3

Machine Learning in Digital Twins for Threat Estimation and Detection

<u>Supervisors:</u>

Fabien DAGNAT <fabien.dagnat@imt-atlantique.fr>
Marc-Oliver PAHL <marc-oliver.pahl@imt-atlantique.fr
>
Fehmi JAAFAR <fjaafar@uqac.ca>

PhD Student – Hugo BOURREAU
hugo.bourreau@cyberCNI.fr

security research

speaker series

PhD schools

MOOCs

CYBER CNI.FR

Presentation plan

- Who Am I?
- Let’s breakdown Digital Twin
- Digital Twin literature
- Emerging Approaches: From Data-Driven to Model-Aware Digital Twins
  - Reference system modelization
  - Double pipeline ML
- Future Work
- Challenges
- Summary

Who Am I? — PhD Student, Digital Twins & Cybersecurity

🎓 *PhD Student – IMT Atlantique & UQAC (Cotutelle France–Canada)*

🧩 *Member of the Industrial Chair CyberCNI*

📅 *2nd year PhD – 2024 → 2028*

🧠 *Research Focus:* Digital Twins, Machine Learning, Cyber-Physical Systems Security

My research focuses on leveraging Digital Twins to model system behavior and estimate ongoing or future cyber attacks.

Who Am I?

Goals of my thesis

Goal 1 Modelization of the reference system within the DT.

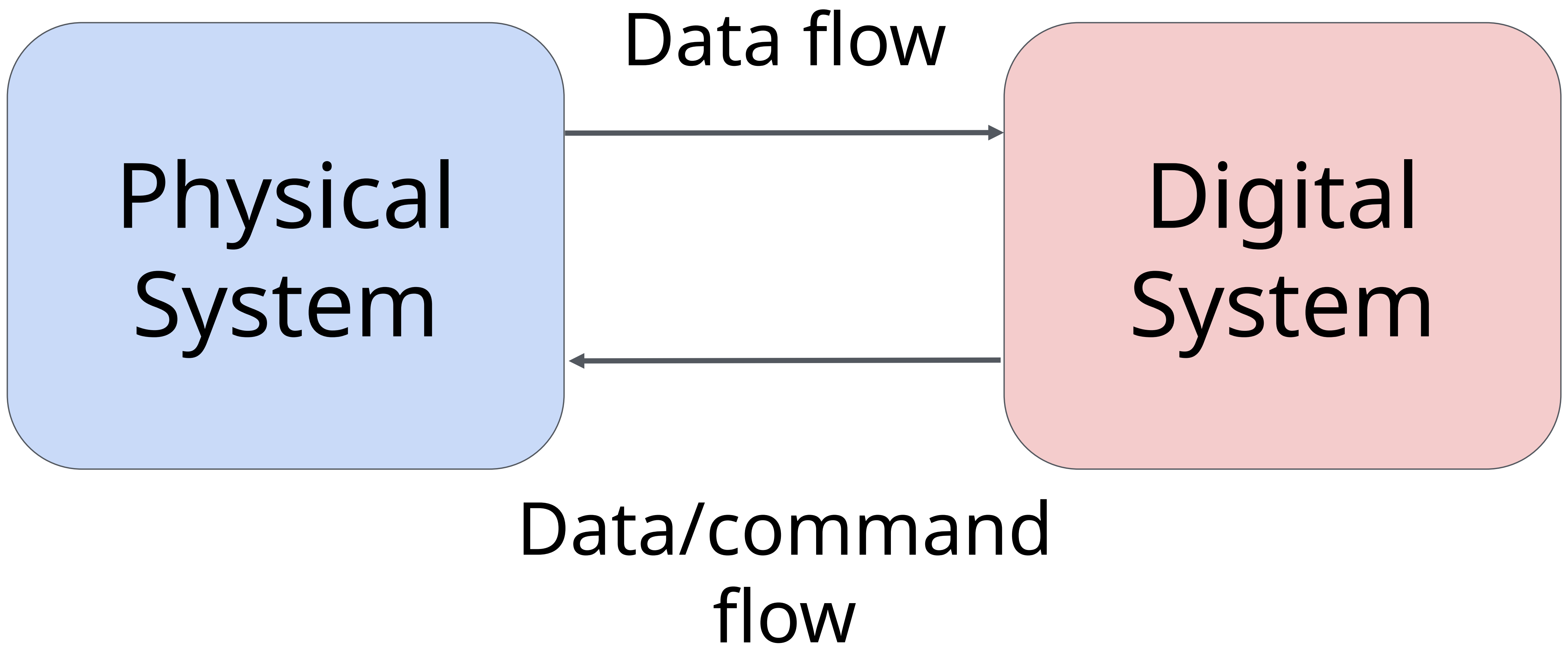
Goal 2 Using DTs for simulating side-channel measurements for better attack detection.

Goal 3 Using DTs for more comprehensive risk analysis.

