

Responsible Research and Technology Assessment

Julia Hahn
Institute for Technology Assessment and Systems Analysis



On the Agenda

- Technology and society
- Technology Assessment
- Challenges of Assessing
- Methods of Technology Assessment
- Responsible Research and Innovation
- Assessing Big Data – Insights from a project

- Effects of technologies are the **outcome of human decisions and activities** in a specific context (“intended or unintended consequences”)
- Assumption: **Technology systems are results of social processes**
- Basis question: How can **societal challenges** be solved by technologies without necessarily increasing **unintended consequences**?

Natural AND Social Sciences

Research AND Advice

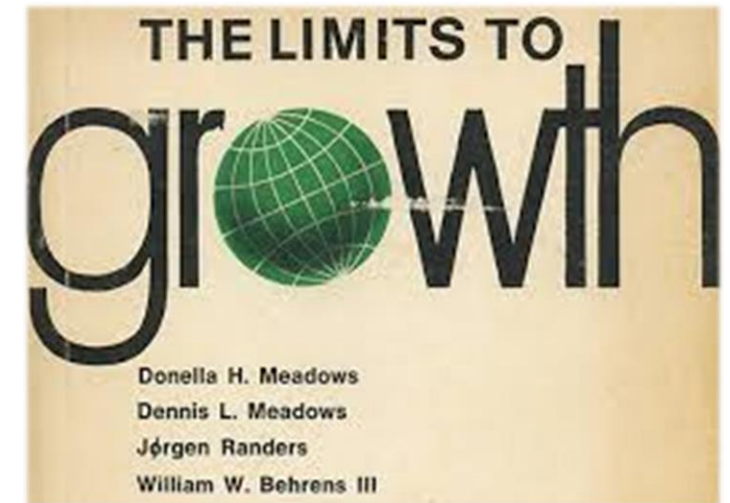


Technology AND Society



A Bit of Historical Context...

- **Technological progress** increasingly criticized, e.g. sustainability
- **Investments** in research and development increase
- Risks often **underestimated or ignored**
- Need for an **early warning system** to identify potential hazards and minimize effects
- “it is essential that, to the fullest extent possible, the **consequences of technological applications be anticipated, understood, and considered** in determination of public policy on existing and emerging problems” (US Senate 1972)



Technology Assessment on the Rise

- 1960s: TA as a 'technocratic' response to technocratic problems: TA as experts, assessments as reaction to progress and complex effects

*"TA is class of **policy studies** which **systematically examine the effects on society** that may occur when a technology is introduced, extended or modified..."* J.F. Coates, 1980

- Science and technology are social processes, embedded in society: TA is problem-oriented

*"TA is a **scientific, interactive and communicative process** which aims to contribute to the **formation of public and political opinion** on societal aspects of science and technology."* TAMI, 2003

- Central for TA:

- Knowledge of possible consequences for various stakeholders
- Knowledge for decision making processes
- 'Better' developments and innovation for society



