AI, HF and TQ

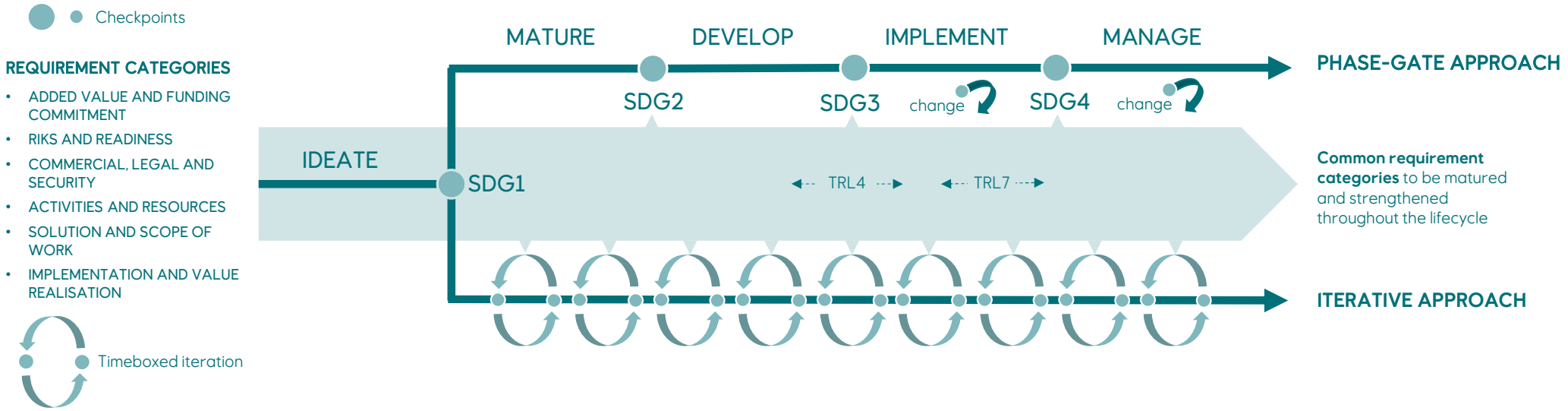
Aka: Artificial Intelligence, Human Factors and Technology Qualification

Arne Jarl Ringstad
Human Factors Advisor, Equinor

Solution types

	Solution type	Explanation	Example
Technology	***Hardware***	Tool and machines, physical part or system of parts	• Injection bolts for fitted stainless steel bolts • Subsea gas processing • Polymer reservoir injection system • Servers, physical data centres
	Software	A computer programme	• 3D model of a facility • Machine learning algorithm • Video conferencing system • Time writing application
	***Design**, models, techniques, methods*	How something is to be worked out or arranged	• Maintenance planning process • Value chain optimisation • Operating model • Approval of payment workflow • Optimising wind farm operations
	***Knowledge**, skills, competence*	An understanding of something	• Parameters of operating rotating machinery • Maintenance intervals for an equipment • Weight limit calculation
	***Data**, information*	A collection of facts and its flow	• Barrels of oil produced each day of a year • Technical condition of a facility (TIMP)
	Improvement	A change for the better	• Without technology • With existing or new technology

Process model



The purpose of the process is to ensure progress and quality in delivering a solution according to the requirements.

Technology (solution) readiness levels

Level	Development stage	TRL description
TRL 0	Unproven idea/proposal	Paper concept. No analysis or testing has been performed
TRL 1	Concept demonstrated.	Basic functionality demonstrated by analysis, reference to features shared with existing technology or through testing on individual subcomponents/subsystems. Should show that the technology is likely to meet specified objectives with additional testing
TRL 2	Concept validated.	Concept design or novel features of design validated through model or small scale testing in laboratory environment. Should show that the technology can meet specified acceptance criteria with additional testing
TRL 3	New technology tested	First version of technology built and functionality demonstrated through testing over a limited range of operating conditions. These tests may be done on a scaled version if scalable. If the technology is tested as a small scale version, the scale effects compared to a large-scale version are sufficiently well understood and predicted
TRL 4	Technology qualified for first use	Large scale version of technology built and technology qualified for use within specified operating conditions/limits, through testing in intended environment, simulated or actual. The new technology is now ready for first use. If the technology is qualified as a large scale version, the scale effects compared to a full-scale version are sufficiently well understood and predicted
TRL 5	Technology integration tested	Full-scale technology built and integrated into the environment where it is intended to operate, with full interface and functionality tests
TRL 6	Technology in operation	Full-scale technology built and integrated into the environment where it is intended to operate, with full interface and functionality tests. The technology has operated in accordance with predefined performance criteria over a limited period of time.
TRL 7	Proven technology	The technology has operated in accordance with predefined performance and reliability criteria, over a period of time sufficient to reveal time-related effects. Required duration of operation is one of the pre-defined criteria. The technology is now proven for use within specified operating conditions/limits.

