Tradeoff Between Data Security and Resilience in AI Pipelines

By: Evan Stosic

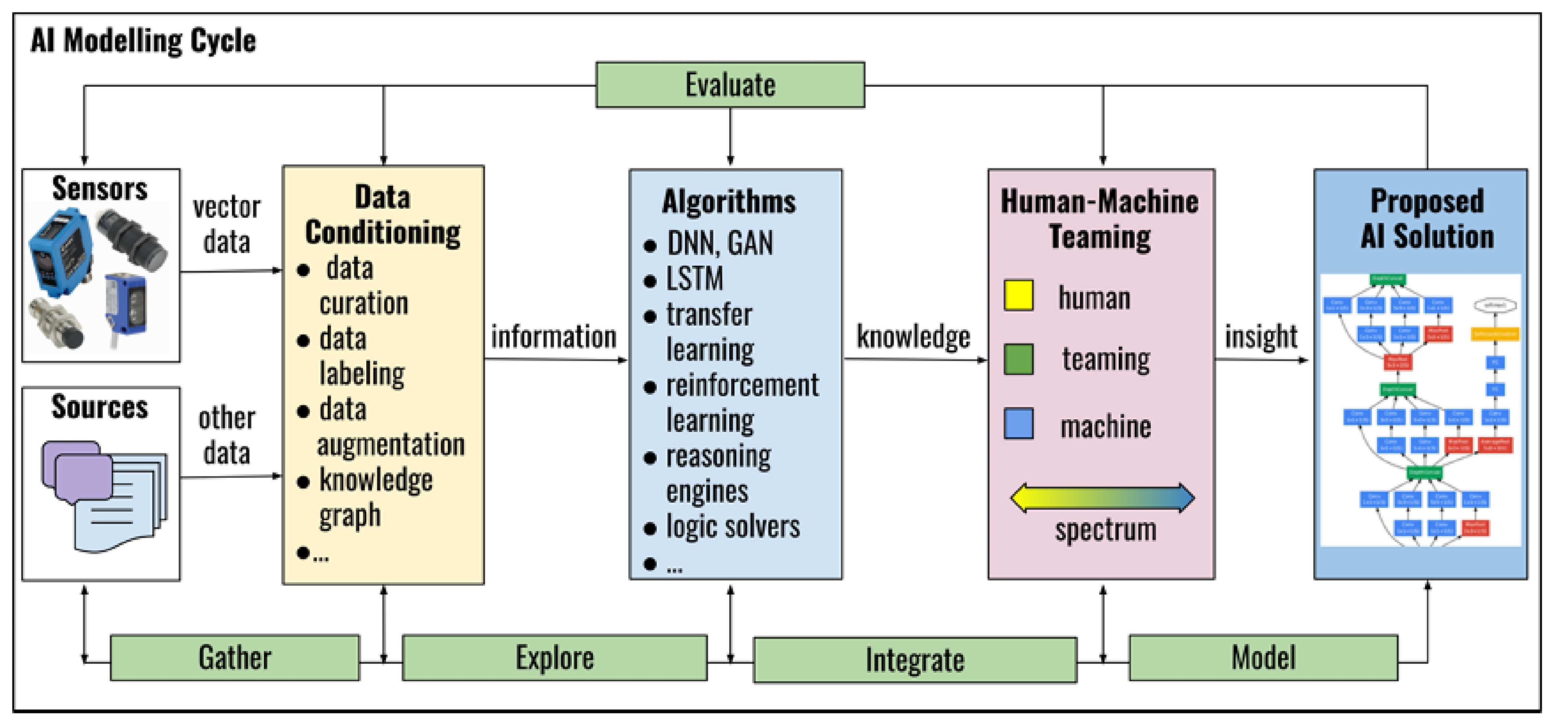
Overview

- What I Keep Finding in Research
- Possible Research Gaps

What I Keep Finding in Research

- When I try to find papers discussing the tradeoff between data security and resilience in AI pipelines, I keep running into papers that center around the following:
  - Resilience in AI Pipelines
  - Data Pre-Processing Methods
  - Tradeoffs for Continuous Integration, not necessarily AI

What I Keep Finding in Research



AI Modelling Cycle

Evaluate

Sensors

vector data

Sources

other data

Data Conditioning

- data curation
- data labeling
- data augmentation
- knowledge graph
- …

information

Algorithms

- DNN, GAN
- LSTM
- transfer learning
- reinforcement learning
- reasoning engines
- logic solvers
- …

knowledge

Human-Machine Teaming

🟨 human

🟩 teaming

🟦 machine

spectrum

insight

Proposed AI Solution

Gather

Explore

Integrate

Model

Source: [1]

Data Pre-Processing Methods

- I’ve only found papers discussing how to mitigate bad data propagation in an AI pipeline
  - Better data pre-processing
  - Data quality checks along a pipeline
  - Train an AI model with some resiliency to bad data
- When these methods are mentioned, trade-off with resilience is not considered
  - Ex: Data quality checks: If a check continually fails in a pipeline, how does this affect the resilience of the pipeline?
- It seems in the literature that data assurance is prioritized without measuring effects on AI pipeline resilience

Possible Research Gaps

- What is the effect of different data pre-processing methods and data-quality checks on the resilience of an AI pipeline?
  - Maybe a comprehensive view?
  - Worst vs. average vs. best case scenario
- When physical parts break down and either stop sending data or send bad data, how does this affect a pipeline’s resilience broadly?

References

- [1]: https://www.mdpi.com/2504-4990/3/1/4
- [2]: https://www.mdpi.com/1999-4893/16/3/165
- [3]: https://www.mdpi.com/2076-3417/13/12/7082
- [4]: https://mir.cs.illinois.edu/marinov/publications/HiltonETAL17TradeOffsInCI.pdf