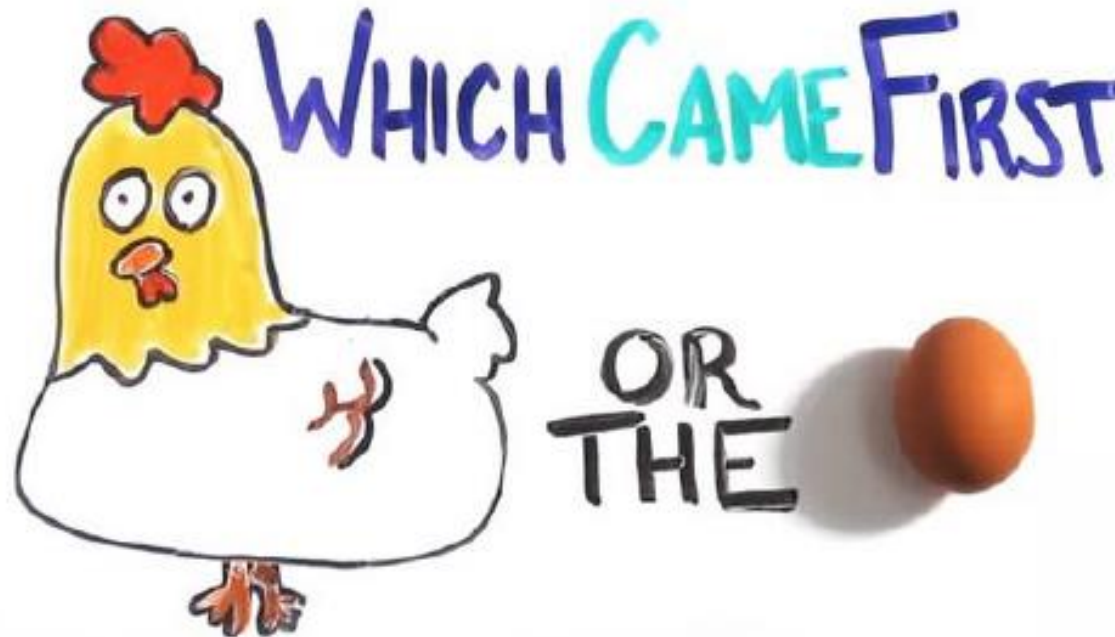
Challenge: Ambivalenz

- TA is **not about** providing recommendations to **morally justify** or a certain technology or not
- Instead it is about **balancing options**, needed for this: knowledge and assessment
- Assessment of **effects** and **weighing of options**



Challenge: Complexity

- The positive and negative effects of technology are determined by **societal frames and contexts**
- Often unclear, which **cause is responsible for which effect**
- **Effects are different** for different groups, individuals, societies



Challenge: Uncertainty

- Technology effects can lead to **very different reactions** by different individuals, stakeholders, etc.
- Next to classic measurement errors **uncertainty also comes from non-knowledge**
- Non-knowledge due to assessment of only **certain systems** or **unknown external factors**
- When do we **start assessing**?



Dealing with Challenges....

- Inclusion of **public and stakeholders** in assessments
- Ethical considerations often based on **culture, values, identity, emotions**, etc.
- Facts don't speak for themselves: **framing of issues and narratives** important for providing evidence-informed policy advice
- Technologies are often connected to “**controversies about risk and benefits** and **ethical disputes** about human dignity, the common good and questions of responsible research” (Zhao et. al 2015: 40)
- Engagement as method to **include ethical considerations and current discussions** based on values, etc. in assessment



Classical TA

- Expert orientated / focus on risks
- Attempt to 'rationalise' debate (threat diffusion)
- One-way relationship between TA and public
- Examples: TAB, ITAS

Participatory TA

- Non-expert / public inclusion
- Focus also on 'value'
- Attempt to create 'coherent' debate
- Variety of experimental models
- Examples: DBT, RATHENAU

Scientific methods:

- Expert discussion
- Delphi method
- Modelling / simulation

Interactive methods:

- Consensus conference
- Citizens' jury
- Scenario workshop



Communication:

- Newsletter
- Articles in press
- Video presentation



Responsible Research and Innovation

- “RRI implies that **societal actors** (researchers, citizens, policy makers, business, third sector organisations, etc.) work together during the **whole research and innovation process** in order to better align both the process and its outcomes with the **values, needs and expectations of society**” (EU Commission)
- Ideally: through engagement problems of acceptance are solved by including all relevant actors during the entire development process
- Stakeholders should “obtain relevant knowledge [...] to effectively evaluate both outcomes and options in terms of **societal needs and moral values** [...] to use these considerations [...] as **functional requirements for design and development** [this] should be established as a **collective, inclusive and system-wide approach**” (EU Commission 2013).



