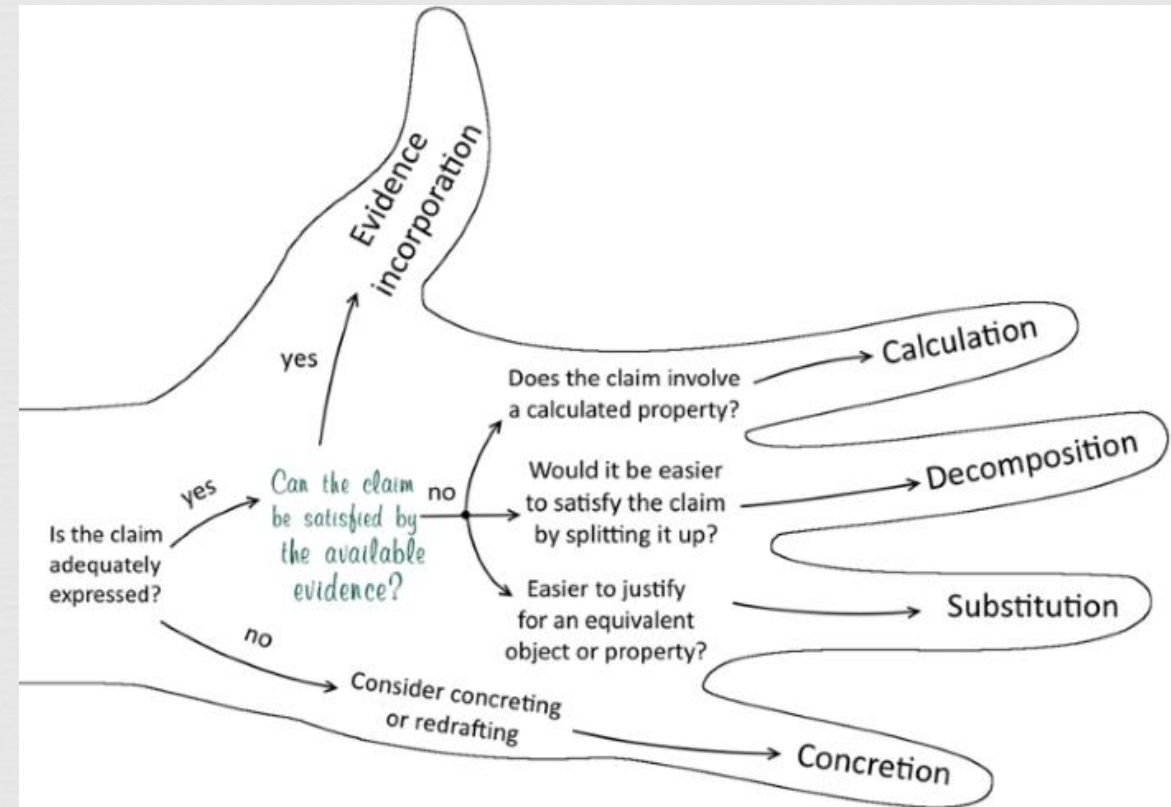
- Example:

