RRI Tools - Anticipate

Dependable and Deployable AI for Robotics 2

Barbara Webb

The AREA framework for RRI

- **A**nticipate outcomes
- **R**eflect on motivation, processes and products
- **E**ngage with stakeholders
- **A**ct responsively

- **"Anticipating** possible outcomes of research and innovation. This includes considering both intended and unintended outcomes, and can help you to think about why the work is worth doing, and also what might "go wrong" or cause problems in the future".
- **"Reflecting** on your (and others) motivations, processes and products. Even within a single research or innovation project we all face different and competing demands on our time and energy. By setting aside time to deliberately reflect we can help to ensure that we are still heading in a good direction and using our time well".
- **"Acting** accordingly to address issues revealed. RRI practices should make a real difference to what we do and how we do it".

https://rai.ac.uk/toolkits/rri-toolkit/

Established tools for anticipation

- Technology assessment/roadmapping
- Cost-benefit analysis
- Life-cycle assessment
- Risk assessment

Each of the above can be extended to include responsibility:

- Incorporate the wider social context and the relevant values in the assessment
- Ensure a wider group of stakeholders are considered in (and ideally, have input into) the analysis
- Allow for diverging opinions, uncertainty, qualitative decision heuristics
- Make processes more continuous to monitor alignment to (changing) social values

Example

How could responsibility considerations be introduced at each step?

Daim et al. (2018) https://www.sciencedirect.com/science/article/pii/S0040162517307734*

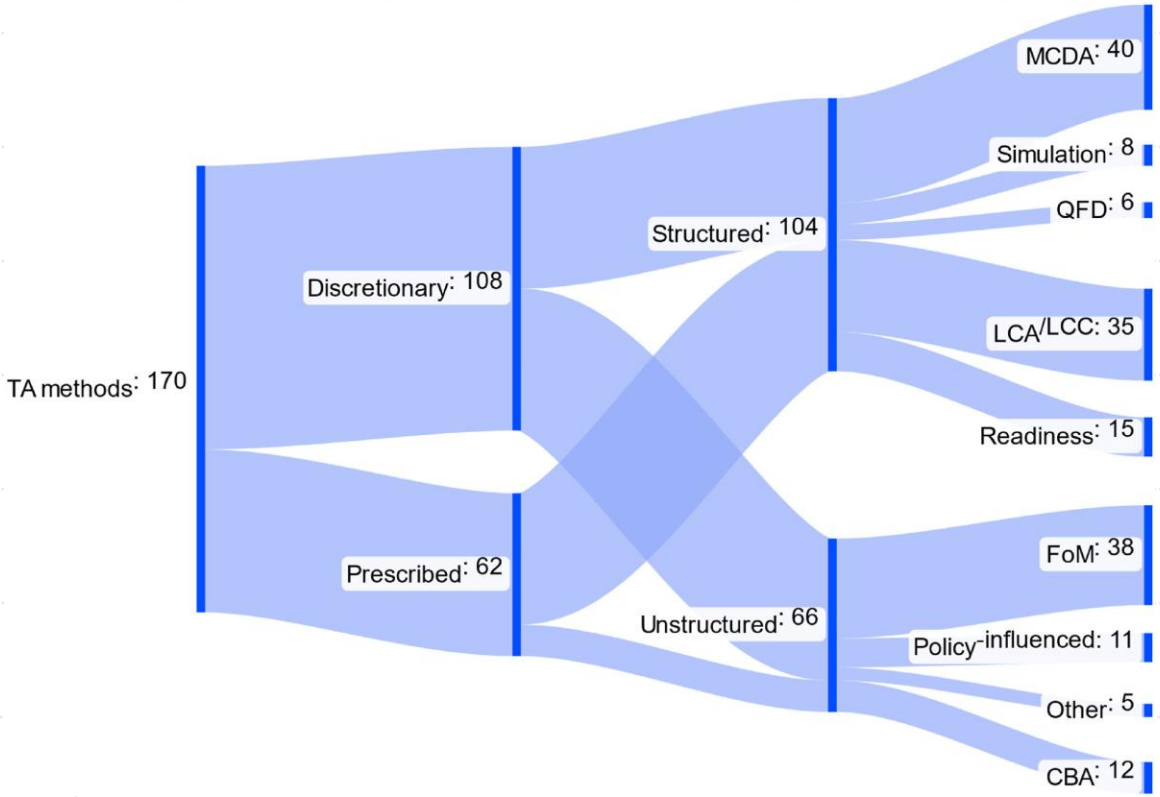
Road-mapping of robotics technology for the power industry:

