Management of the monitoring data acquisition uncertainties within deep geological repositories

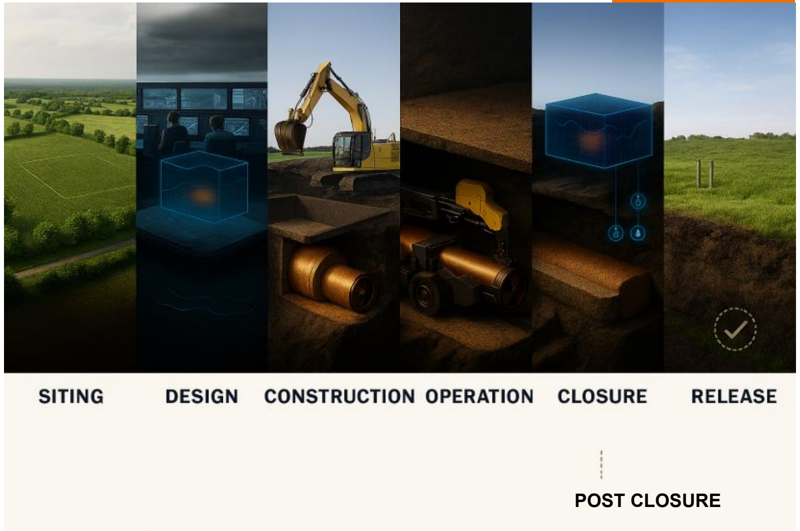
Arto Laikari and Pirjo Hellä
VTT Technical Research Centre of Finland Ltd., Finland

19.9.2025 SafeND 2025 - Workshop W7
Uncertainties in geological 3D-Modells: A Challenge in the nuclear waste repository site selection procedure

Lifecycle of the repository

- DGRs are still in early phase of their lifecycle

- Technologies and modelling tools evolve (fast)
  - 2D maps and diagrams
  - 3D models
  - Digital twins
  - AI and machine learning



- In Finland ONKALO repository early site selection (siting) investigations began in the 1980s.
- Similarly in Sweden surveying work for an URL also started in the 1980s.
- Early site selection also done in other countries.

EURAD MODATS

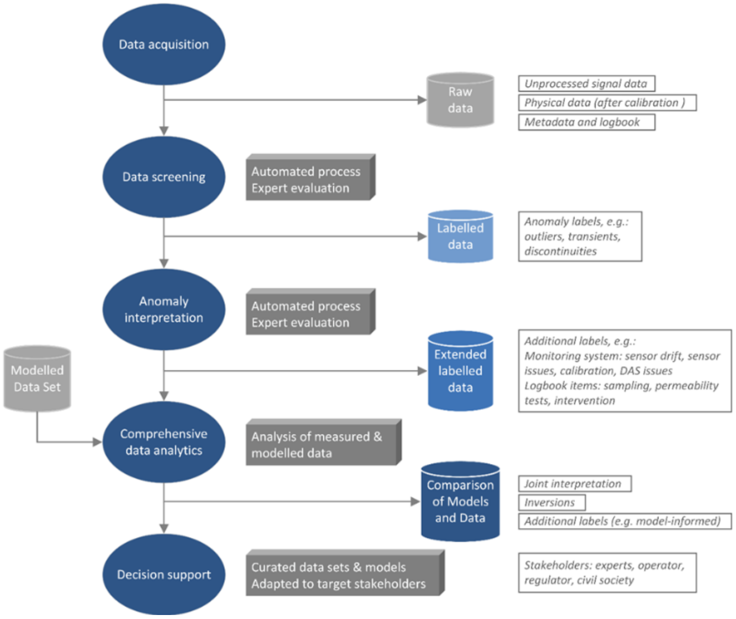
- **MODATS** - MOnitoring equipment and DAta Treatment for Safe repository operation and staged closure (2021 – 2024)
- The objective of MODATS was to evaluate, develop and describe methods and technologies, and to provide the means to measure, treat, analyse and manage data in a consistent manner.

