Responsible Research and Innovation (RRI)

Adjunct Professor &
Principal Scientist
Mika Nieminen
mika.nieminen@vtt.fi

Contemporary democratic theory and governance

RESPONSIBLE AI AND ROBOTICS
AN ETHICAL FRAMEWORK

MATHEMATICS AND STATISTICS

Top 10 innovations

1. Printing press
2. Light bulb
3. Airplane
4. Personal computer
5. Vaccines
6. Automobile
7. Clock
8. Telephone
9. Refrigeration
10. Camera

The Benefits of Digital Health

Digital health helps Canadians access better quality care, more efficiently. Electronic medical records help doctors and nurse practitioners manage patient difficulties, such as chronic diseases, more efficiently and effectively. Electronic health records provide a comprehensive, digital view of a patient's health history. This means duplicate tests are avoided because authorized clinicians can access past results of tests ordered by other health care providers. This ultimately leads to faster diagnosis and treatment.

3 key AI benefits for the future of work

Most threats to humans come from science and technology, warns Hawking

Speaking ahead of his BBC Reith Lecture on black holes, Stephen Hawking discusses the danger inherent in progress and the chances of disaster on Earth

▲ Stephen Hawking on falling into a black hole - audio

Risks from Artificial Intelligence

Recent years have seen dramatic improvements in artificial intelligence, with even more dramatic improvements possible in the coming decades. In both the short-term and the long-term, AI should be developed in a safe and beneficial direction.

Social Media has become a Direct Threat to Democracy

Social Engineering

- Using knowledge of human behavior to elicit a defined response.
- *Put simply…getting you to willingly do something for me which is likely not in your best interest.*

Integrity

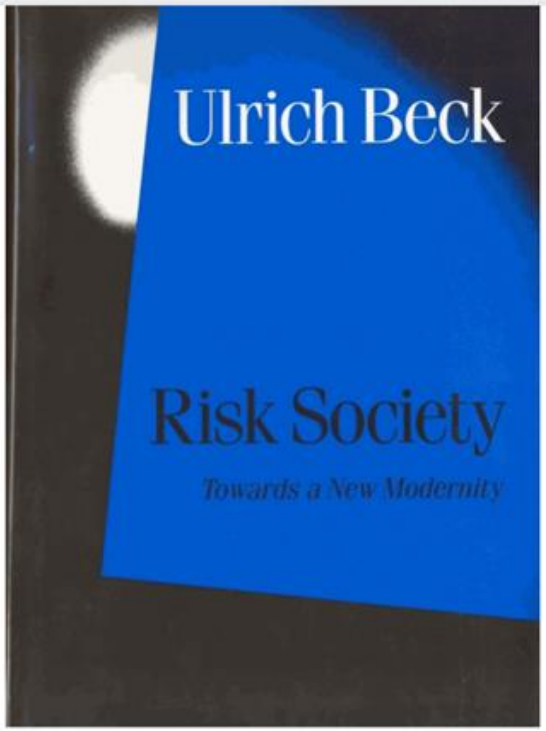
The giant garbage vortex in the Pacific Ocean is over twice the size of Texas — here's what it looks like

Follow @EINordic | Follow @EINordic | 3,008 followers

Kevin Loria | 23 Mar 2018 6:48 PM | 2285

The five biggest threats to human existence

1. Nuclear war
2. Bioengineered pandemic
3. Superintelligence
4. Nanotechnology
5. Unknown unknowns



- The society should make use of science and technology so that it increases the quality of life and does not cause harm to anyone
- Melvin Kranzberg's (1986) first law of technology ”Technology is neither good nor bad; nor is it neutral” Nor science and innovation are absolutely neutral arenas.
- Future is not predestined. Instead we are continuously designing our future. The decisions made today formulate our future in all levels.

