Trust and safety in the context of complex AI systems: challenges related to verification and validation

November 9th , 2023

Rodica Mihai

NORCE Energy

ADRForum– Versailles, 2023

Agenda

- Introduction
- Challenges for verification and validation in complex AI systems
- Lessons learned from implementation of a complex AI system in a high-risk environment
- Conclusion

Introduction

- Trust
- Safety
- Assumptions
- Domain related constraints
- Interoperability: data and functionalities

Introduction

Trust in the system:

- ✓ Trust that the algorithm makes the correct decision under the correct assumption.
- ✓ Understand what happened when the algorithm failed.
- ✓ Predictions in line with expectations.
- ✓ Knowing that the algorithm performs badly on a specific set of data is very useful (thorough knowledge of limitations is very important).

Introduction

NORCE

Safety:

- ✓ Safety of the system.
- ✓ Safety of the environment.
- ✓ Safety of the process.

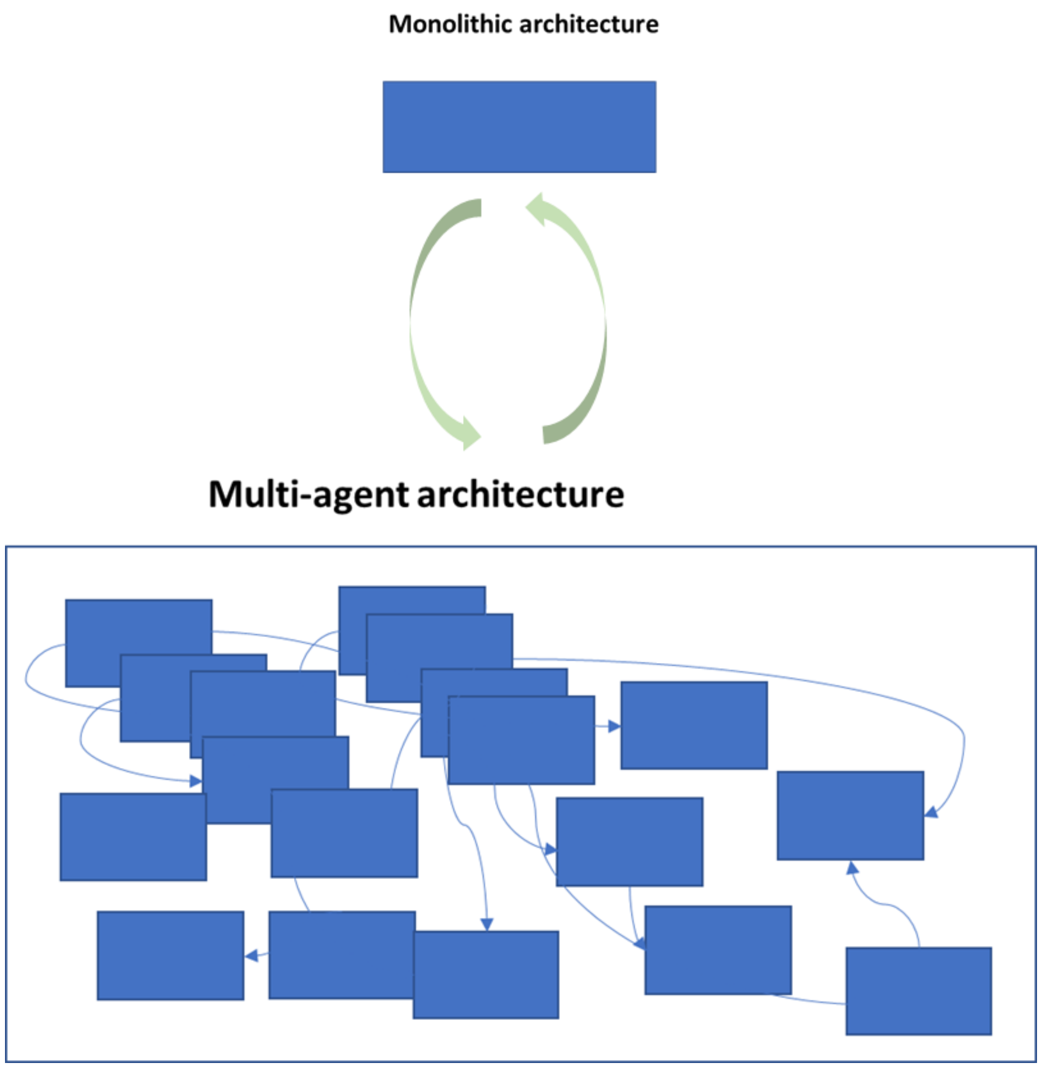
Introduction

- How much data is needed for the method to work.
- Specific method features (assumptions: example – population remains the same).
- React to abrupt changes in environment.
- Sensitive decision.
- Context is not fixed (different companies, sensors, etc.).
- Time and space sparsity of the data.
- Continuous real-time functioning without interruptions.