- Identify experts using bibliometric analysis
- Identify technology alternatives (four robotic systems)
- Develop evaluation model based on: functionality, design (e.g. size, robustness, flexibility), technology (e.g. for positioning, precision, processing time), usability (including interface, upgradeable, maintainable), "electronic" (e.g. remote operation, visual capability, data transmission)
- Ask experts to rate alternatives (including estimate for future advances) against criteria
- Construct hierarchical decision model and calculate technology value of each alternative

**Note this was chosen as an illustrative example for the field, rather than as exemplary practice*

Technology assessment tools

- Parolin et al (2023) How can technology assessment tools support sustainable innovation? https://doi.org/10.1016/j.technovation.2023.102881



MCDA: Multicriteria decision analysis

QFD: Quality Function Deployment

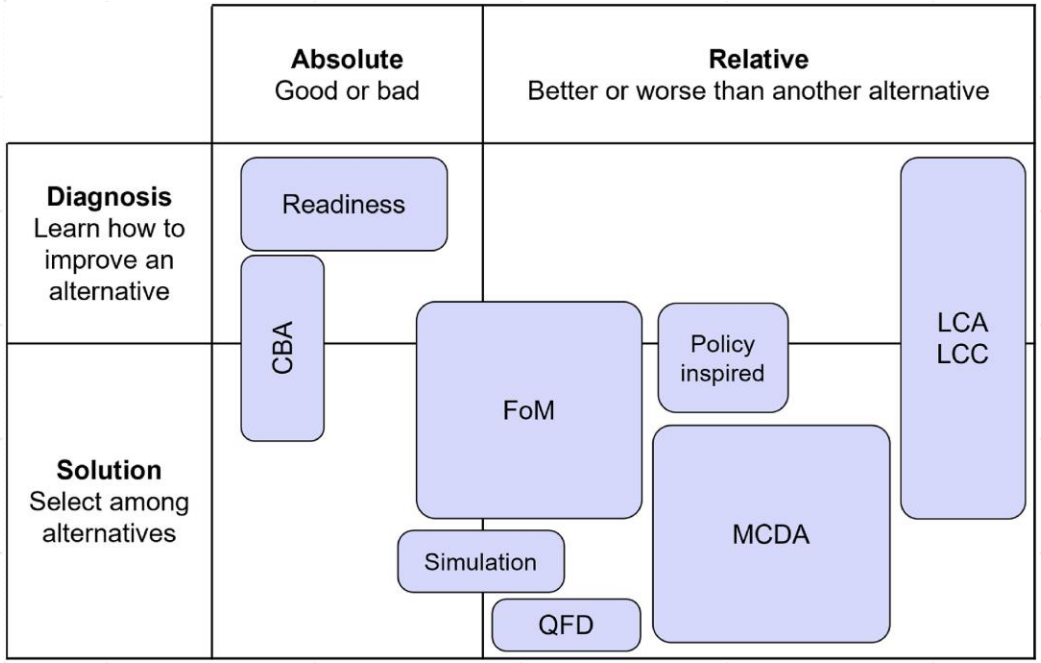
LCA/LCC: Life Cycle Assessment/Costing

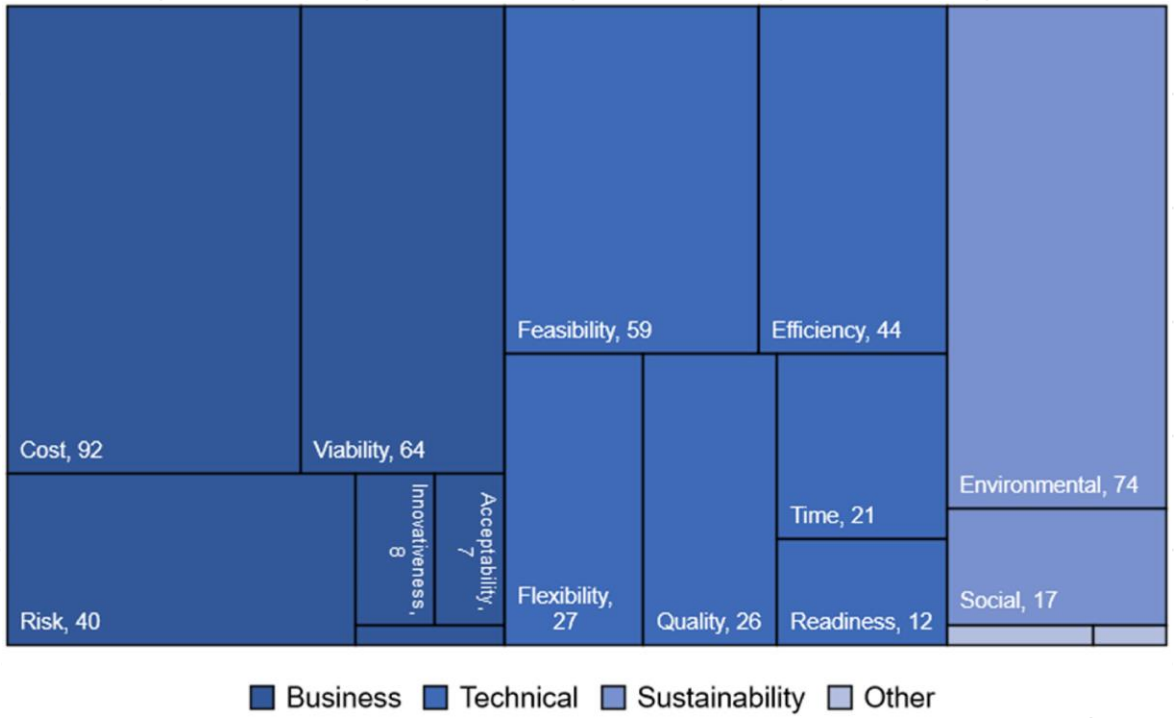
FoM: Figures of Merit (Key Performance Indicators)

CBA: Cost Benefit Analysis

Technology assessment tools

- Parolin et al (2023) How can technology assessment tools support sustainable innovation? https://doi.org/10.1016/j.technovation.2023.102881





Exercise: Tarot cards of Tech

- https://tarotcardsoftech.artefactgroup.com/

RRI Tools - Reflect

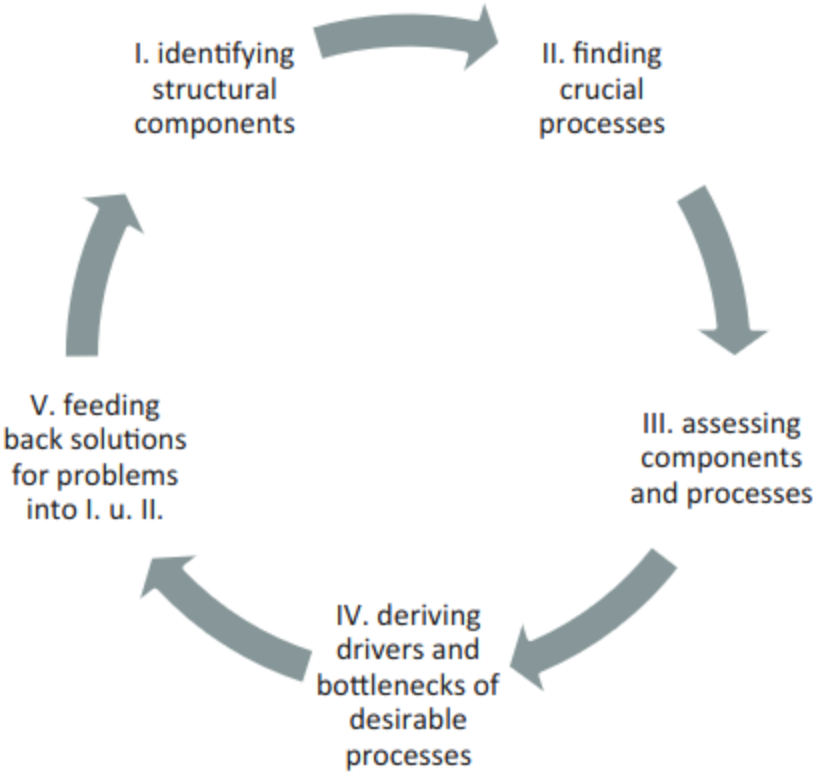
http://self-check-tool.innovation-compass.eu/faq