Some History for RRI

- Continuous discussion on social impacts and risks of science goes hand in hand with modern science (e.g. Mary Shelley´s Frankenstein 1816 & debates on science-society relationships in pre 2nd WW UK & debates on atomic bomb after the war)
- More systematic attempts to assess risks of research and innovation emerge in the 1960s. In the U.S. was developed the idea of Technology Assessment (TA)
- The term “TA” was then defined as “sociotechnical research that discloses the benefits and risks to society emanating from alternative courses in the development of scientific and technological opportunities.” (Wong 2014, 223) Later various versions of TA were developed (e.g. parliamentary, expert, participatory, and constructive TA).
- There are also various other approaches like bioethics, technology ethics, AI ethics, ethical technology design, and Ethical, Legal, and Social Aspects (ELSA) research of genomics and nanotechnology research

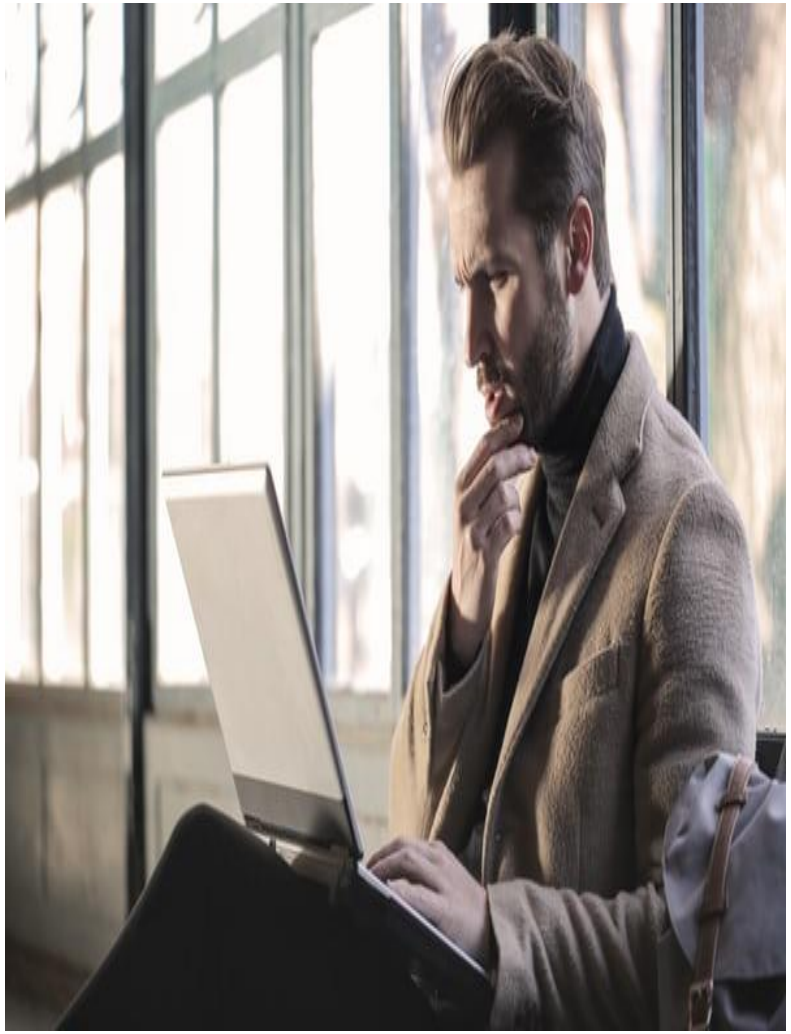
Some History for RRI

- While various forms of TA etc. continue to exist, during the recent decade more visibility has gained Responsible Research and Innovation (RRI)
- Term is used especially by EU , but the idea of ethical and responsible reserch and innovation gathers together wide community of researchers with various journals & academic positions
- EU has attempted to operationalize the idea especially through Horizon2020 framework program
- Thus, during recent years in a number of EU projects specific “toolboxes” for applying RRI have been developed, ideas of RRI governance have been studied, and evaluation frameworks for RRI have been created