High performance of the AI system was shown in clinical trial no. aaa-1234



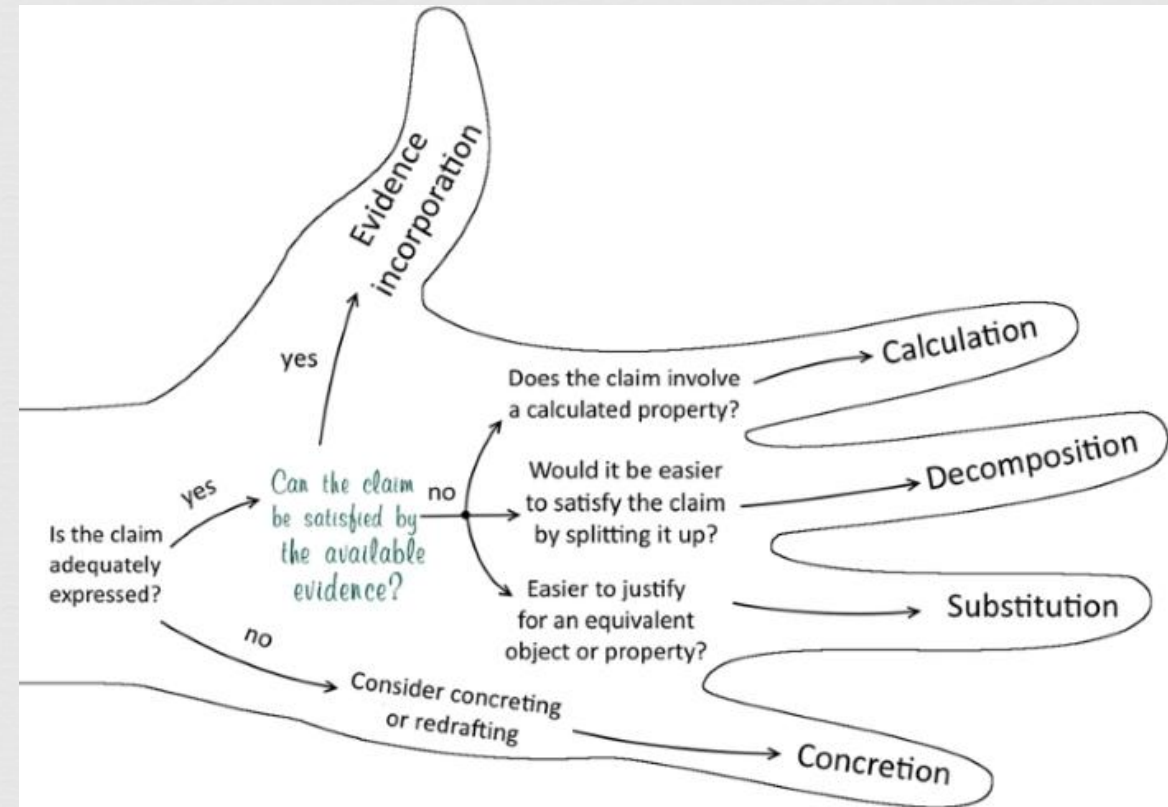
How to verify Claims with Evidence

- Difficulty of verification depends on claim and available evidence
 - Claim can be too complex
 - Claim can be expressed inadequately
 - Evidence can only prove part of the claim
 - ...
- “Helping hand” as a guideline



Evidence incorporation

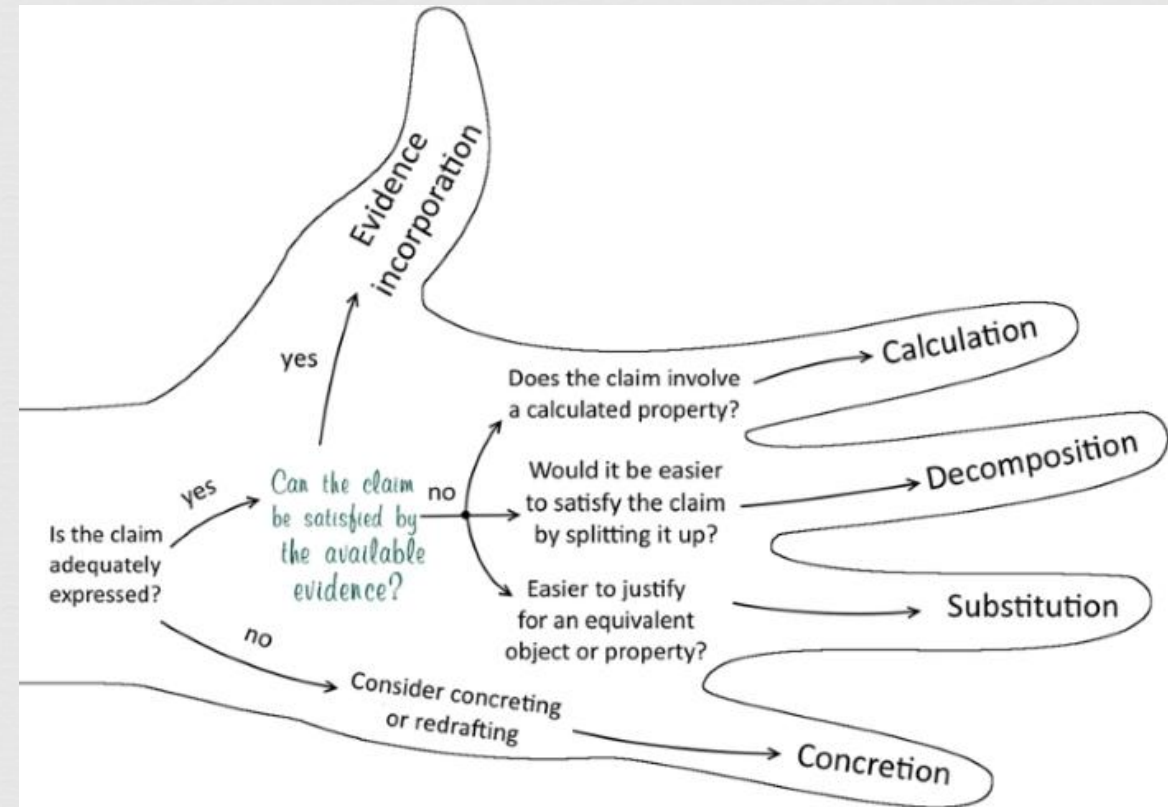
- Easiest case
- Claim is adequately expressed
- Available evidence can directly show that the claim is true / false
 - evidence still needs to be trustworthy
- Example Claim: AI system can detect condition with higher specificity than the average doctor
 - Randomized clinical trial shows this is true