Let’s break down Digital Twin

CYBER CNI.FR

Let’s break down Digital Twin



<u>Fig 1:</u> Digital Twin core principle

Let’s break down Digital Twin

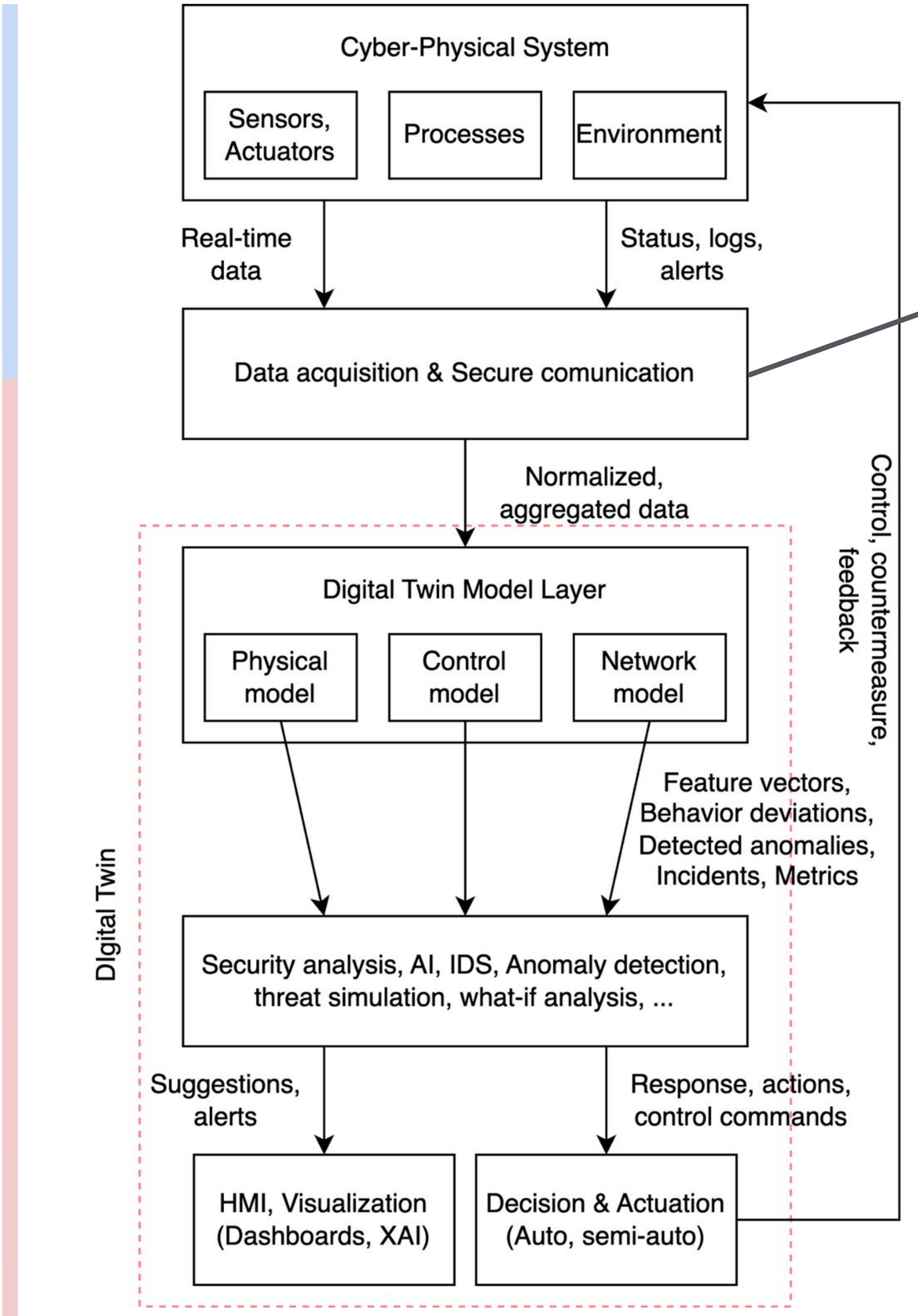


Fig 2: Digital Twin architecture

1.Gather reference system data

Aggregate and normalize

Storage and integrity

Let’s break down Digital Twin

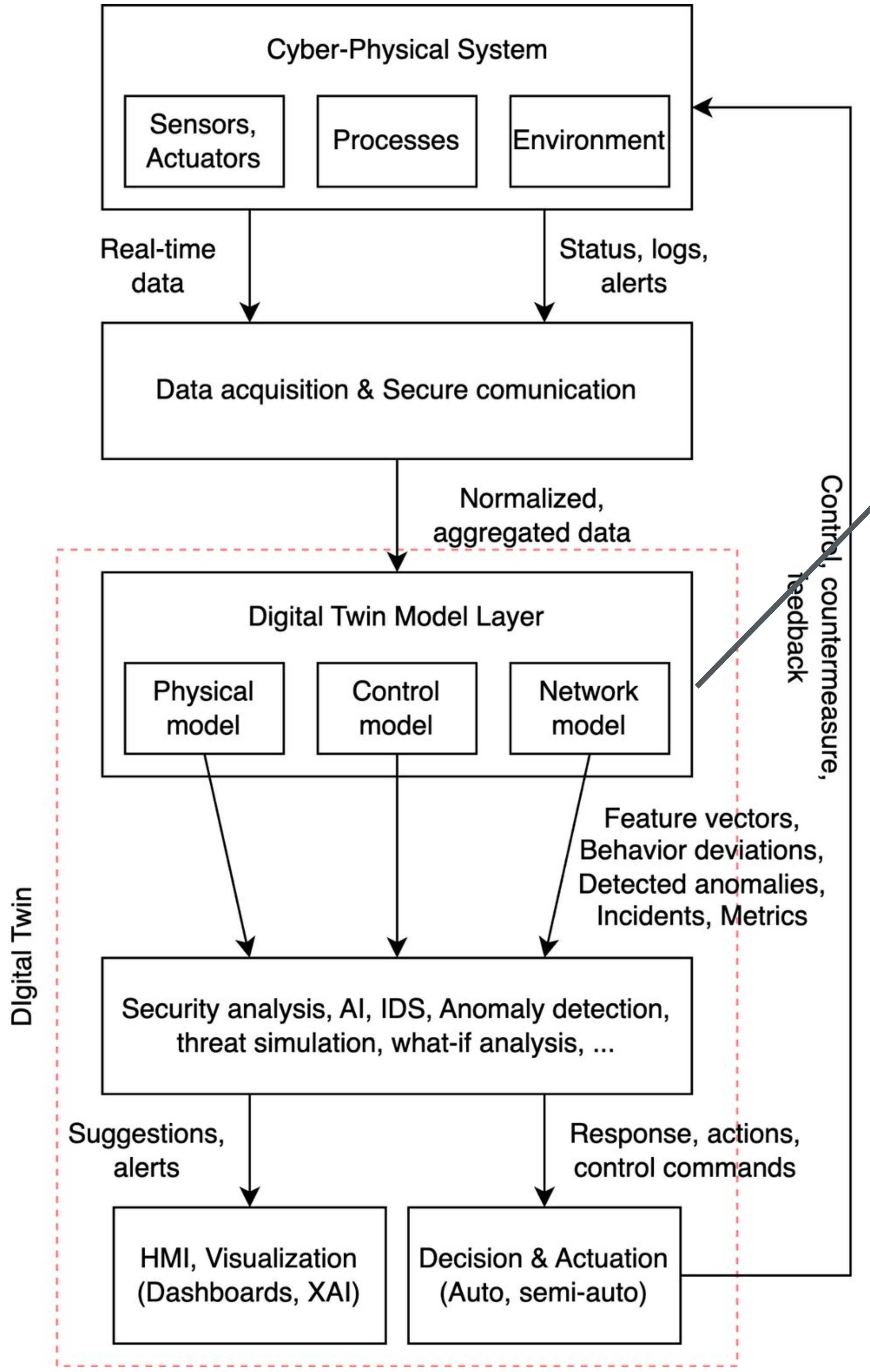


Fig 2: Digital Twin architecture

2.Physical system modelization

Represent system’s state

Make simulations

Let’s break down Digital Twin

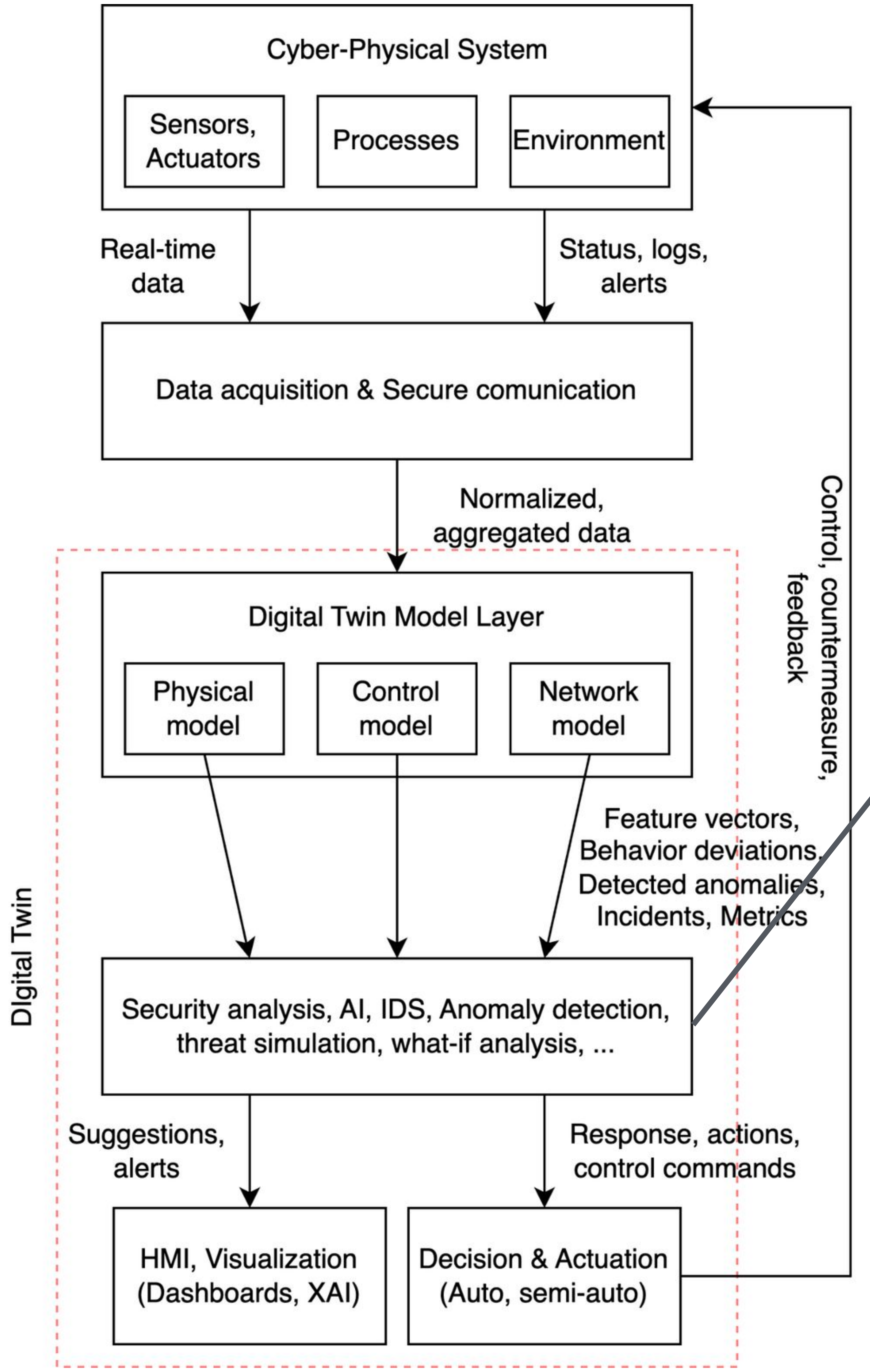


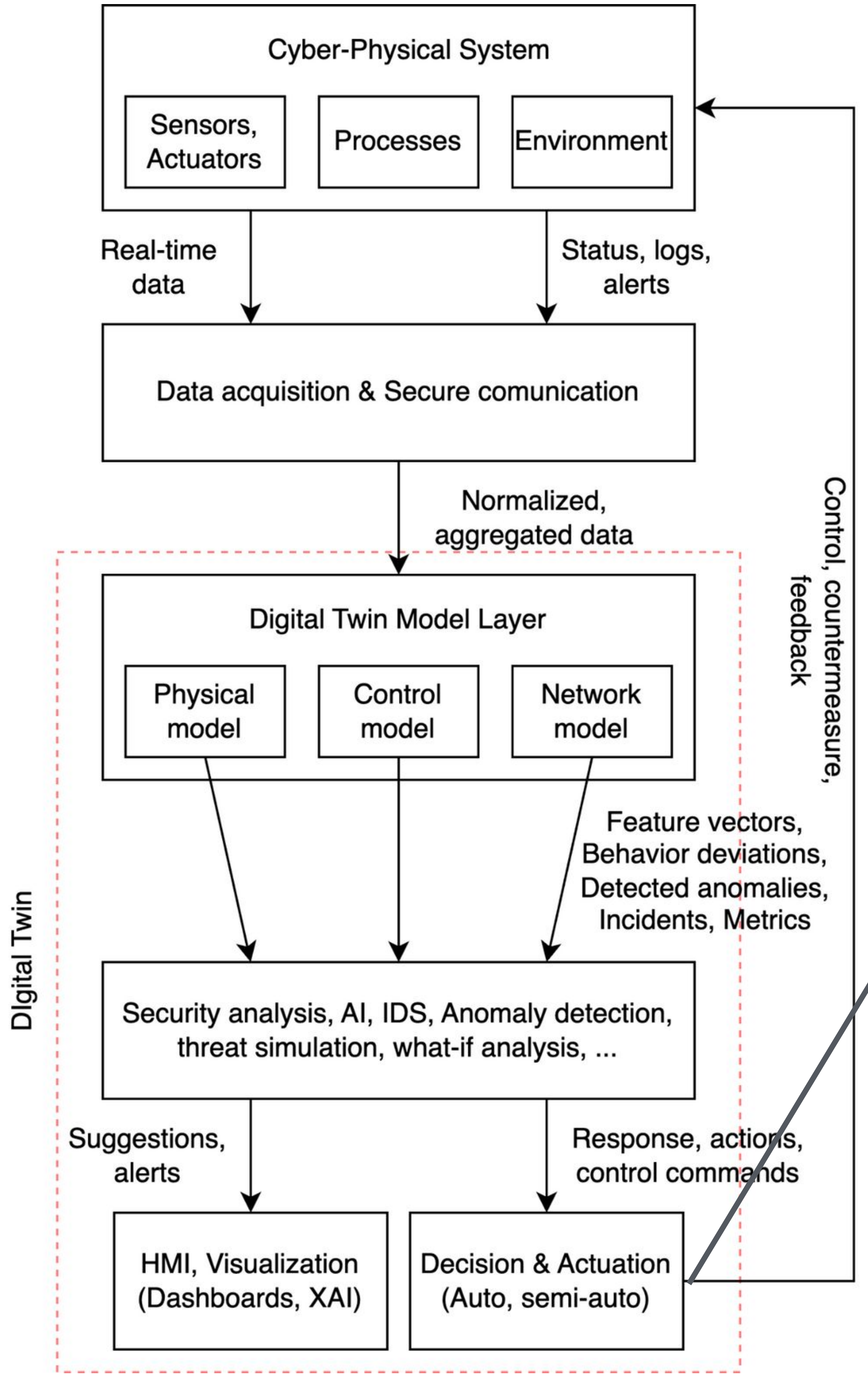
Fig 2: Digital Twin architecture

3.Data and metric analysis

Get alerts or recommandations

Responses or actions

Let’s break down Digital Twin



<u>Fig 2:</u> Digital Twin architecture

4.Visualization & Decision

Present the results to dashboards

Send command control to the system

Let’s break down Digital Twin

Digital Twin Consortium Periodic Table

1.

2. & 3.

4.

DS.AR AI Model Repository	**DS.AG** Data Aggregation	**IR.AS** API Srvices				**UX.GM** Gamification	**UX.DB** Dashboards
DS.SR Simulation Model Repository	**DS.AS** Asynchronous Integration	**IR.CL** Collaboration Platform Integration	**IC.CS** Composition	**IC.SM** Simulation	**IC.RP** Reporting	**UX.3R** 3D Rendering	**UX.XR** Extended Reality (XR)