The Rationale

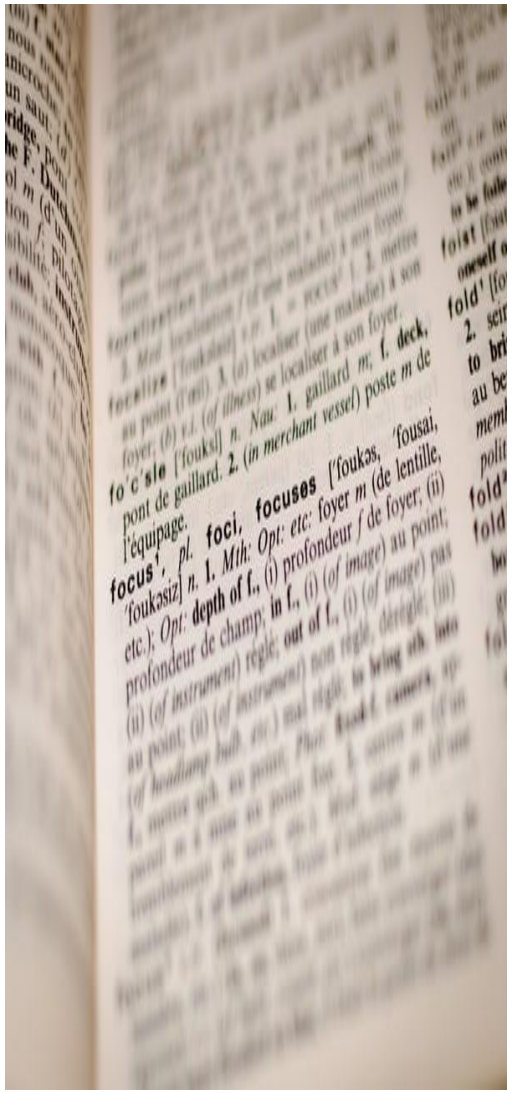
- Strengthens **acceptability, desirability and ethicalness** of R&I; improves **social effectiveness**; brings in **new perspectives**
- Strengthens **dialogue on the desirable society** and its characters
- Changes people´s position in innovation process from users or consumers to **active political and moral subjects**

- **In innovation and business:**
  - Supports socially and environmentally sustainable products and services
  - Increases trust on products and their desirability and acceptability
  - Decreases business risks and undesirable impacts
  - Strengthens business and company brand