Challenges for verification and validation in complex AI systems

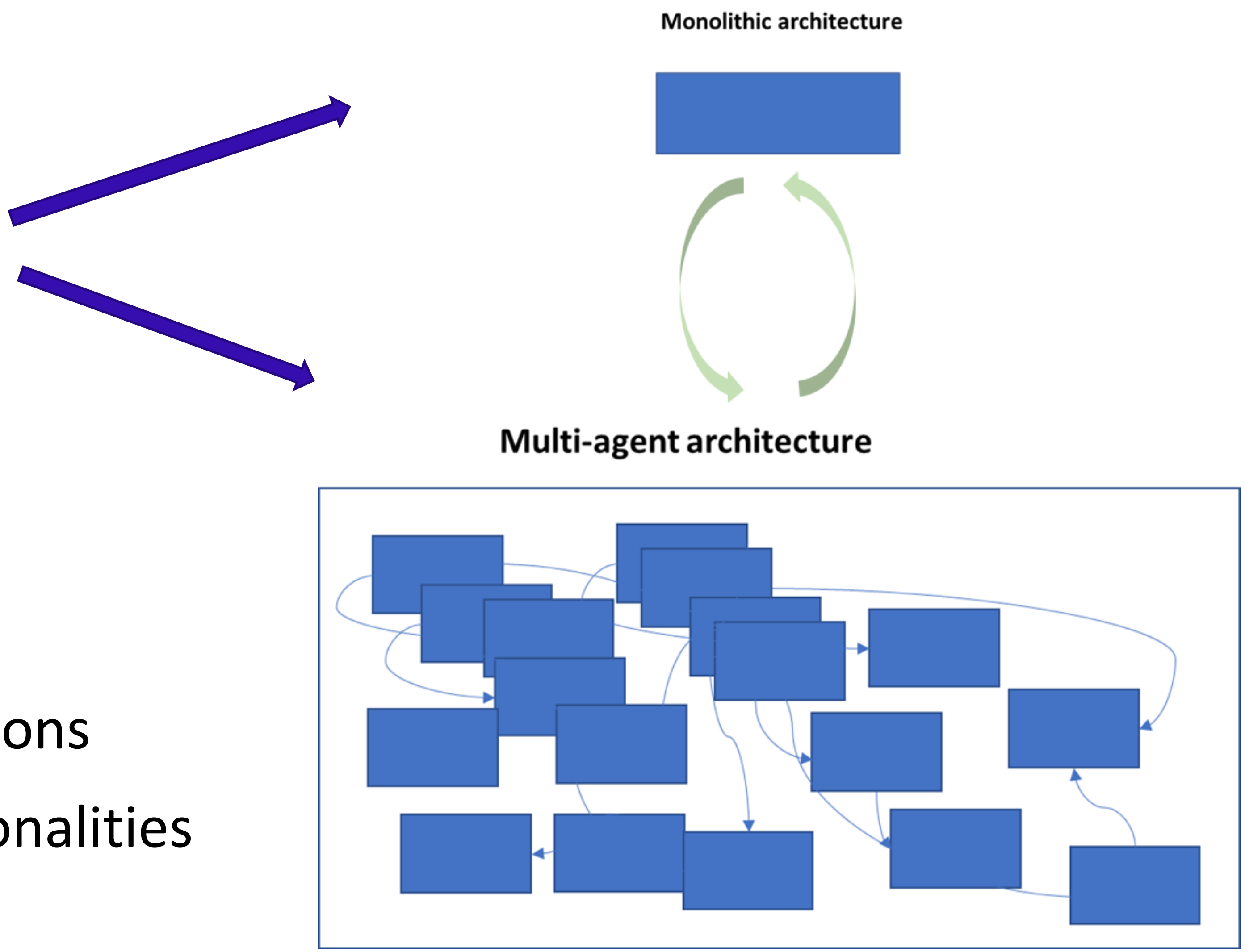
NORCE

- Multiple providers
- Proprietary systems
- Black box behaviour
- Rigid versus adaptive behaviour
- Information on training data
- Online training capabilities
- Information on specific assumptions
- Interoperability: data and functionalities



Lessons learned from implementation of a complex AI system in a high-risk environment

- ✓Multiple providers
- ❖Proprietary systems
- ❖Black box behaviour
- ✓Rigid versus adaptive behaviour
- ✓Information on training data
- ✓Online training capabilities
- ✓Information on specific assumptions
- ✓Interoperability: data and functionalities

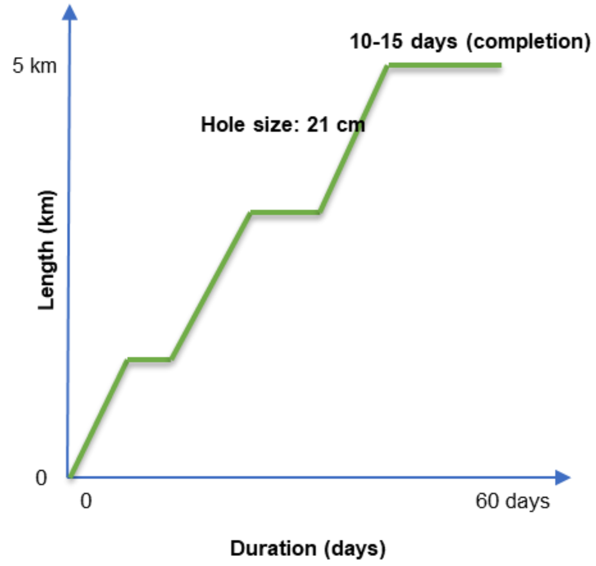


Lessons learned from implementation of a complex AI system in a high-risk environment

NORCE

Drilling operations:

- Oil and gas
- Geothermal
- CO_2 storage
- Hydrogen storage, production, etc.



Main goal of drilling operation is:

- To reach the desired depth as fast as possible.
- To avoid incidents.

Incidents cause delays in the process and can lead to damaging situations (hole collapse, drill-string getting mechanically stuck, etc.)

- Continuous estimation of downhole conditions.
- Proactive actions (weak symptoms)

Lessons learned from implementation of a complex AI system in a high-risk environment

- Context: **dynamic**
- Data: varying **quality and availability**
- Safety of the process: **highly relevant**
- Criticality of the process: **highly critical process**
- Level of human involvement in the decision making: **no human intervention**

Motivation&Objectives:

- ✓ **Increase safety**
- ✓ **Improve performance**
- ✓ **Achieve consistency in the process**

Lessons learned from implementation of a complex AI system in a high-risk environment