Dependable and Deployable AI for Robotics 2

Barbara Webb

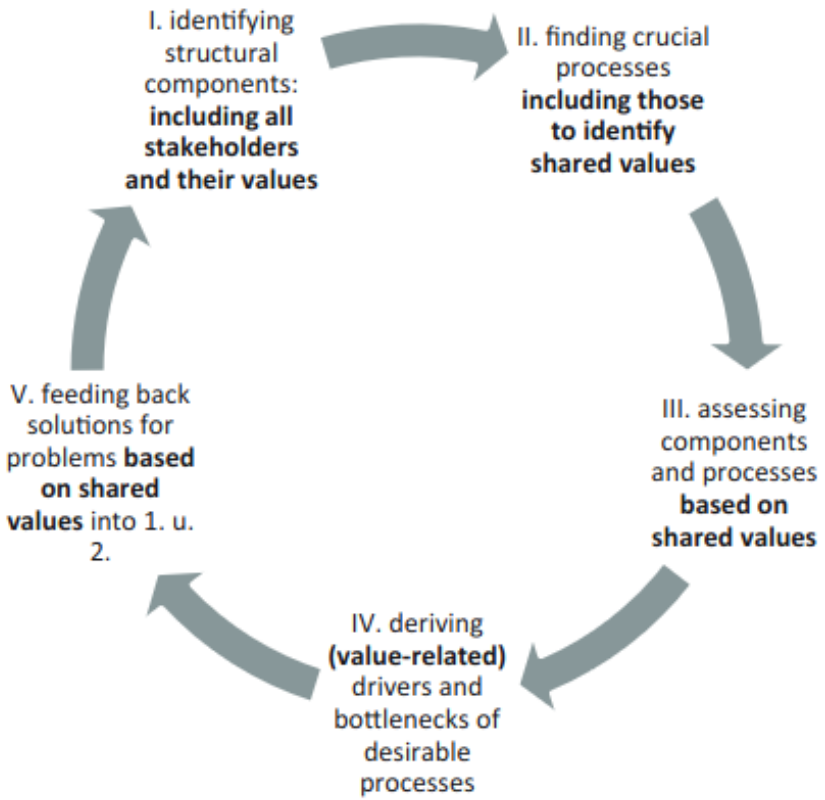
Assessing an innovation system

- From Werker 2020 https://www.taylorfrancis.com/reader/download/43374892-69da-4b07-a730-3b4c3a0d26a2/chapter/pdf
- Components includes actors, relationships, institutions
- Activities include knowledge advance and sharing, methods selection, development of markets
- Assessment might consider: actors missing, or lack capability, relations too weak or strong, missing formal institutions; lack of knowledge development or sharing; lack of processes to identify best structures or solutions