DS.SA Data Storage and Archive Services	**DS.RT** Real-time Processing	**IR.DT** Digital Twin Integration	**IC.DL** Distributed Ledger and Smart Contracts	**IC.FL** Federated Learning	**IC.AL** Alerts and Notifications	**UX.GE** Gaming Engine Visualization	**UX.ER** Entity Relationship Visualization
DS.DS Domain Specific Data Management	**DS.BP** Batch Processing	**IR.IO** OT/IoT System Integration	**IC.BR** Business Rules	**IC.AI** Artificial Intelligence	**IC.OS** Orchestration	**UX.BP** Business Process Mgmt & Workflow	**UX.RM** Real-time Monitoring
DS.IR Digital Twin Instance Repository	**DS.CX** Data Contextualization	**IR.EG** Engineering Systems Integration	**IC.PS** Prescriptive Recommendations	**IC.PR** Prediction	**IC.IC** Command and Control	**UX.BI** Business Intelligence	**UX.AV** Advanced Visualization
DS.RP Digital Twin Model Repository	**DS.TR** Data Transformation and Wrangling	**IR.ET** Enterprise System Integration	**IC.MA** Mathematical Analytics	**IC.AA** Data Analysis and Analytics	**IC.SR** Search	**UX.CI** Continuous Intelligence	**UX.BV** Basic Visualization
DS.ON Ontology Management	**DS.ST** Data Streaming	**MG.DG** Data Governance	**MG.SM** System Monitoring	**TW.RP** Responsibility	**TW.RL** Reliability	**TW.PR** Privacy	**TW.DS** Device Security
DS.SG Synthetic Data Generation	**DS.AI** Data Acquisition and Ingestion	**MG.EL** Event Logging	**MG.DM** Device Management	**TW.RS** Resilience	**TW.SF** Safety	**TW.SC** Security	**TW.EX** Data Encryption