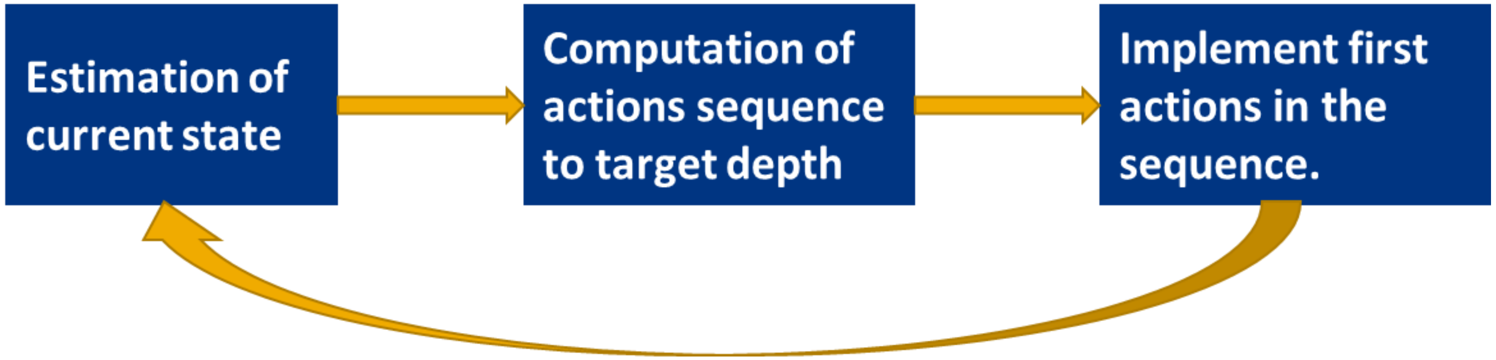
- Aim

To demonstrate autonomous drilling.

- How

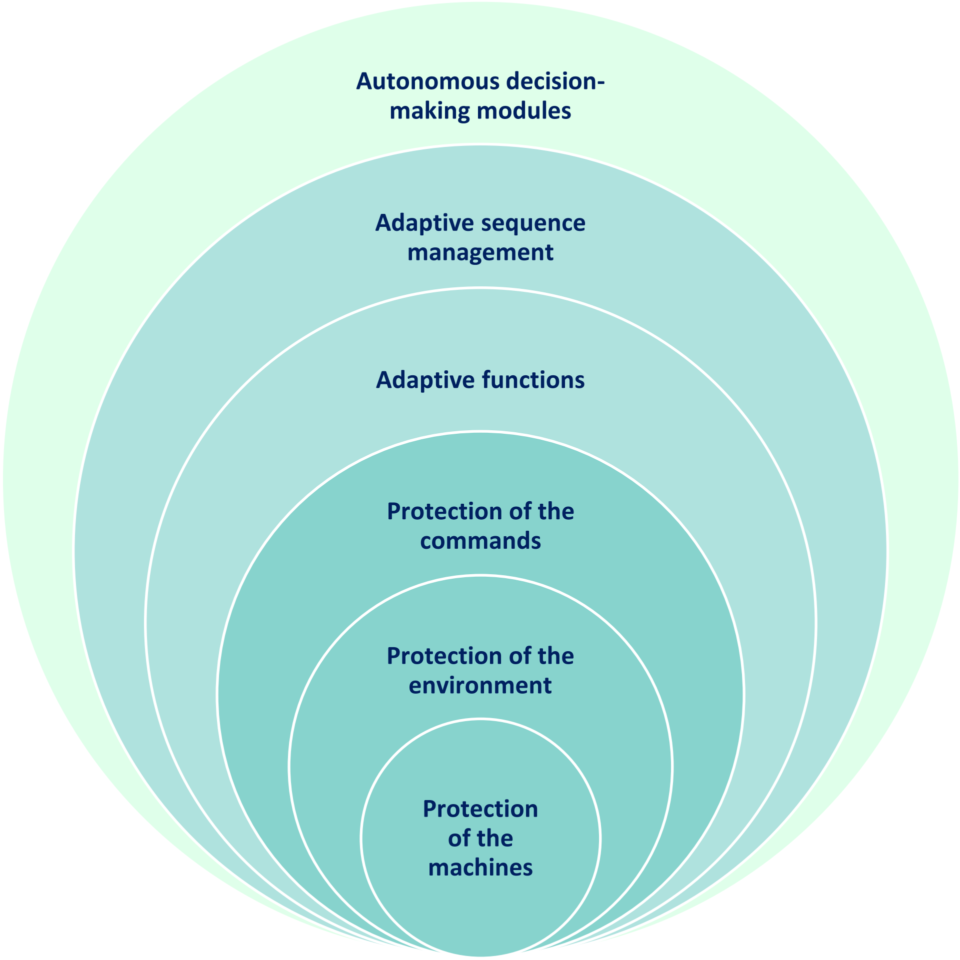
1. **Autonomous decision-making on the fly, handling of unexpected situations.**
2. **Executing these decisions automatically.**