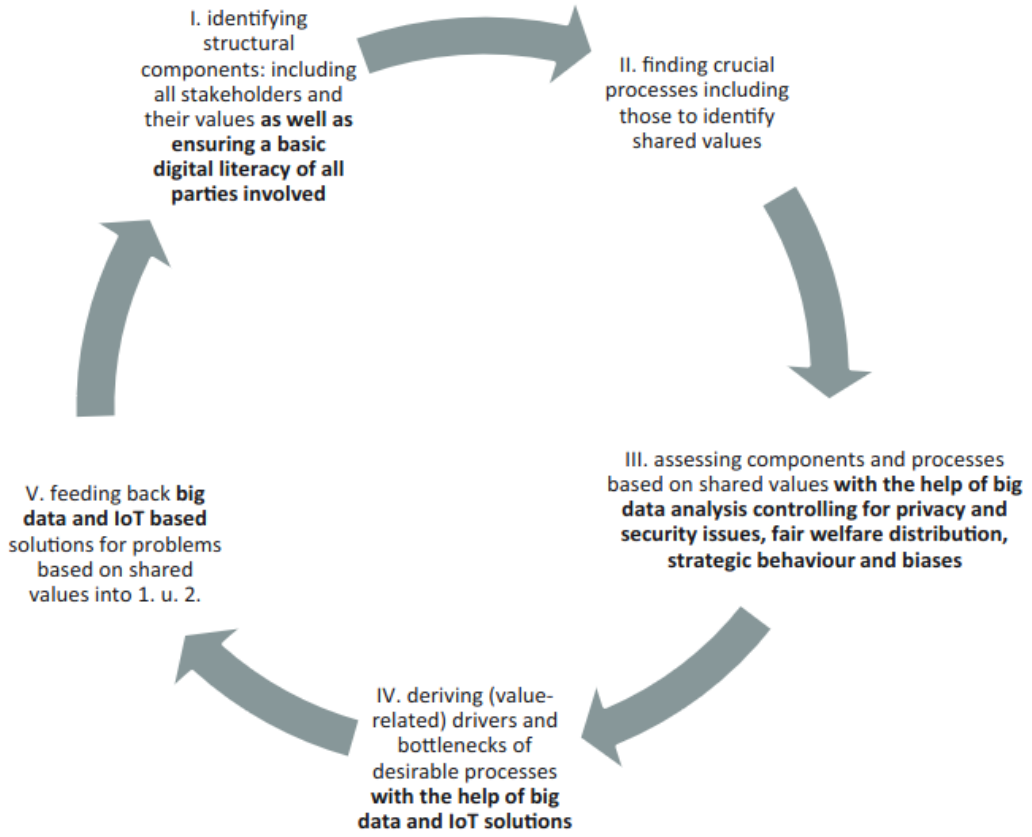
Assessing Responsible Innovation

- From Werker 2020 https://www.taylorfrancis.com/reader/download/43374892-69da-4b07-a730-3b4c3a0d26a2/chapter/pdf
- Additional problems:
- Not including all stakeholders and all their values
- Lack of processes to collaborate towards shared values
- Failure to compensate stakeholders disadvantaged by the innovation
- Solution-finding approaches don't account for values



Using AI tools to assess Responsible Innovation

- From Werker 2020 https://www.taylorfrancis.com/reader/download/43374892-69da-4b07-a730-3b4c3a0d26a2/chapter/pdf
- The process of assessment could be enhanced by AI tools for analysing complex systems
- But use of AI tools introduces further considerations related to their *responsible* use



Informal reflection on your own research

- Why are you doing this research? Why is important that you carry on? Who is it for?

Exercise: write a letter to a relative or friend who does not understand what you do, explaining and justifying your research.

- Are there people/groups or organisations who might be opposed to the work? Are there reasons not to do it? Think in advance how you might respond to such objections.
- Who might be excluded from the benefits of the work? Can it be made more accessible?
- What priorities, privileges and biases might you be bringing to the work?
- Self-assessment: COMPASS http://self-check-tool.innovation-compass.eu/faq

Formal methods

https://www.gov.uk/guidance/portfolio-of-ai-assurance-techniques

	Technique type								
Stage	Impact assessment	Compliance audit	Certification	Bias audit	Conformity assessment	Impact evaluation	Formal verification	Performance testing	Other ongoing testing
Scoping	Yes	Yes							Yes
Data gathering and preparation	Yes	Yes	Yes						Yes
Modelling and design	Yes	Yes	Yes		Yes		Yes	Yes	Yes
Development	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Deployment	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Live operation and monitoring	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Retirement		Yes				Yes			Yes

RRI Tools – Act

Dependable and Deployable AI for Robotics 2

Barbara Webb

Barriers to action

Lukien et al (2024) Embedding responsible innovation into R&D practices: A case study of socially assistive robot development https://doi.org/10.1016/j.jrt.2024.100091

- Project actively engaged in anticipation, e.g., identified need to safeguard human dimension, avoid unintentional deception, ethical and legal aspects of recording in home environment; awareness that AI algorithms might exacerbate issues of human oversight, privacy and transparency.
- Few follow-up actions in the research activity: 'too late' to do much in the design to address issues; lack of explicit obligations (team structure diluted engagement); other tasks prioritised, not embedded in workplan, postponed for later consideration.
- Some codesign activities but tended to focus on substantiating existing assumptions about design, with lack of reflexivity about researcher's values and interests.

Practical considerations

- What is the right balance of time/effort/money to spend on anticipation, reflection and engagement?
- How can these be efficiently and effectively integrated into your research plan?
- Within a team, how directly should *you* be involved in these processes?
- How should you monitor, assess and report the effect of anticipation, reflection and engagement on your research?

Potentially useful:

https://thinkingtool.eu/tool

https://tas.ac.uk/responsible-research-innovation/rri-prompts-and-practice-cards/

Preparing a responsible action plan

Early stages:

- What additional information do you need and how will you get it?
- Are you up-to-date on the relevant background and context?
- Are you aware of the relevant sociotechnology configuration and the actors within it?
- Have efforts been made to give a role to societal values, needs and expectations in defining the research aims?
- Have diverging opinions on the problem/approach been considered?
- Have you tried to identify long-term as well as short-term impacts?
- Can you identify the most suitable methods for anticipation, reflection and engagement?

Later stages:

- - How can you build in reflection on impacts; and alignment to and changes in societal values, needs and expectations during the project?