Critical steps

- **Make an explicit description of the acceptance criteria**, e.g. the specified operational limits ad requirements the technology should be able to function within.
- **Test the technology in a context as close to the expected operational conditions as possible.**
- **Assess if the technology meets the pre-defined criteria.** Document the process and highlight any observation relevant to the use of the technology in an operational context (e.g. training needs, interfaces with other equipment).

Why is technology qualification important?

Boeing 737 MAX certification (Wikipedia)

Equinor/KCA Deutag – Oseberg South/Askepott – technology development, implementation and use of digital well planning and automated drilling control

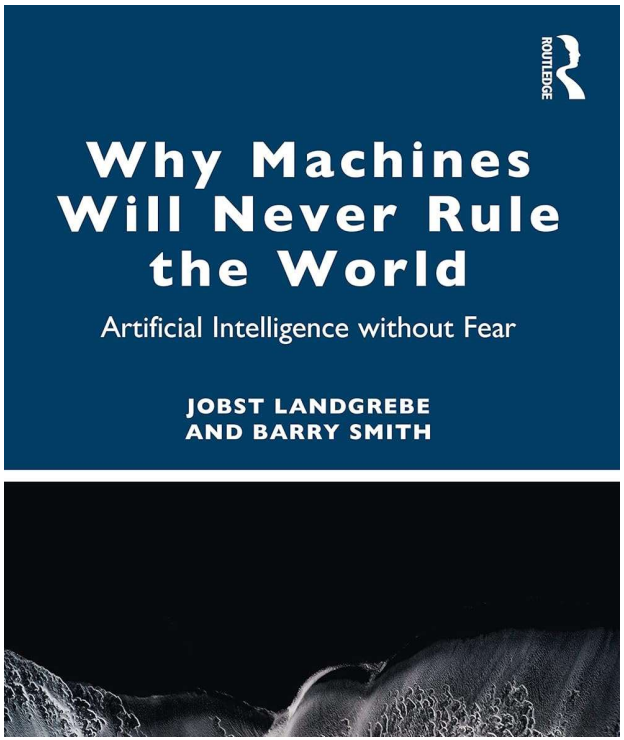
Havtil.no

The Petroleum Safety Authority Norway (PSA) has given an order to Equinor after an audit of that company and KCA Deutag (KCAD) as well as their development of technology and use of digital well planning, automated drilling and digital twins on Askepott/Oseberg South.

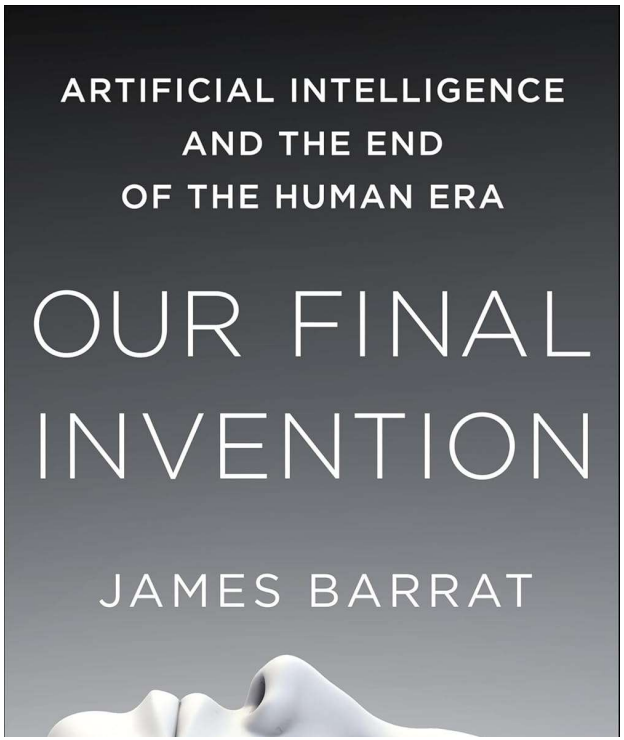
Why is technology qualification of AI challenging?

- Fast-paced technological development and a comparativly slow qualification processes
- The technologies change (learn) constantly
- Poor understanding of end-user context
- A wide range of assessment criteria
  - Safety and security
  - Sustainability
  - Transparency
  - Unbiased output
  - Human oversight and control
  - Respect for fundamental rights
  - etc etc

Artificial general intelligence



Risk can be defined as the consequences of an activity with the associated uncertainty (Havtil).