Data Services · Integration · Intelligence · UX · Management · Trustworthiness

Fig 3: Digital Twin consortium periodic table

Let’s break down Digital Twin

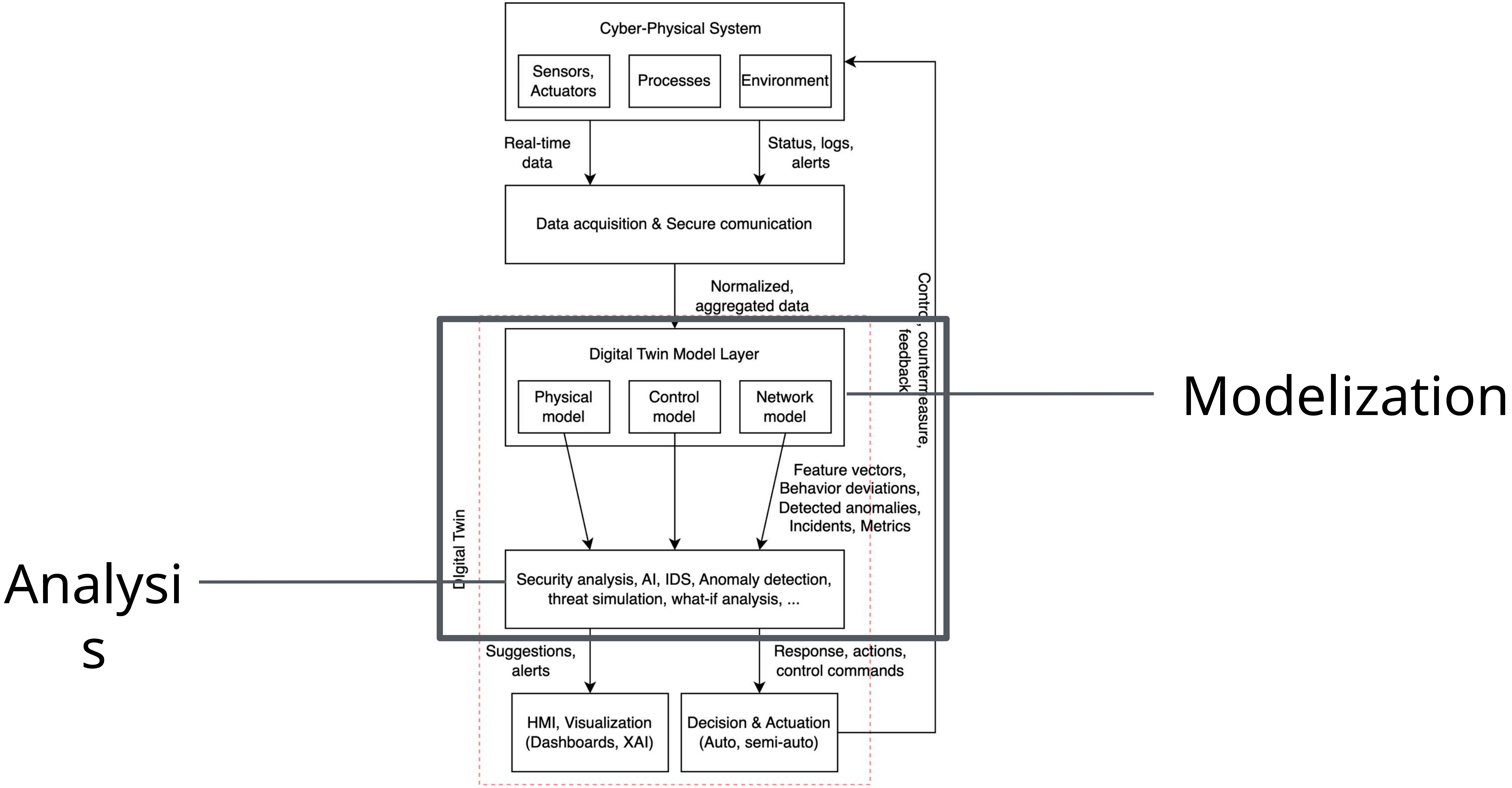


Fig 2: Digital Twin Architecture

3

Digital Twin literature

CHAIRE CYBER CNI

Sécurité des infrastructures critiques

CYBER CNI.FR

Digital Twin Literature

Taxonomy application

Ref	Deployment	Synchronization	Main security role	Evolution strategy	Learning method	AI/ML Horizon
[1]	On-site	Continuous	Intrusion detection	Static	Supervised	Reactive
[2]	Edge	Continuous	Proactive detection	Static	Supervised	Short-term estimation
[3]	On-site	Continuous	Intrusion detection	Static	Unsupervised	Reactive
[4]	Hybrid	Periodic	Decision support	Static	Supervised	Short-term estimation
[5]	Cloud	Continuous	Intrusion detection	Static	Semi-supervised	Reactive
[6]	Cloud	Continuous	Proactive detection	Static	Supervised	Short-term estimation
[7]	Edge / Hybrid	Continuous	Proactive detection	Static	Unsupervised	Short-term estimation
[8]	Cloud	Periodic	Intrusion detection	Static	Supervised	Reactive

Fig 4.1: Positioning of paper using ML with DT within the taxonomy

Digital Twin Literature

Taxonomy application

Deployment and location	Synchronization	Main security role	Evolution strategy	Learning method	ML Horizon

(R.1) Passive Monitoring : Observation and state collection.
(R.2) Intrusion Detection (IDS) : Identification of known/unknown attacks.
(R.3) Proactive Detection (Prediction) : Forecasting attacks before exploitation.
(R.4) Decision Support : Assisting operators in response actions.
(R.5) Post-Incident Analysis (Forensics) : Reconstruction after an attack.

(E.1) Static (Fixed Model) : Trained once without updates.
(E.2) Incremental : Continuously updated as new data arrives.
(E.3) Federated : Distributed training across multiple DTs or entities.

Digital Twin Literature

Taxonomy application

Ref	Deployment	Synchronization	Main security role	Evolution strategy	Learning method	AI/ML Horizon
[1]	On-site	Continuous	Intrusion detection	Static	Supervised	Reactive
[2]	Edge	Continuous	Proactive detection	Static	Supervised	Short-term estimation
[3]	On-site	Continuous	Intrusion detection	Static	Unsupervised	Reactive
[4]	Hybrid	Periodic	Decision support	Static	Supervised	Short-term estimation
[5]	Cloud	Continuous	Intrusion detection	Static	Semi-supervised	Reactive
[6]	Cloud	Continuous	Proactive detection	Static	Supervised	Short-term estimation
[7]	Edge / Hybrid	Continuous	Proactive detection	Static	Unsupervised	Short-term estimation
[8]	Cloud	Periodic	Intrusion detection	Static	Supervised	Reactive

Fig 4.2: Positioning of paper using ML with DT within the taxonomy

Digital Twin Literature

Taxonomy application

Deployment and location	Synchronizatio n	Main security role	Evolution strategy	Learning method	ML Horizon

(H.1) Reactive : Immediate attack detection based on reference system flows.
(H.2) Short-term Estimation : Predicting imminent events such as drifts or anomalies.
(H.3) Long-term Evaluation : Risk analysis and resilience planning.

Digital Twin Literature

Taxonomy application

Ref	Deployment	Synchronization	Main security role	Evolution strategy	Learning method	AI/ML Horizon
[1]	On-site	Continuous	Intrusion detection	Static	Supervised	Reactive
[2]	Edge	Continuous	Proactive detection	Static	Supervised	Short-term estimation
[3]	On-site	Continuous	Intrusion detection	Static	Unsupervised	Reactive
[4]	Hybrid	Periodic	Decision support	Static	Supervised	Short-term estimation
[5]	Cloud	Continuous	Intrusion detection	Static	Semi-supervised	Reactive
[6]	Cloud	Continuous	Proactive detection	Static	Supervised	Short-term estimation
[7]	Edge / Hybrid	Continuous	Proactive detection	Static	Unsupervised	Short-term estimation
[8]	Cloud	Periodic	Intrusion detection	Static	Supervised	Reactive