Calculation

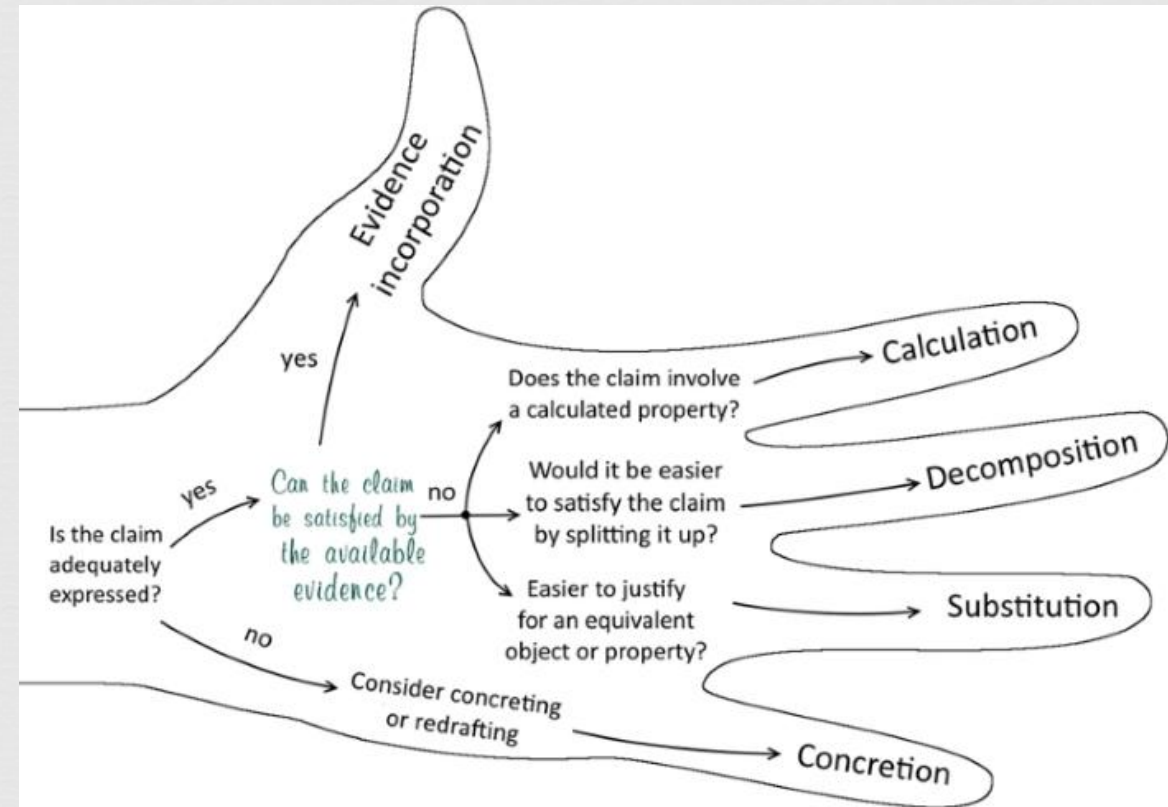


- Available evidence does not directly support the claim
- Claim contains a computed property
 - calculate the property
- Example Claim: on average, the AI system can detect condition faster than a medical professional
 - database of detection times available