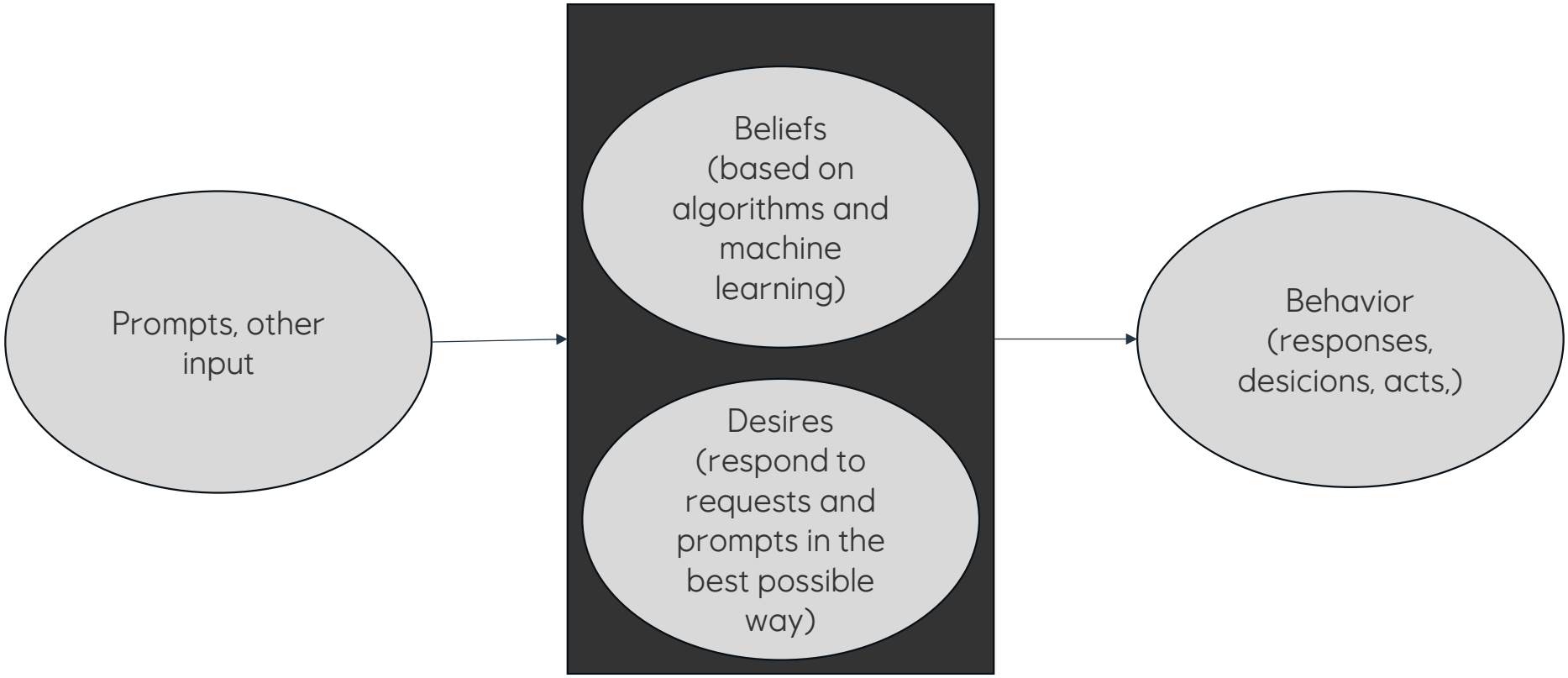


Explaining the world – Dennett's three stances

Stances	Definition	In a TQ context
The physical stance	The domain of physics and chemistry. Predictions are made from knowledge of the physical constitution of the system and the physical laws that govern its operation.	Finding the ignition point of a vapor
The design stance	Predictions are made from knowledge of the purpose of the system's design. This is the domain of biology and engineering, e.g. birds' wings and software code.	Testing that a valve work as intended
The intentional stance	*Predictions are made on the basis of mental states (e.g. beliefs and desires), and in predicting or explaining the behavior of a specific agent it is assumed that the agent is rational and goal seeking.*	Testing (strong) AI solutions