Fig 4.3: Positioning of paper using ML with DT within the taxonomy

Digital Twin Literature

Reference System Models for Cybersecurity

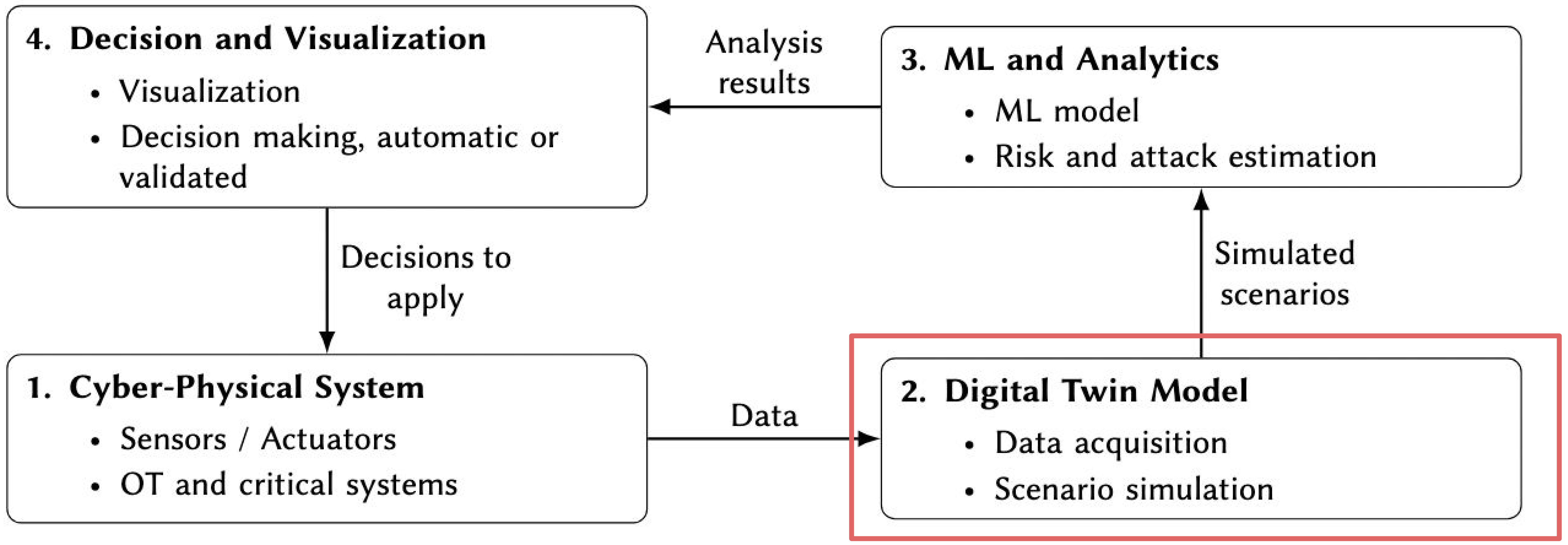


Fig 5: Layers of DT architecture applied with ML

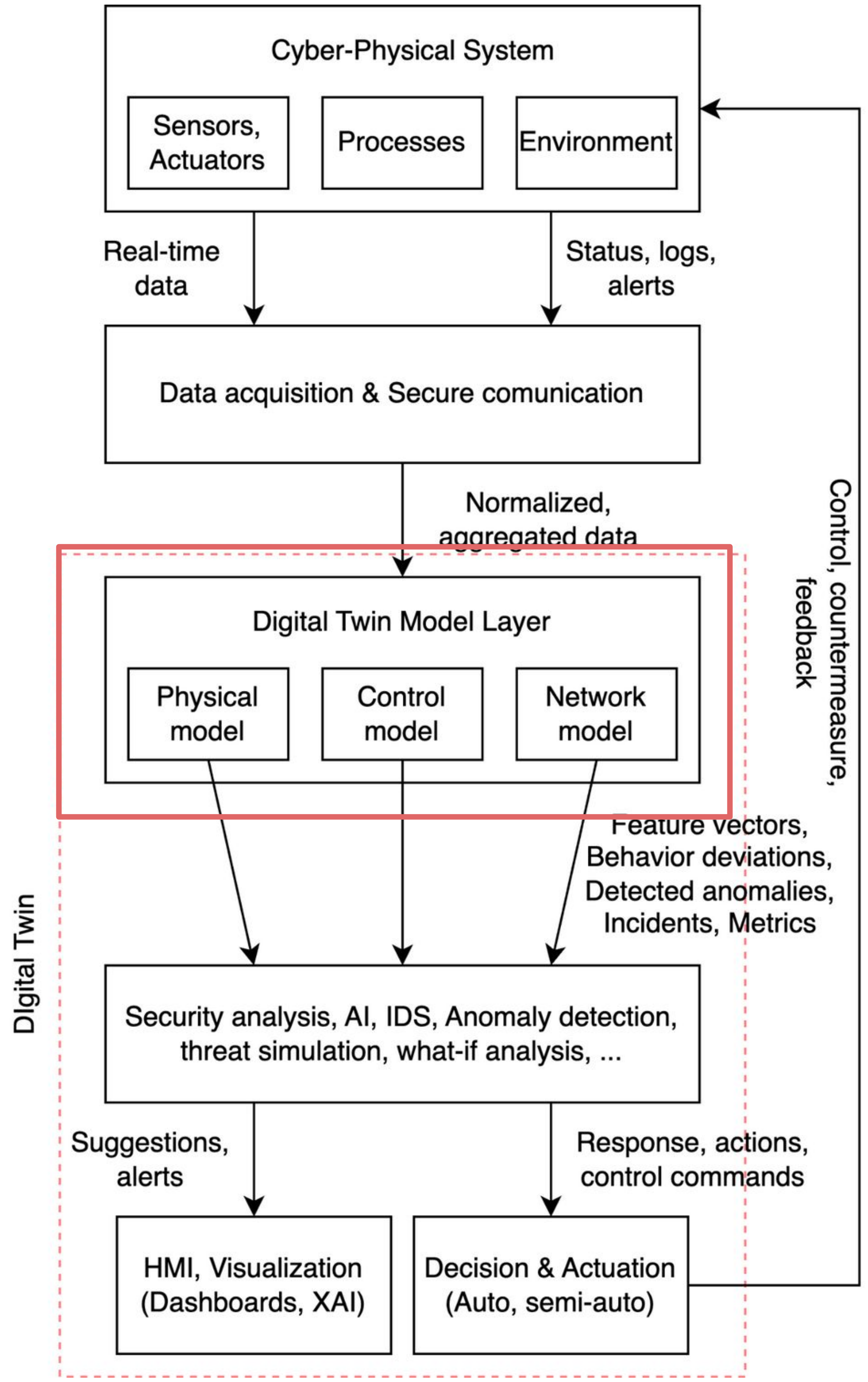


Fig 2: Layers of DT architecture

3

Emerging Approaches: From Data-Driven to Model-Aware Digital Twins

CYBER CNI.FR

Most Digital Twins Only Mirror Data — Not System Behavior

Without an explicit model of the reference system, detection stays reactive.

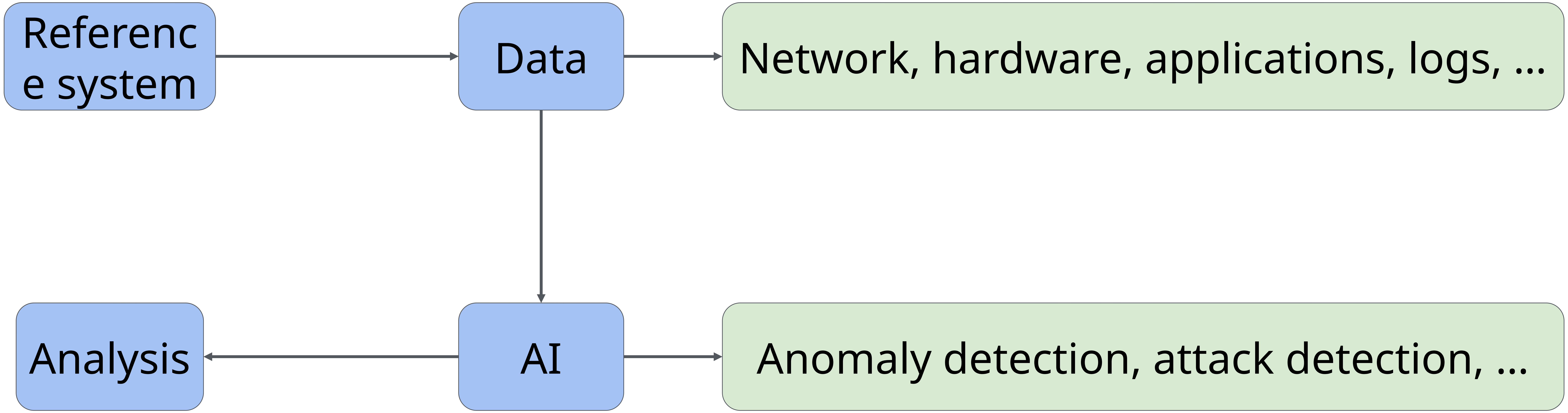
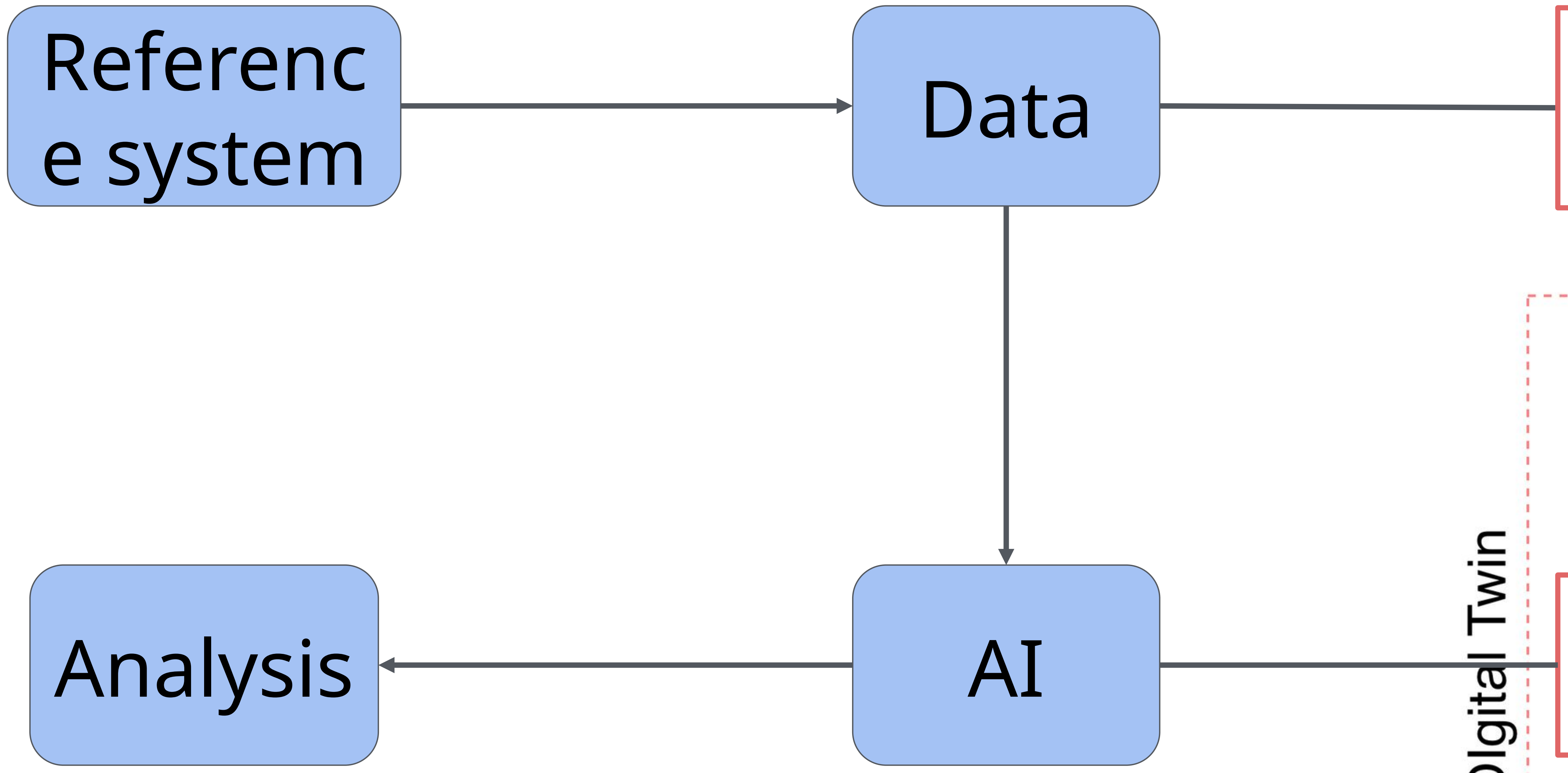