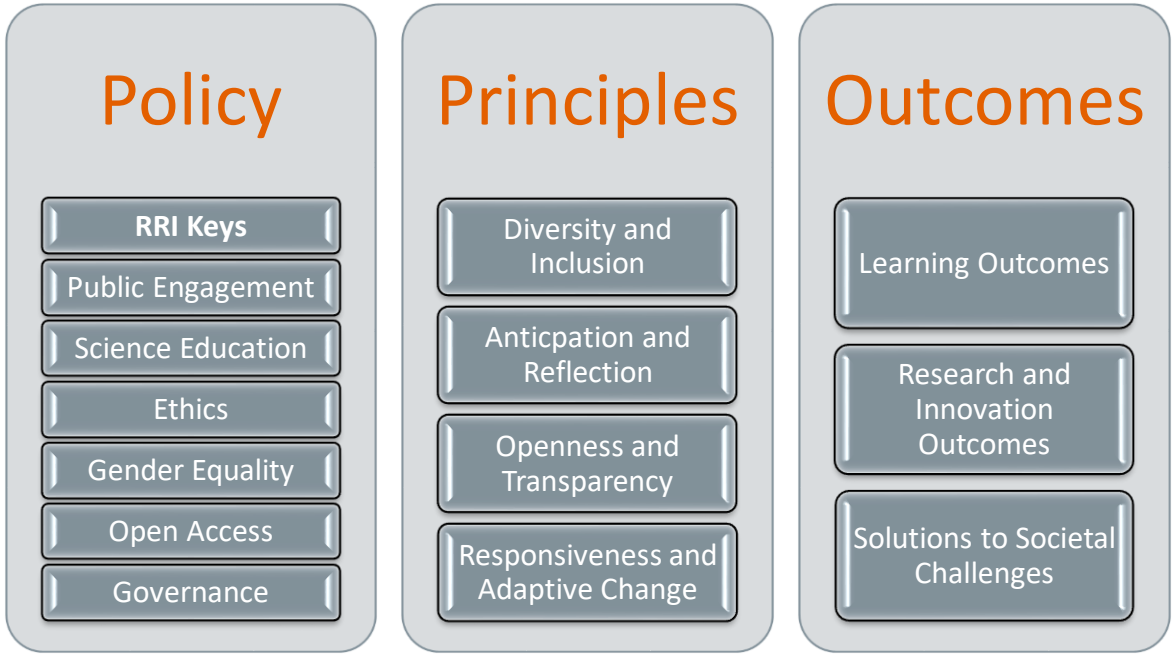


The Concept

- There are various definitions of RRI, but they each share a number of common characteristics, including e.g. focus on social challenges, engagement of stakeholders, opening up of research and innovation to society, and risk avoidance (Smallman 2018; Gurzawska 2017)
- For instance: “RRI is an inclusive approach to research and innovation (R&I), to ensure that societal actors work together during the whole research and innovation process. It aims to better align both the process and outcomes of R&I with the values, needs and expectations of European society.” (European Commission, 2013)
- Owen & al. (2013) see that responsible innovation includes:
  - **Anticipation**: Analysis of the social, economic and environmental impacts of innovation activity
  - **Reflexivity:** Each actor should consider their own underlying motivations and purposes for participating in the innovation activity openly
  - **Inclusiveness:** Brings into the common discussion various stakeholder and citizen interests, values and perspectives.
  - **Responsiveness**: Learning and changing of target-setting and operative practices.