- The work on data management; modelling and visualisation was based on following test cases:
  1. ALC1605 Experiment – Bure URL, France
  2. Full-Scale Emplacement (FE) Experiment – Mont Terri URL, Switzerland
     1. Test case 1 (NAGRA)
     2. Test case 2 (PSI)
     3. Test case 3 (UFZ)
  3. Posiva Plug (POPLU) Experiment – ONKALO, Finland
  4. Prototype Repository II Experiment – Äspö Hard Rock Laboratory, Sweden
  5. PRACLAY Experiment – HADES URL, Belgium

https://www.ejp-eurad.eu/implementation/monitoring-equipment-and-data-treatment-safe-repository-operation-and-staged-closure

EURAD MODATS
Monitoring, 3D modelling and Digital Twins

- Existing long term URL experiment monitoring programmes provided datasets which were used in the iterative development of example tools and data management methods for the safety case to support decision making.
- One part of MODATS focused on the technical development of modelling and visualisation tools, including early digital twin prototypes and data curation methods with machine learning.
- Another part of MODATS focused on synthesis of these efforts, stressing their role in confidence building, decision support, and stakeholder engagement.
- Monitoring data can be used both for follow-up and verifying and improving modelling results.
- Combining monitoring, modelling, digital twins and AI can provide powerful toolset for decision making and stakeholder communication.



MODATS workflow for data handling
From acquisition to decision support.

Monitoring uncertainties

- Sources of Uncertainty:
  - **Sensor reliability**: Issues like calibration drift, fouling, breakage and data transfer errors can compromise data quality.
  - **Environmental variability**: Geological heterogeneity and long-term changes in repository conditions introduce uncertainty in interpreting measurements.
  - **Data gaps**: Long monitoring timescales increase the risk of incomplete datasets due to equipment failure or obsolescence.
  - **Context loss**: When event logs are incomplete or missing.
  - **Evolving data management concepts**: Old concepts are missing many modern options or metadata definitions.
- Mitigation Approaches:
  - **Quality Assurance (QA/QC)**: Standardized procedures for calibration, redundancy in sensor networks, and robust data validation.
  - **Data curation**: Detecting and correcting anomalies before integration into models.
  - **Uncertainty quantification**: Statistical methods and sensitivity analyses applied to monitoring datasets to assess confidence levels.

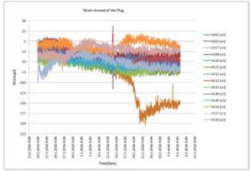
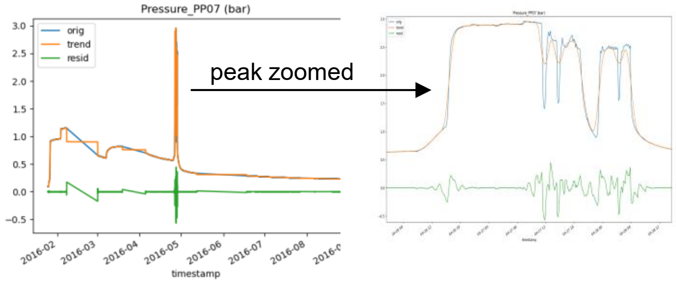
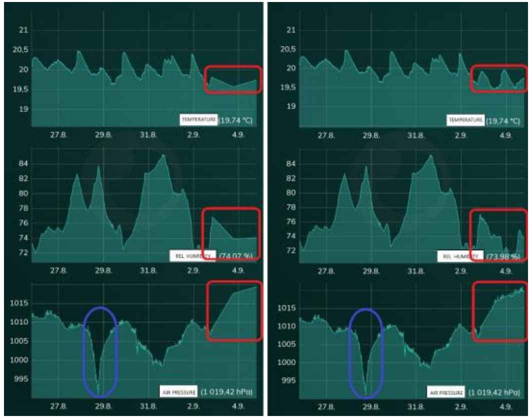


Figure 7.13. Strain Gauges in around the plug.