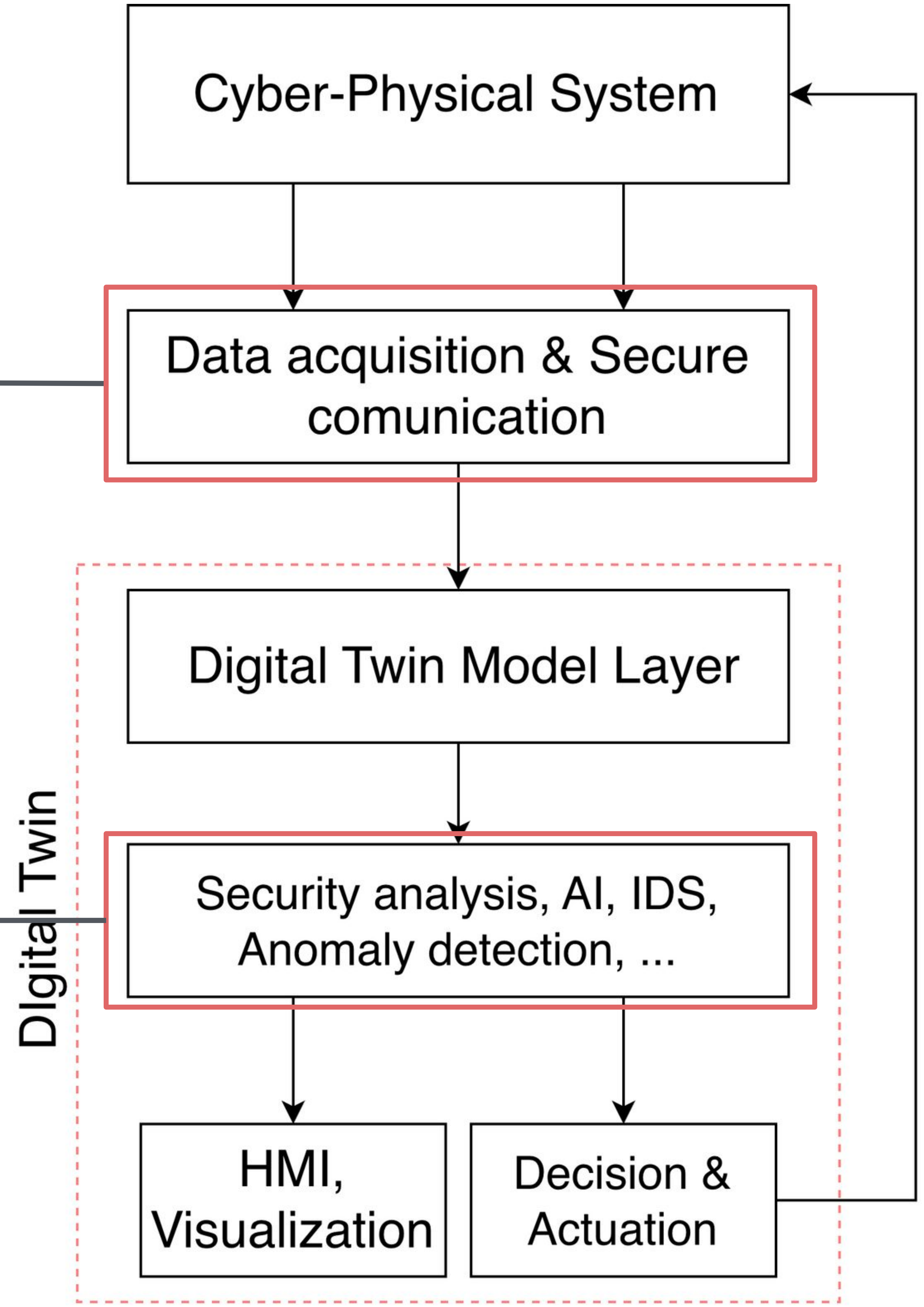
Fig 6: Dataflow using ML model

Most Digital Twins Only Mirror Data — Not System Behavior

Without an explicit model of the reference system, detection stays reactive.