Normative concept: E.g. »Six Keys», RRI Principles & Outcomes

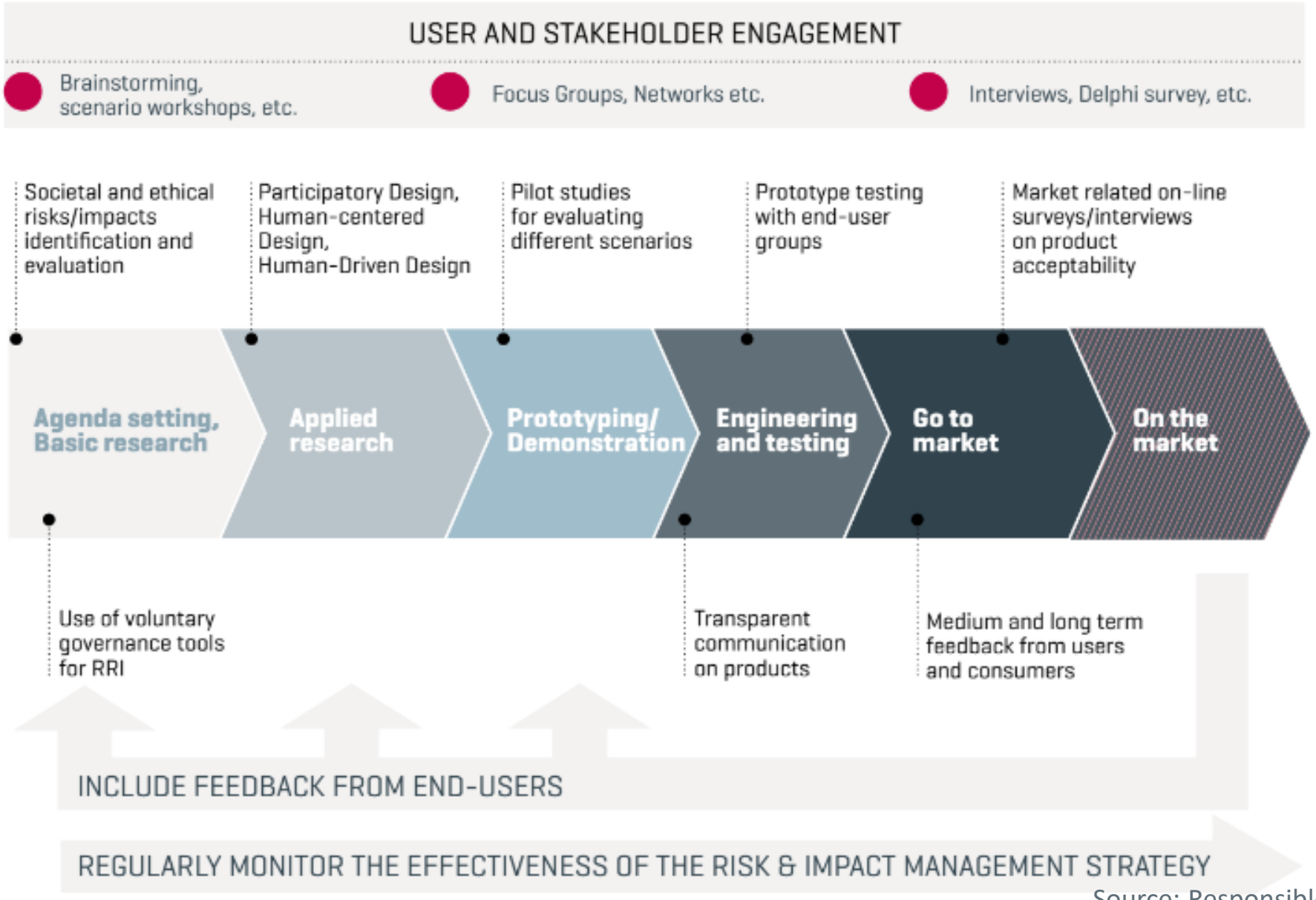


An example of normativity:

- For science and innovation actors different aspects important; different targets, values, economic and other aspects can be in contradiction…