Decomposition

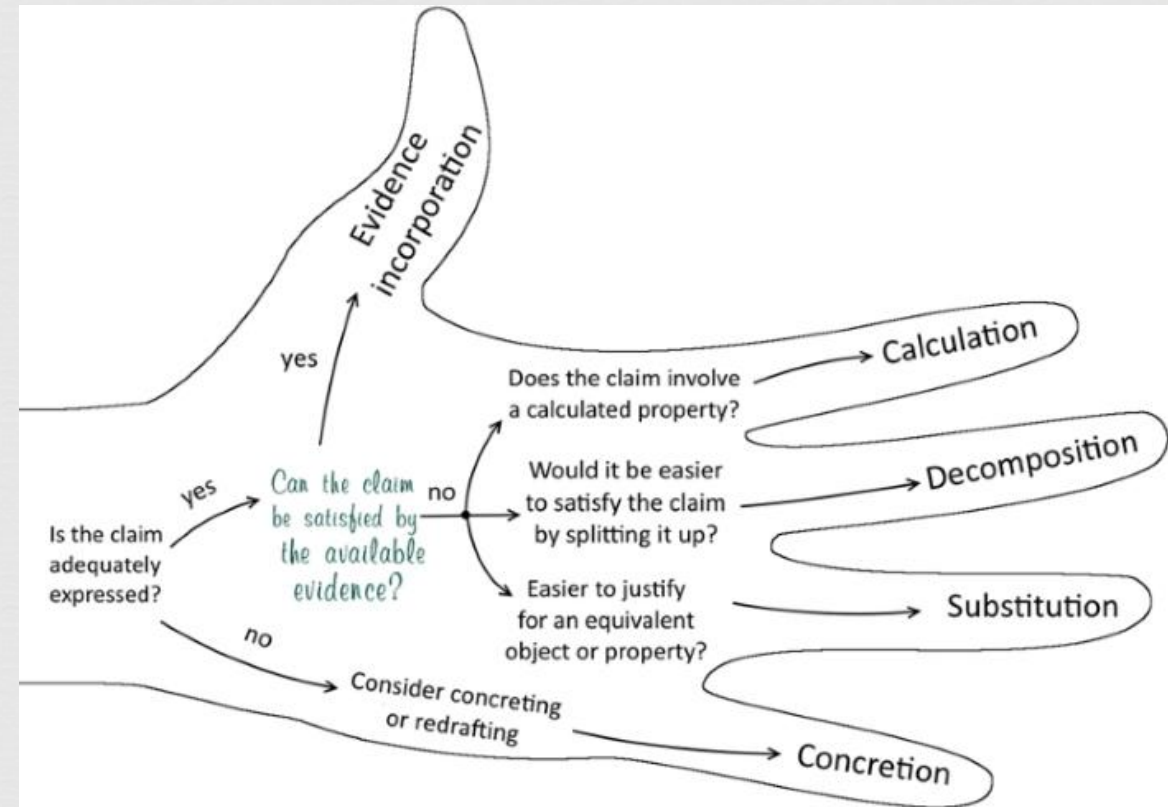
- System has a property if all parts have the property?
 - show property for all parts
 - smaller claims, easier to incorporate evidence
- Example Claim:
AI system is trustworthy
 - complies with laws and regulations
 - adheres to ethical principles
 - robust and does not cause unintended harm