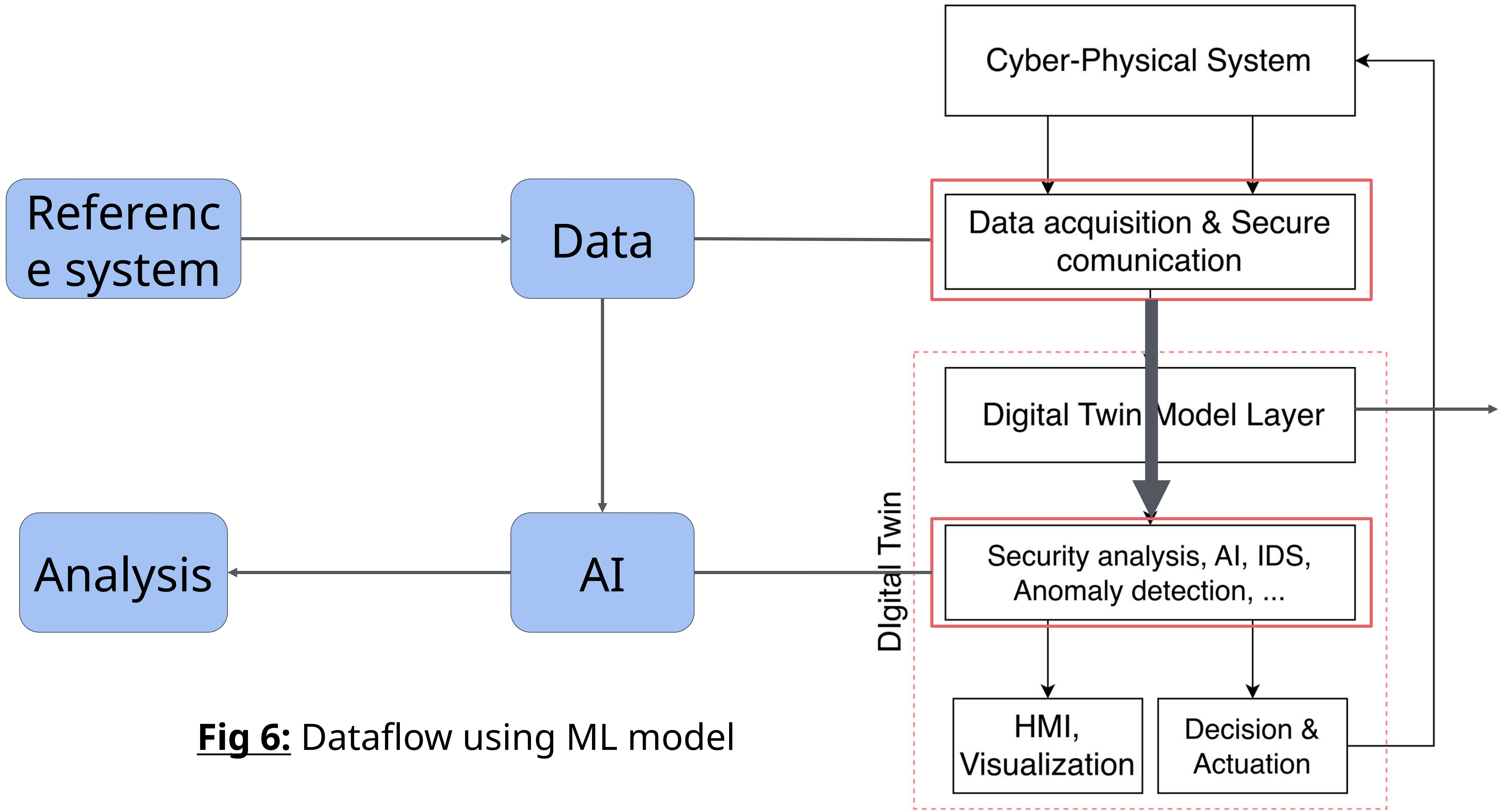


<u>Fig 6:</u> Dataflow using ML model



Most Digital Twins Only Mirror Data — Not System Behavior

Without an explicit model of the reference system, detection stays reactive.



This limits detection to reactive analysis and prevents forecasting.

<u>Fig 6:</u> Dataflow using ML model

Modelizing the Reference System Enables Proactive Estimation

From observation → to estimation → to anticipation.

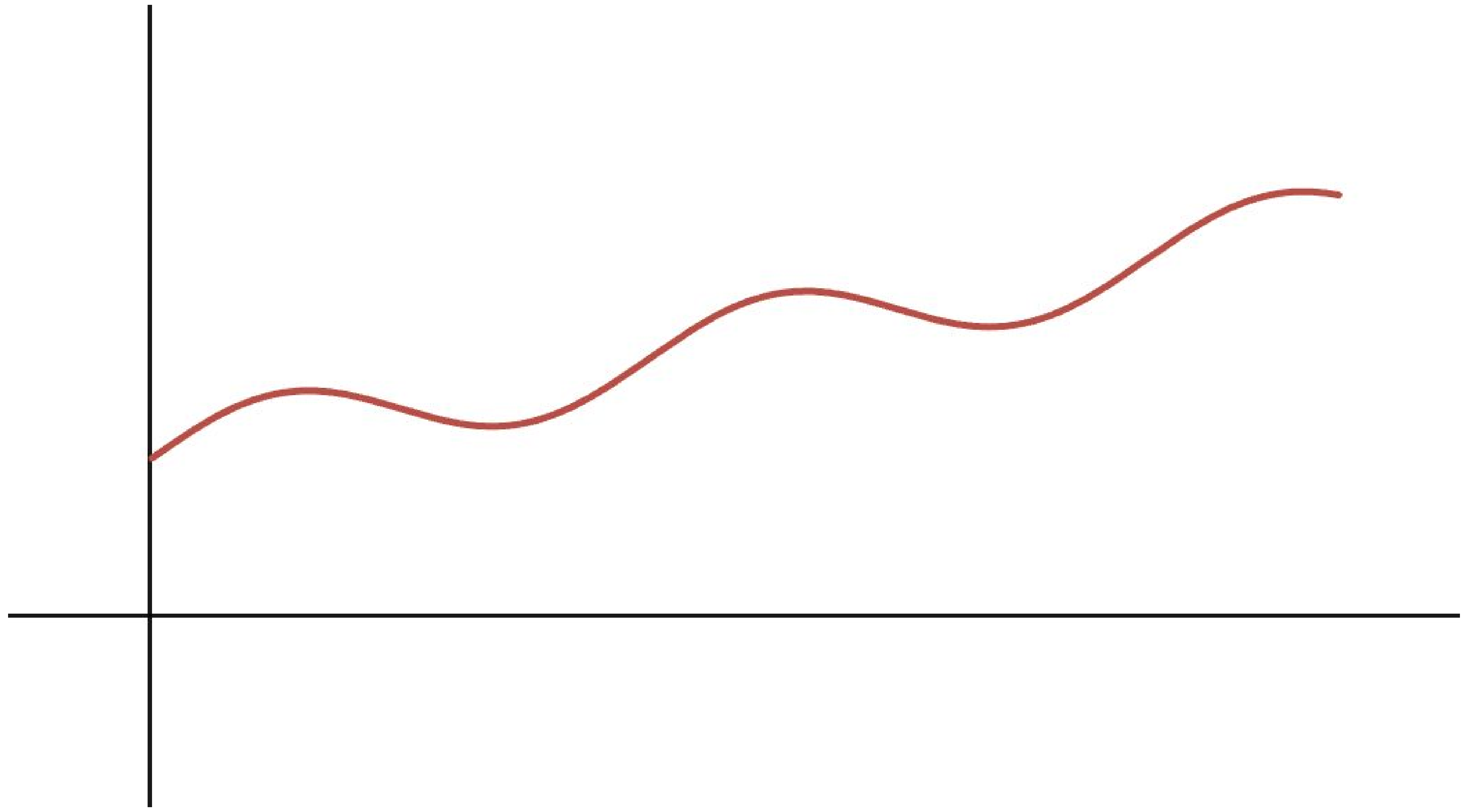
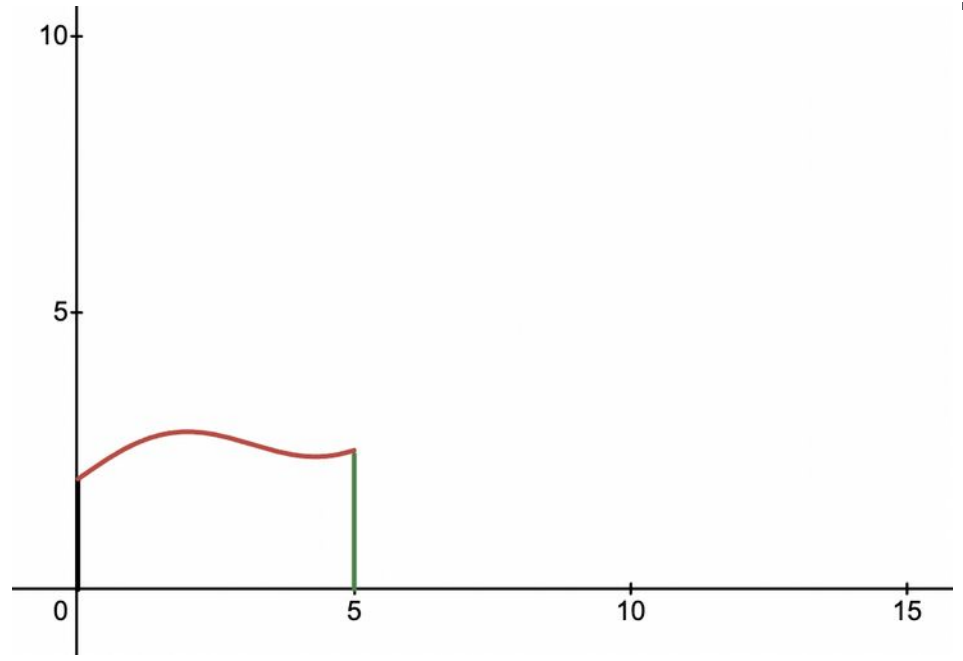
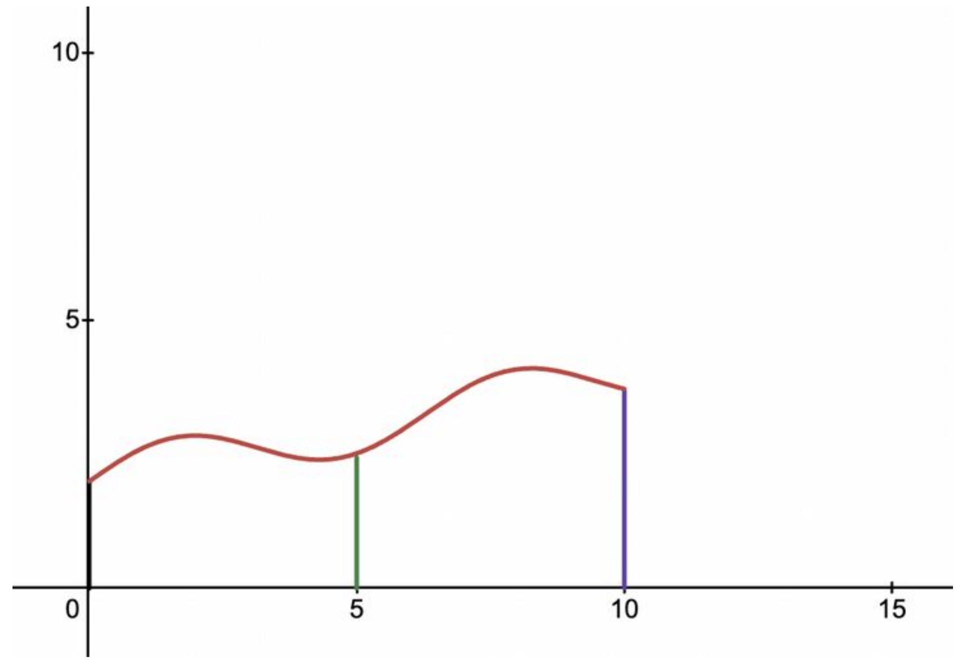
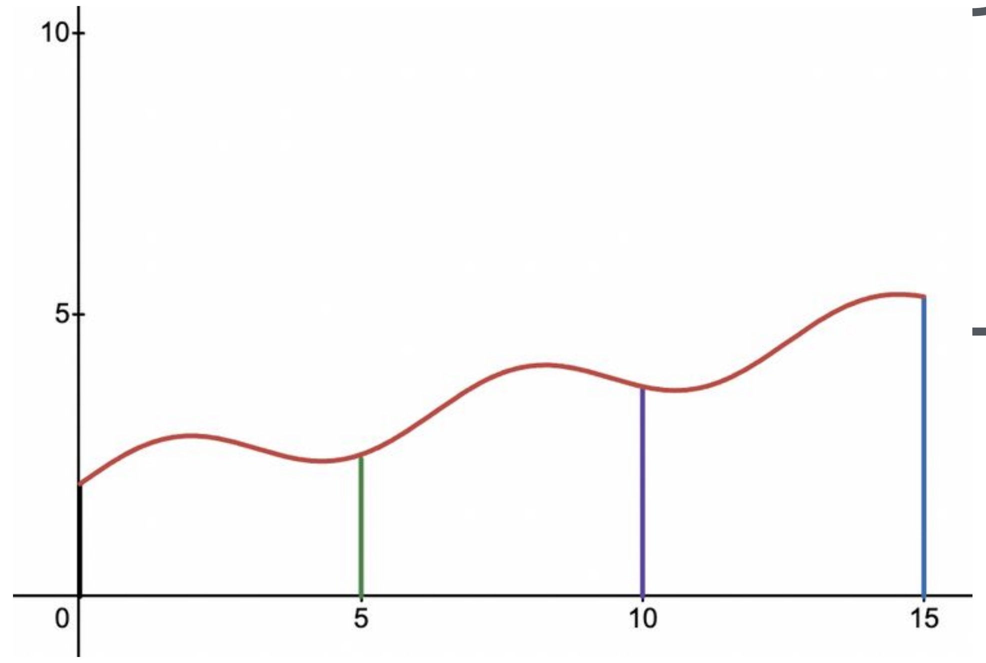