- Can also be considered as a ”Eurocentric” concept
- Has faced criticism and provoked suggestions for better societal ”contextualization” e.g. economic situation, indigenous knowledge…

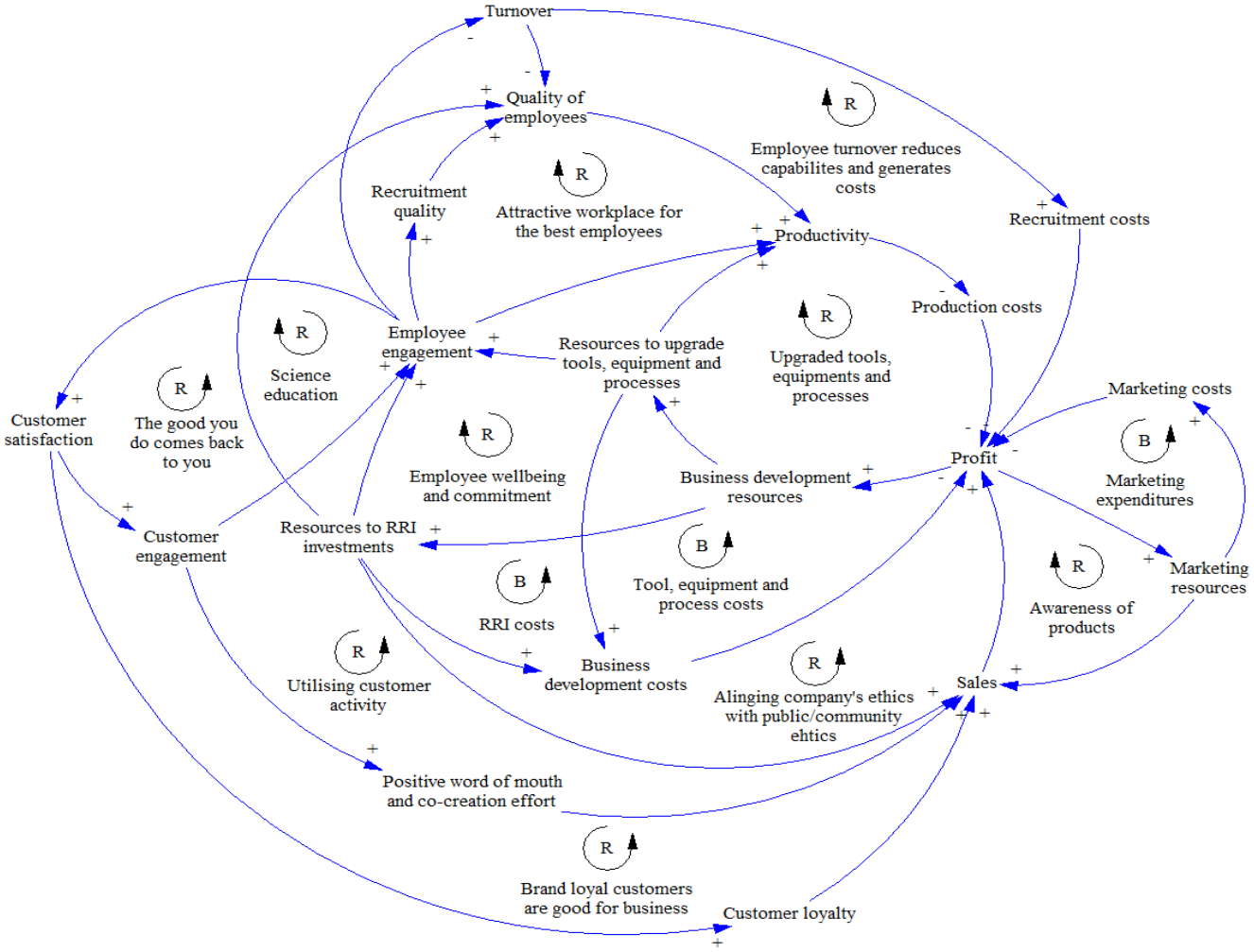
RRI - Areas to consider to build trustworthiness?