Substitution



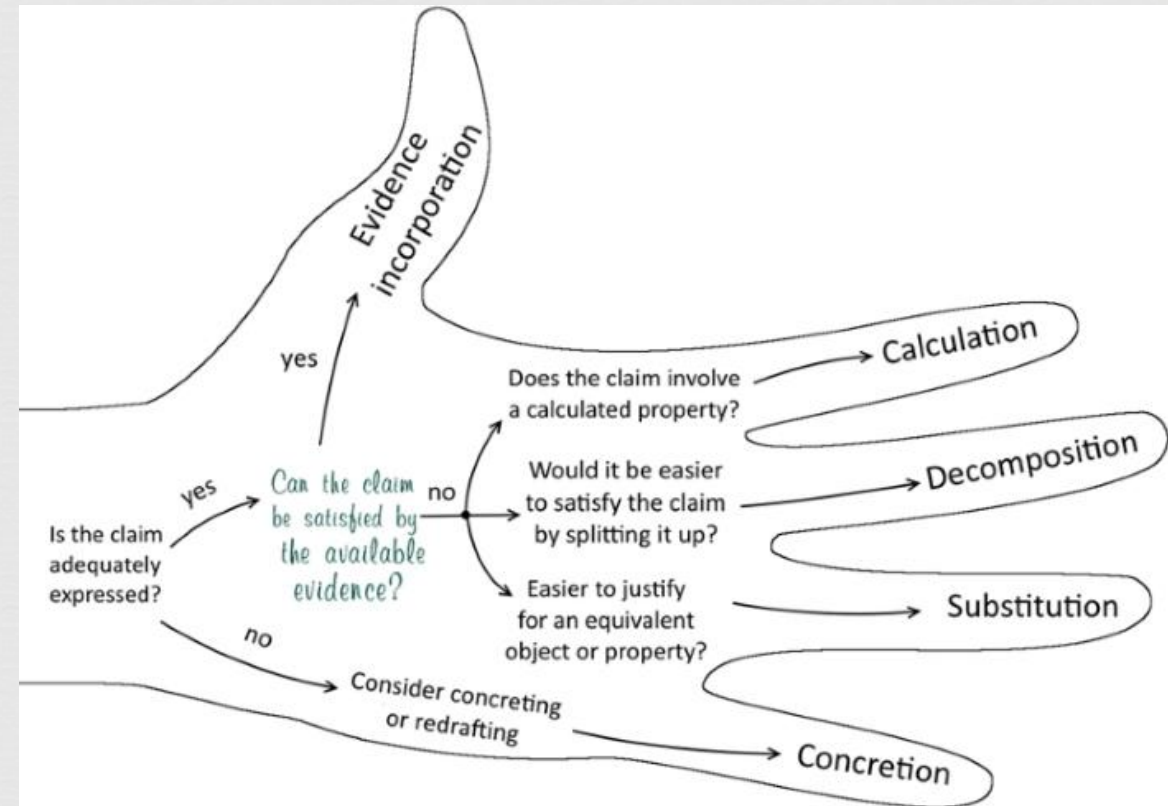
- No evidence for current system / property available
 - evidence for an equivalent system / property is available
 - use equivalent evidence
 - equivalence must be justified
- Example Claim:
AI system can detect skin cancer from images
 - Clinical study performed in the Netherlands
 - Should work as well in Germany
 - Both central European countries, skin colors are similar



Concretion



- It is not sure how evidence could be evaluated regarding the claim
 - better definition
 - better interpretation
 - more context
- Example Claim:
AI system is better than doctors
 - AI system can detect condition X on average faster and with higher accuracy than a trained doctor



How can I use this?



- Check what claims the vendor makes
 - organize them as CAE Tree in a “Mind-Map”
 - link to the relevant documents
 - use PowerPoint, paper, ...
 - Can the claims be verified / refuted?
 - What arguments and evidence do they provide to support the claims?
 - Do they provide evidence at all?
 - Do you think the evidence is trustworthy? Why?
- Critically evaluate what claims the evidence supports and what the evidence does not support
 - I.e. how generalizable are the results from a study?
- If no (good) evidence is available, that is a result!

Using CAE with the ALTAI tool



- Questions can be rephrased as claims
 - **Q.** Could the AI system affect human autonomy by generating over-reliance by end-users
 - **C.** The AI system does not affect human autonomy by generating over-reliance by end-users.
- Questions provide additional context (blue phrases, (?)s)
- Bottom up approach: Questions at the end are the important claims
 - Questions before are subclaims
 - Arguments for combining the sub-claims would be needed
- Good way to identify problems and missing information

References



- Adelard LLP, *Claims, Arguments and Evidence (CAE)*. Retrieved April 16, 2021 from <https://www.adelard.com/asce/choosing-asce/cae.html>
- Adelard LLP, *CAE Framework*. Retrieved April 16, 2021 from <https://claimsargumentsevidence.org/>
- My NASA Data, *Introduction: Building Claims from Evidence*. Retrieved April 16, 2021 from <https://mynasadata.larc.nasa.gov/lesson-plans/introduction-building-claims-evidence>