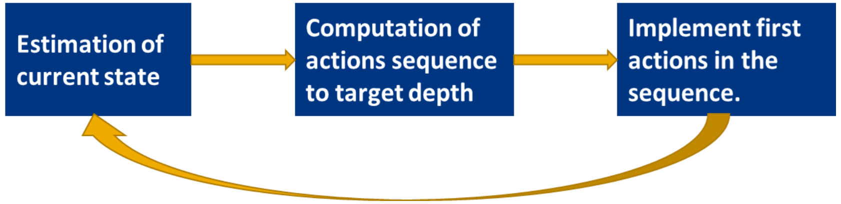
- Strategy



Lessons learned from implementation of a complex AI system in a high-risk environment

NORCE





1. ***On Transitions Functions Model for Decision-Making in Offshore Operations***, R. Mihai, B. Daireaux, A. Ambrus, E. Cayeux and L. Carlsen, NORCE, 2022 IEEE 17th International Conference on Control & Automation (ICCA), 2022, pp. 309-314, doi: 10.1109/ICCA54724.2022.9831911
2. ***Autonomous Decision-Making While Drilling***, E. Cayeux, B. Daireaux, A. Ambrus, R. Mihai, and L. Carlsen, NORCE, Energies, vol. 14, no. 4, p. 969, Feb. 2021, doi: 10.3390/en14040969.
3. ***Demonstration Of Autonomous Drilling On A Full-scale Test Rig***, R. Mihai, E. Cayeux, B. Daireaux, L. Carlsen, A. Ambrus, P. Simensen (NORCE), M. Welmer (SEKAL), M. Jackson (NOV), SPE Annual Technical Conference and Exhibition, Houston, USA, 2022, https://doi.org/10.2118/210229-MS.

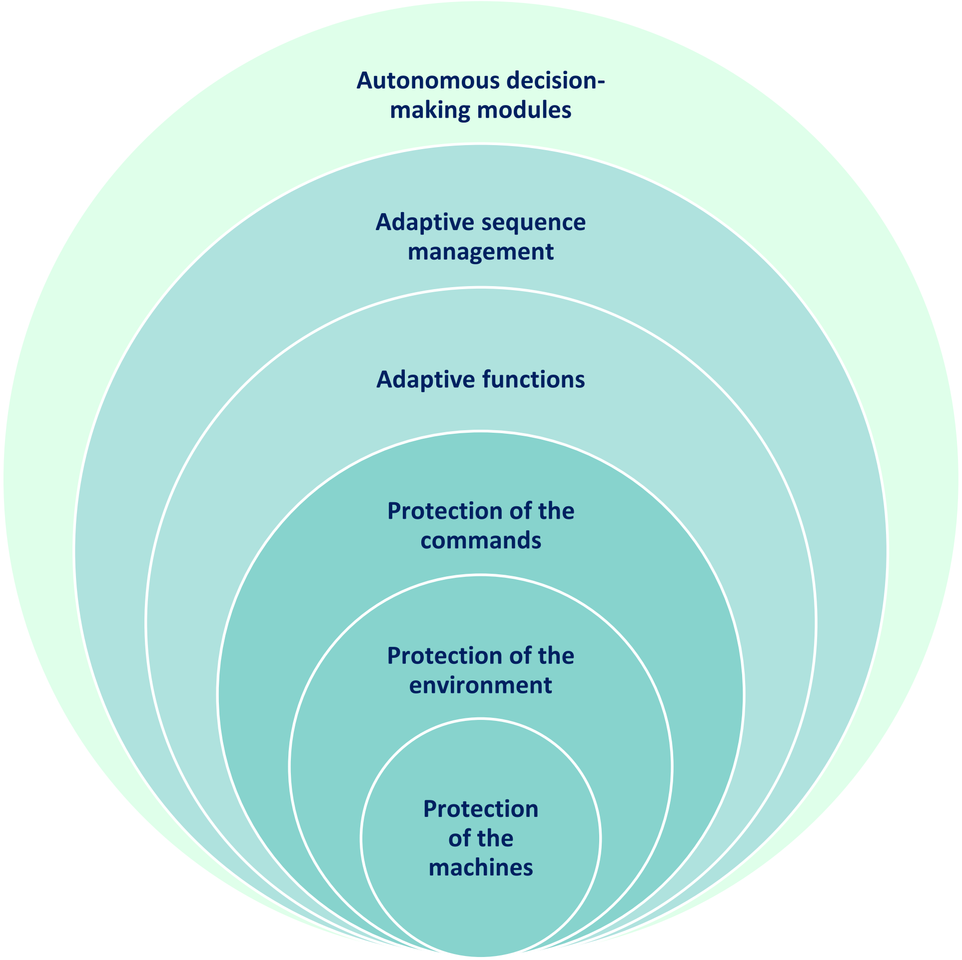
Autonomous drilling: protection features