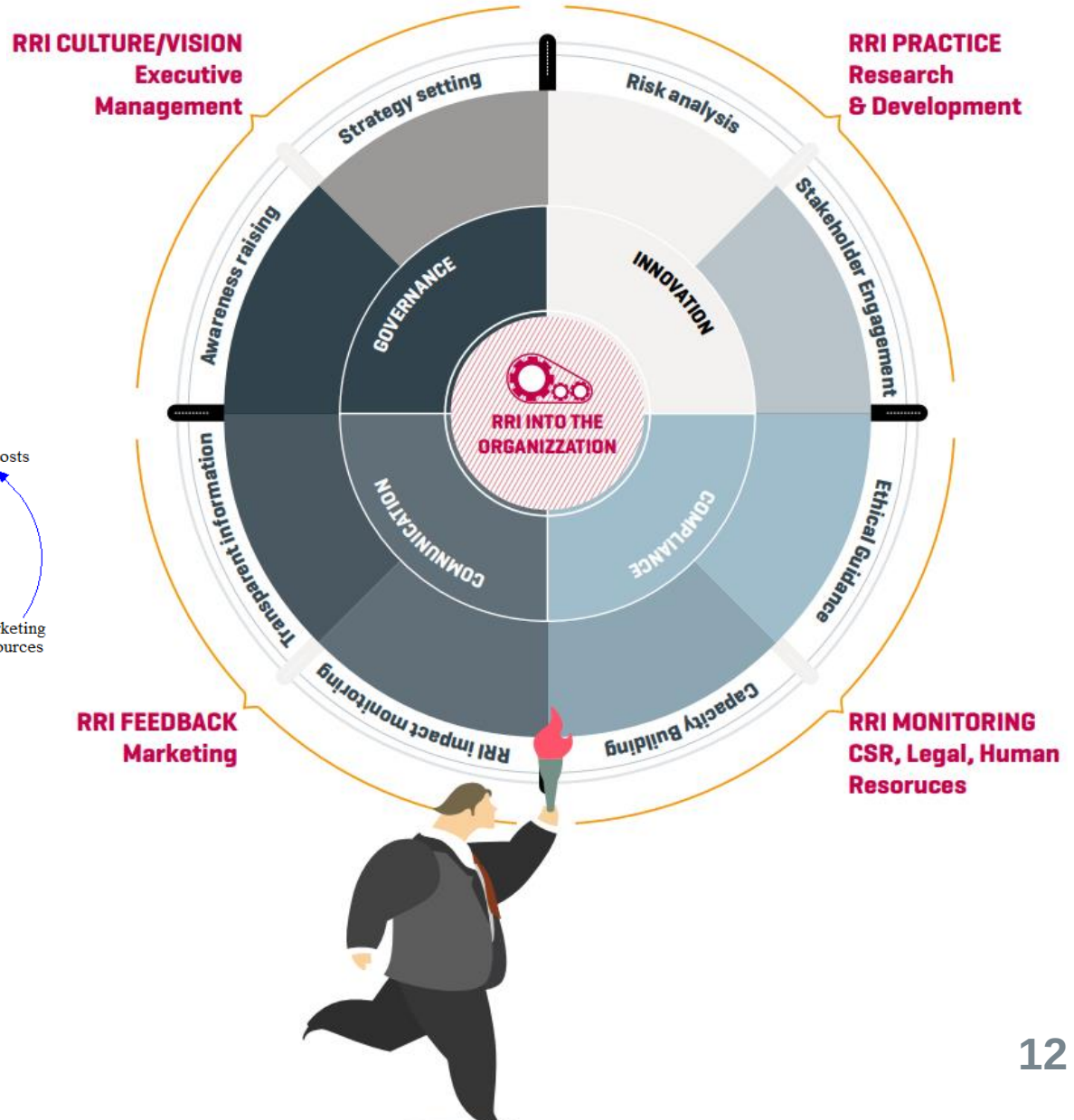
People	People/society centric design & problem solving	New & existing health & safety risks/issues	Inclusion, diversity, gender, equality, respect	Livelihoods, jobs, wealth distribution, labour standards	Human rights, inc respect, data security, labour standards	Collaboration of stakeholders in design & dev
Environment	Eco benefit / superiority by design	Eco safety - tox, impact, by design	Resource use & reduction	Biodiversity threats	Reversibility and uncertainties?	Potential cumulative impacts
Society	Impact on social cohesion/justice	Distribution of benefits	Cumulative potential impacts	Livelihoods, jobs, dignity, employment	Unintended consequences	Creating resilience to negative impacts
Culture	Cultural benefits considered	Negative cultural impacts considered	National /cultural differences / impacts considered	Religious differences / impacts considered	Human dignity / privacy considered	Animal rights considered
Economic	Economic benefit clarified	Negative Econ impacts considered	Cumulative econ impact considered	Negative econ impacts - response?	Economic implications of RRI issues considered	Resilience to negative econ impacts considered
Ethical	Does it do what you say it will? Show evidence	Wider social, ethical impacts considered	Dilemmas & trade offs considered	Openness re trade offs & priorities	Involvement of others	Collaboration of stakeholders in ethical decisions
Governance	Explore new needs for regulation, standards, insurance considered	Risk assessment thorough & broad & responsive	Oversight mechanisms adaptive, anticipatory, appropriate, adequate	Planning/action where oversight not fit for new tech	Openness & transparency eg-use of tech, HSE, research findings, data, access	Whistleblowing encouraged & supported