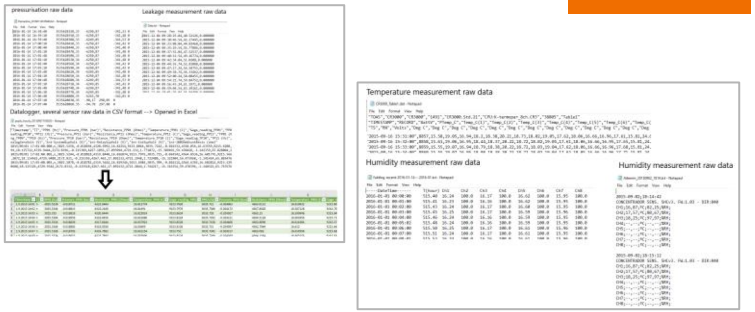


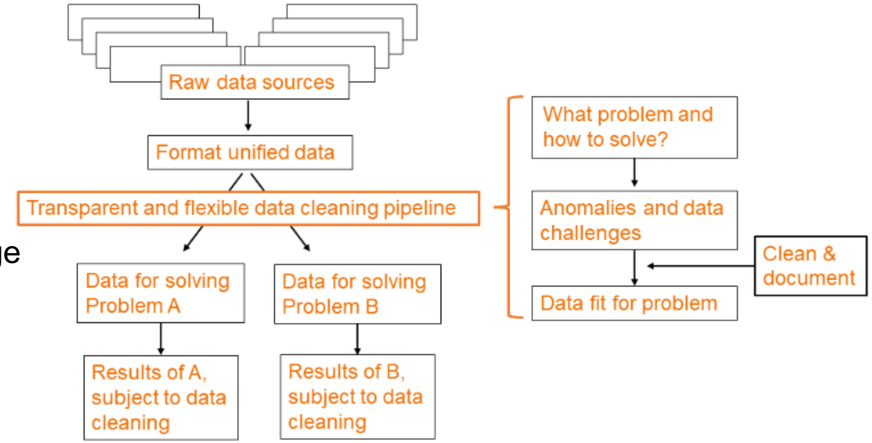


missing data interpretation – how to do?

Data management & Data curation

- Data is acquired in various formats and many sources.
- Data management need to be planned. Many aspects to be taken into account. E.g.
  - Ontology and metadata for sensors, time perspectives, location information, events and curation processes.
  - How to collect, store and secure data.
  - How to document everything.
  - Data formats and their evolution.
  - Version control.
  - Data validation and curation & Data analytic methods.
- Data needs to be processed differently for different usage scenarios!
  - Keep raw data as much as possible.
- Automated processes to be developed.
  - Use of machine learning and AI tools.
- Curated data can be used with models, AI and decision-making support.



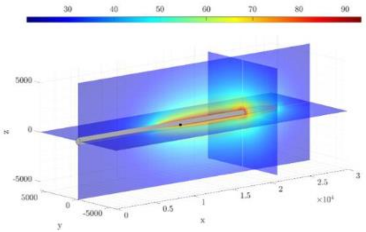


Simplified framework of using sensor data

EURAD MODATS

3D Models

- Role in Monitoring:
  - 3D models are used to simulate repository evolution (thermal, hydraulic, mechanical, and chemical processes) and compare predictions with monitoring data. They help e.g. to visualize sensor placement and optimize monitoring strategies.
- Integration with Data:
  - Models are updated iteratively as new monitoring data becomes available, improving predictive capability.
  - Models are used for scenario testing to evaluate the impact of uncertainties on repository performance.