USER AND STAKEHOLDER ENGAGEMENT



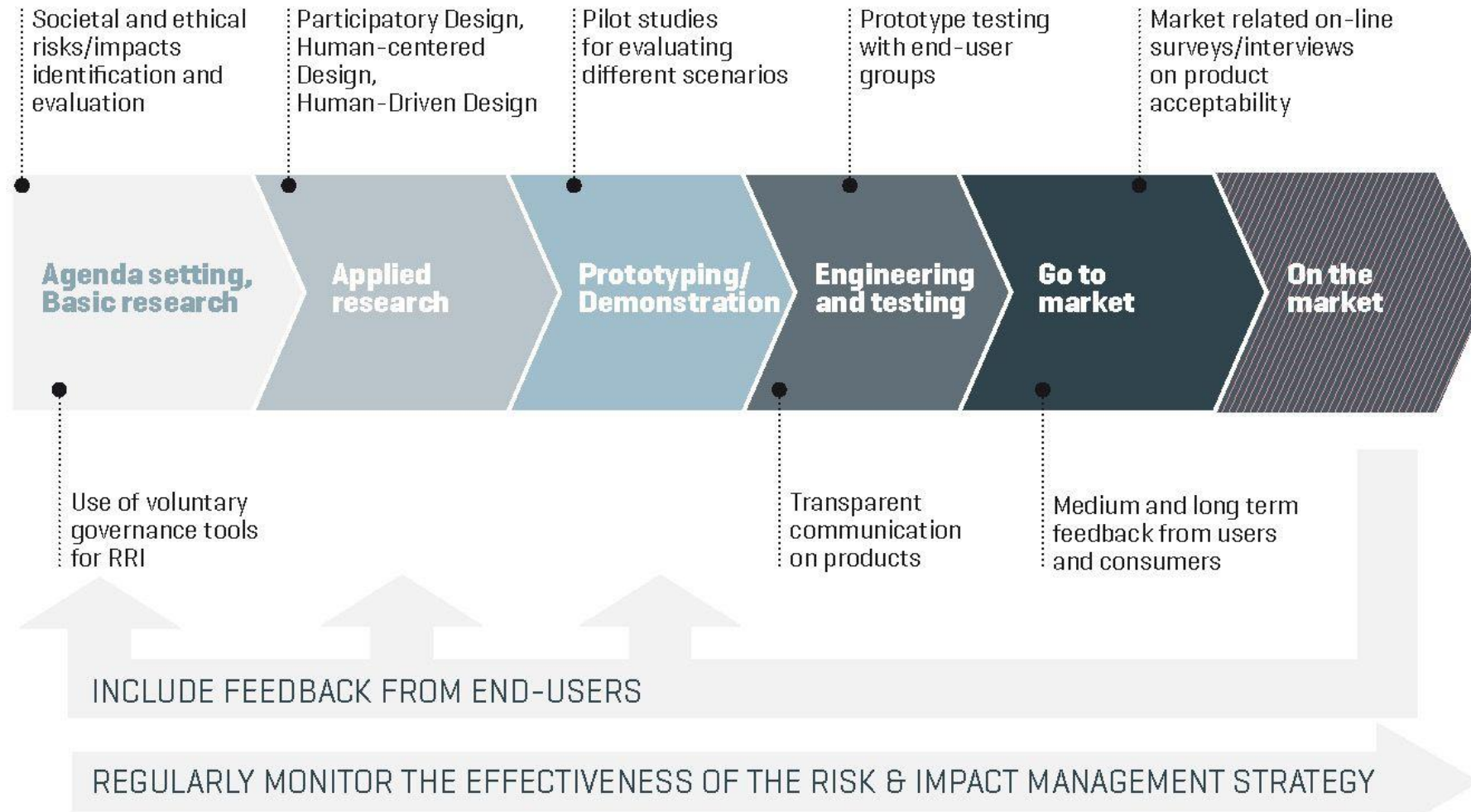
Brainstorming,
scenario workshops, etc.



Focus Groups, Networks etc.



Interviews, Delphi survey, etc.



- What are the **developments in Big Data** technologies and applications and **what trends** can be expected in the near future? What **changes** will be brought about in the **economy, in services, in government action, or private life**?
- What do **citizens** know about Big Data and what is their assessment of the developments?
- What are the **relevant issues of social change** triggered by Big Data? What do **individual disciplines** know about it, and how are Big Data developments judged by them?
- What conclusions can be drawn for **options for political and social action**?