Source: Responsible Industry

Responsible Industry





Open Questions

- Who has the ”final word”, and what kind of compromises are legitimate (e.g. in medical research): Challenge for the freedom of research and self-organization? Is this a restricting procedure?
- Is it causing biases for research and innovation activities?
- Who has a possibility to participate in the process? What are the possibilities of “ordinary citizens”, is there a danger to become an elitist process (associations etc. represent citizens)?
- Is there a danger that research and innovation processes will be ”politicized”? Who has an opportunity to steer research and innovation? Can it bypass democratic steering of R&I system?

Various sources for further studies

- RRI in H2020: https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation

- Links to RRI projects and their results: https://www.parrise.eu/other-rri-projects/

- When ja how to use RRI: http://satoriproject.eu/
- RRI-Tools: https://www.rri-tools.eu/
- Responsible industry: http://www.responsible-industry.eu/
- Prisma toolkit: http://www.rri-prisma.eu/toolkit/

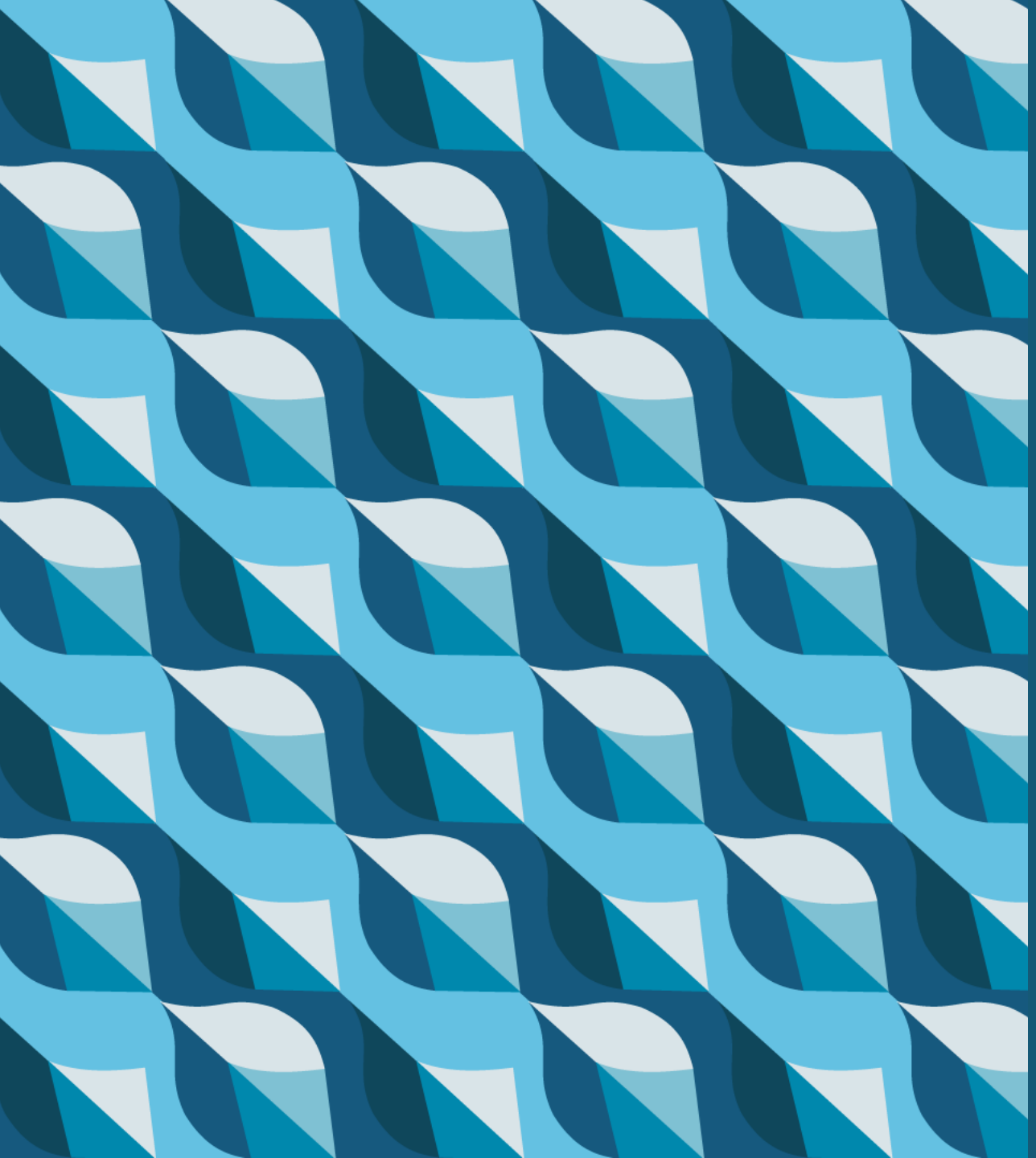
- Journal of Responsible Innovation: https://www.tandfonline.com/loi/tjri20

- Some books to start with:
- Owen R., Bessant J., Heintz M. (Eds.) (2013) Responsible Innovation. Oxford: Wiley.
- Saariluoma, P., Cañas, J.J., & Leikas, J. (2016) Designing for Life - A human perspective on technology development. London: Palgrave MacMillan.

Practical exercise

https://www.thinkingtool.eu/

- Pairs or you can be also on your own
- Use as a concrete thinking piece your own study
- Prepare to discuss on your ideas/findings on RRI in relation to your own work after the session
- Go to the thinking tool and proceed step by step as the program advices: Click “try the tool”, “create new project”, name it (copy the code if you want to continue later), choose gate 1 or 2, choose an entry point, choose key and conditions, choose questions and answer them on the basis of your own study



Thank you!