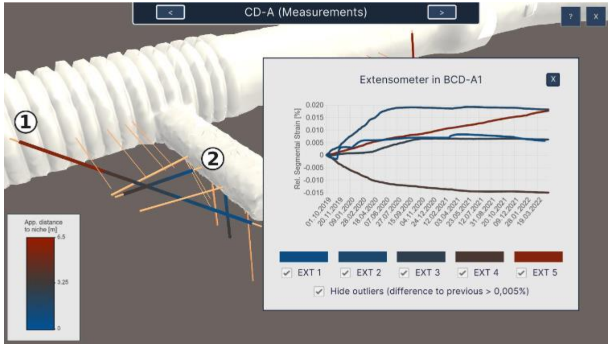


Digital twins

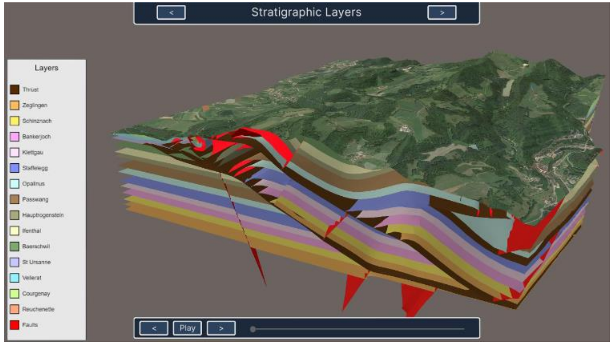
- Concept:
  - A digital twin is a dynamic, virtual representation of the repository system that integrates real-time monitoring data with physical-based models.
- Applications in MODATS:
  - Prototyping: Partners developed preliminary digital twin frameworks using datasets from underground research laboratories (URLs).
  - Data-driven updates: Digital twins continuously assimilate monitoring data, enabling predictive analytics and early anomaly detection.
  - Decision support: Facilitates scenario analysis for operational decisions.

Several model types available for site selection

1. Geological and Geomechanical Models
   - Purpose: Understand the structural integrity and stability of the host rock and surrounding formations.
2. Hydrogeological and Transport Models
   - Purpose: Evaluate groundwater flow and radionuclide migration pathways.
3. Performance Assessment (PA) Models
   - Purpose: Quantify long-term safety and compliance with regulatory criteria.
4. Thermo-Hydro-Mechanical-Chemical (THMC) Coupled Models
   - Purpose: Capture interactions between heat from waste, groundwater movement, mechanical stress, and chemical reactions.
5. Safety Assessment and Scenario Models
   - Purpose: Support the safety case by analyzing potential future evolutions and disruptive events.
6. Integrated Digital Models
   - Purpose: Combine geological, hydrogeological, and engineering data into a unified platform.



Screenshot from the interactive visualisation system for Mont Terri URL



Screenshot from the interactive visualisation system for Mont Terri URL

TERMINOLOGY – MODELS vs. DIGITAL TWINS
WHAT IS A DIGITAL TWIN?

- Link to the presentation *"Advancing Digital Twin Technology for Optimized Nuclear Waste Management: Bridging Disciplines and Enhancing Safety"* in the SafeND 2025 session T5g.
  - Presentation based on EURAD-2 DITOCO2030 work package work.
- Comparison of the functionalities of models vs. digital twins.
- For this domain, the Digital twins are tools.

1. For what is the digital twin used for?
2. What size the digital twin is?
3. For whom is the digital twin made for?
4. Who is using the digital twin?
5. Who is making the digital twin?

Feature	BIM (Building Information Model)	Digital Twin
Purpose	Design, construction planning, and documentation	Real-time monitoring, simulation, and predictive analytics
Data Source	Static design and construction data	Live operational data + historical data
Update Frequency	Updated during design and construction phases	Continuously updated throughout the asset lifecycle
Connectivity	Typically standalone or linked to project databases	Connected to IoT sensors, SCADA, and other real-time systems
Lifecycle Coverage	Primarily design and build phases	Full lifecycle: design → operation → maintenance → decommissioning
Simulation Capability	Limited to design scenarios	Advanced simulations (what-if, predictive maintenance, etc.)
Decision Support	Supports design and construction decisions	Supports operational, maintenance, and strategic decisions

Conclusions

- Reliable data treatment and analysis require extensive additional information beyond the acquired sensor values from one or a few sensors.
  - Datasets from multiple sensors need to be treated and analysed together, incorporating cross-checking and adding information about known events.
- Applying domain knowledge of the entire environment and the used sensors and data acquisition setup is crucial to ensure correct interpretations.
- Over a long timeframe, the amount of data accumulates, and data management concepts evolve, making it possible to revisit the old datasets and find new interdependencies between unforeseen phenomena.
- Use of Artificial Intelligence (AI) is rapidly increasing and bringing new tools to manage acquired data and build and improve models.
  - AI requires extensive data cleaning to avoid outcomes reflecting sensor problems. Each AI result is subject to data pre-treatment.
- Defining Digital Twins (DT) involves determining its scope, purpose, and whether it is used for "as is representation" or predicting future behaviour of an entity, as well as identifying the target group of users.
  - There is no one-size-fits-all solution, but there need to be definitions for specific purpose DTs.

VTT

beyond
the obvious

Thank you!
Danke schön!

Arto.Laikari@vtt.fi

vttresearch.com