RESEARCH QUESTIONS



MONITORING



SCENARIOS, TRENDS, NARRATIVES



WORKING GROUPS



CONFERENCE



DELPHI SURVEY



CITIZENS' CONSULTATIONS



POPULATION SURVEY



IN-DEPTH STUDIES



EXPERT WORKSHOPS



FOCUS GROUPS



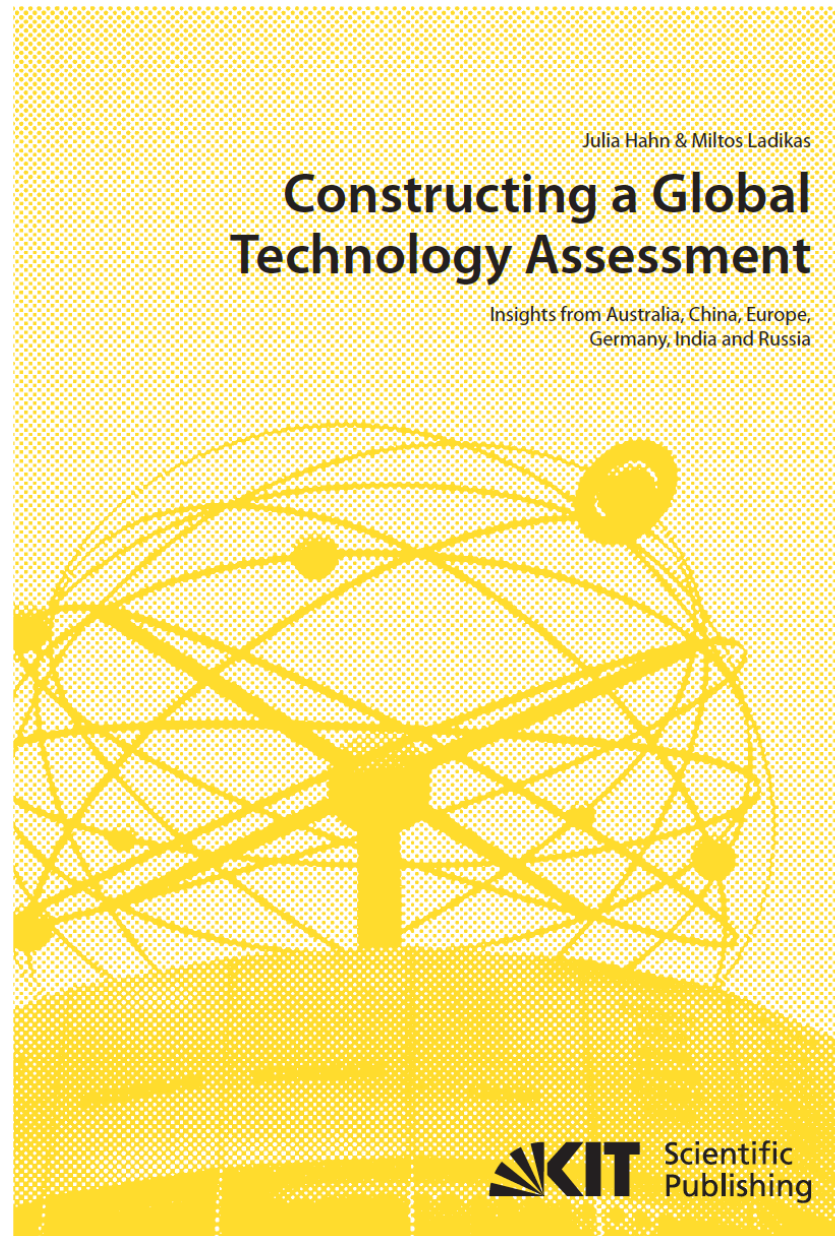
SYNTHESIS REPORT



- Expert workshop: **“Discrimination through algorithms in AI and eRecruiting”**
 - **More transparency required** regarding comprehensiveness of factors, not entire algorithm but instead decisions made and the **parameters** used
 - Algorithms can help **show interdependencies and uncover structural discrimination**
 - **Often unclear what companies regard as a “good employee”** (soft factors) and **correlation between prejudices and confidential factors** (gender, religion, etc.)
- Expert workshop: **“Digital Demos - Big Data in political campaigns”**
 - **Journalistic coverage ambivalent:** big data either as **new power** encompassing all or as **exaggeration**
 - Often unclear whether fascination is with **content or function of big data**, e.g. presidential campaign in USA: often reported how heterogenic the data sets and the resulting prognosis were → shift from actual content of prognosis to a **judgement of outcome** (big data use for Clinton positive, for Trump negative)
 - Need for a **dialogue-form of education** with inclusion of public for open dialogue
 - **Big data itself a form of politics** encompassing and behind each algorithm or digital claim there is a **interest or intension**
 - Need for **minimal consensus and discourse on democracy** in the context of social media

What Now?

- Insight: **ethical, social, cultural aspects** should be part of technology assessments and reflection
- Challenges such as complexity, ambivalence, uncertainty require **different forms of gaining knowledge**
- TA and RRI are approaches that aim to find answers → **mix of methods** and **problem-oriented focus**
- All-encompassing developments such as big data require the **inclusion of an array of disciplines** (interdisciplinarity) and **actors** (transdisciplinarity)



<https://publikationen.bibliothek.kit.edu/1000085280/23225323>