- **Adaptive functions**

- **Protection functions**:
  - **Safe operating envelopes (SOE)**:
    - Acceptable setpoints limits to respect operational constraints

  - **Fault detection Isolation and Recovery (FDIR)**:
    - Quick events are quite common
    - Fast detection and reaction is often essential to avoid an escalating problem.

- **Safe mode management: transition from autonomous to manual drilling** (to address decreased situational awareness).



Reference: ***A Technical Approach to Safe Mode Management for a Smooth Transition from Automatic to Manual Drilling***. Proc., SPE/IADC International Drilling Conference and Exhibition. E. Cayeux, R. Mihai, L. Carlsen et al. 2021. https://doi.org/10.2118/204114-MS.

Addressing low situational awareness from a technological perspective – autonomous setting

NORCE

- **What**: process state which is stable for a certain amount of time.
  - Give the user sufficient time to get into the current situation
- **When**: transition from autonomous to manual control.
  - Because a recovery procedure has not succeeded.
  - Because of an internal problem which imped continuing in autonomous mode.
- **Why**: safe transition to manual control
  - Because situation awareness of user is reduced while running autonomous processes.
  - It might be a different user than the one who engaged the autonomous mode.
- **How**: Estimate which set of actions that can bring back the process in a stable condition
  - This set of action is context dependent.

Reference: ***A Technical Approach to Safe Mode Management for a Smooth Transition from Automatic to Manual Drilling***. Proc., SPE/IADC International Drilling Conference and Exhibition. E. Cayeux, R. Mihai, L. Carlsen et al. 2021. https://doi.org/10.2118/204114-MS.

Conclusion

NORCE

- Some relevant aspects for verification and validation of complex AI systems were illustrated (also from an implementation of such a system).
- Situational awareness and an example how to cope with reduced situational awareness from a technological perspective.
- Information on specific assumptions and training data relevant to be shared for a smooth interplay in complex AI system.
- Interoperability of data and functionalities is crucial for complex AI systems (multiple providers perspective).

NORCE

Acknowledgement

The authors acknowledge the Research Council of Norway, AkerBP, Equinor, Repsol Norge AS, TotalEnergies, VårEnergi and Petrobras for financing the work through the Demo2000 funding program. The presented work is a research cooperation between Canrig Drilling Technology, Odfjell Drilling, NORCE, NOV and Sekal.

This work is part of the Center for Research-based Innovation DigiWells: Digital Well Center for Value Creation, Competitiveness and Minimum Environmental Footprint (NFR SFI project no. 309589, https://DigiWells.no). The center is a cooperation of NORCE Norwegian Research Centre, the University of Stavanger, the Norwegian University of Science and Technology (NTNU), and the University of Bergen. It is funded by Aker BP, ConocoPhillips, Equinor, TotalEnergies, Vår Energi, Wintershall Dea, and the Research Council of Norway.

sfi Centre for Research-based Innovation

The Research Council of Norway

Thank you!