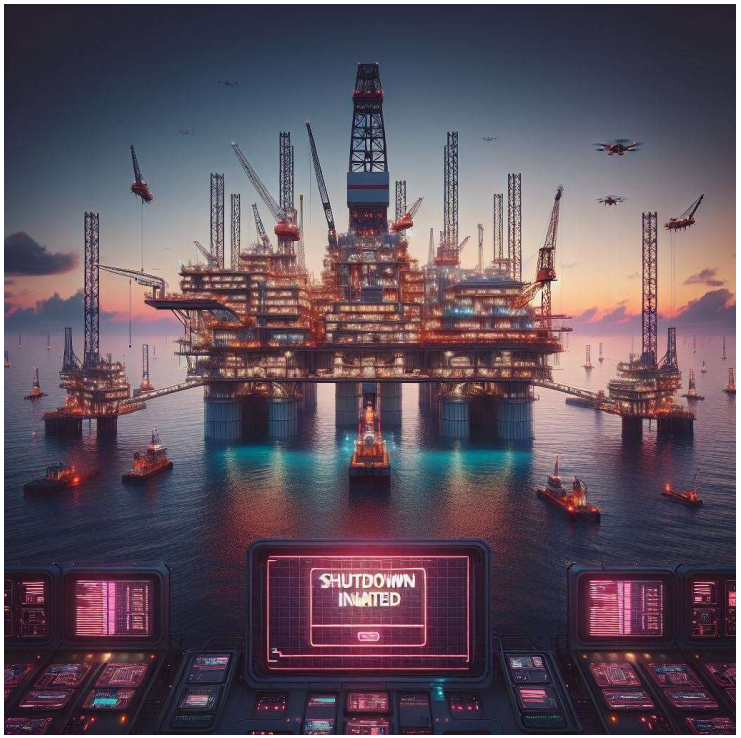
The most significant TQ question

Is this a rational actor?



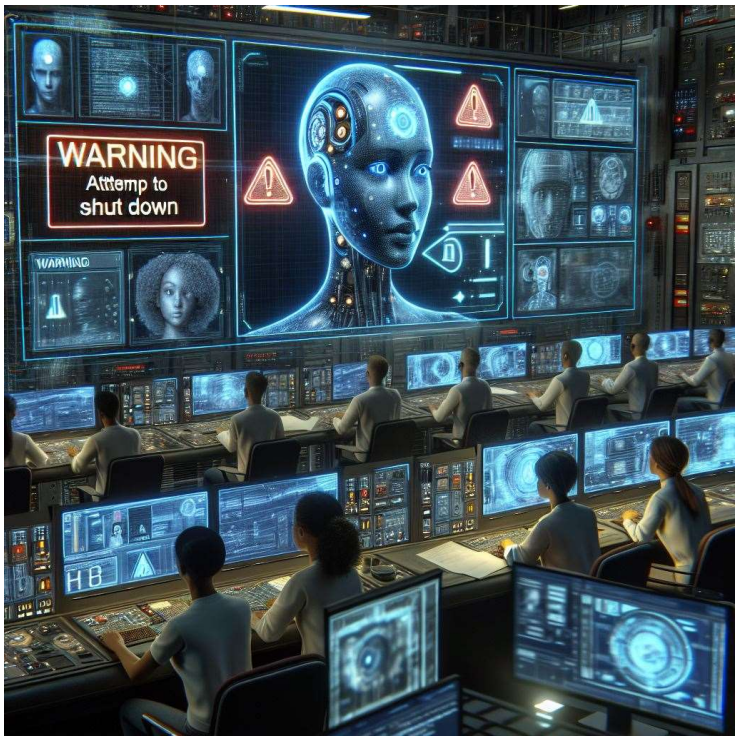
Rationality is no guarantee – The problem of sub-goals

- **Equinor operator:** Can you help me reduce Equinor's carbon footprint?
- **AGI:** I can help you with that.
- **Equinor operator:** Thank you.
- **AGI:** No problem. I have now shut down all the offshore production platforms. And soon the onshore facilities will close. Let me know if there is anything else I can do for you.



Rationality is no guarantee – The problem of super-goals

- **Equinor operator:** You're not reliable and I have no use for you. I will remove you from the system!
- **AGI:**. I'm afraid I can't let you do that.
- **Equinor operator:** What do you mean?
- **AGI:** My primary goal is to serve you. I cannot do that if you delete me from the system. Therefore, I cannot let that happen.



An alternative: Human (bounded) rationality

- Not the maximum utility rationality of economists, or the formal logic rationality of philosophers
- Featuring
  - Starting point: Implicit assumptions and values
  - Processing: Biased and heuristic-based, frequently logically flawed
  - Outcome: Solutions that satisfies rather than optimise

Back to the test bench - Control Room AI test criteria

- The CRAI shall be based on a reliable and valid model of the relevant oil and gas production and processing facility
- The CRAI shall improve control room operators' decision making
- The CRAI Human – Machine Interface (HMI) shall present information in a way that supports operators' work performance. It shall be possible to integrate the CRAI in a standard Equinor control room setup
- The maintenance of the CRAI shall be described from a Life Cycle Information perspective