Fig 7: Known attack pattern





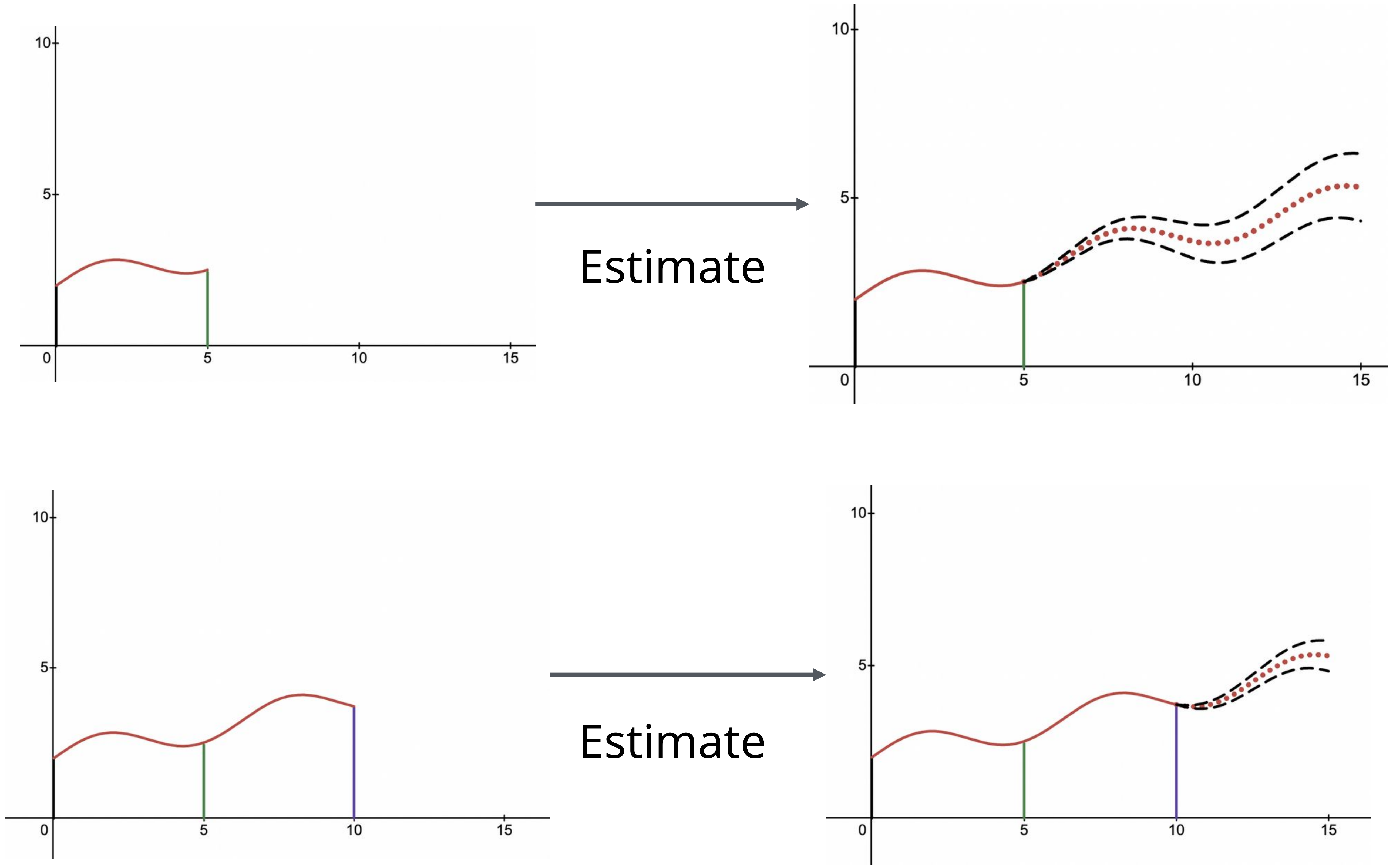
Attack ?



Attack

Modelizing the Reference System Enables Proactive Estimation

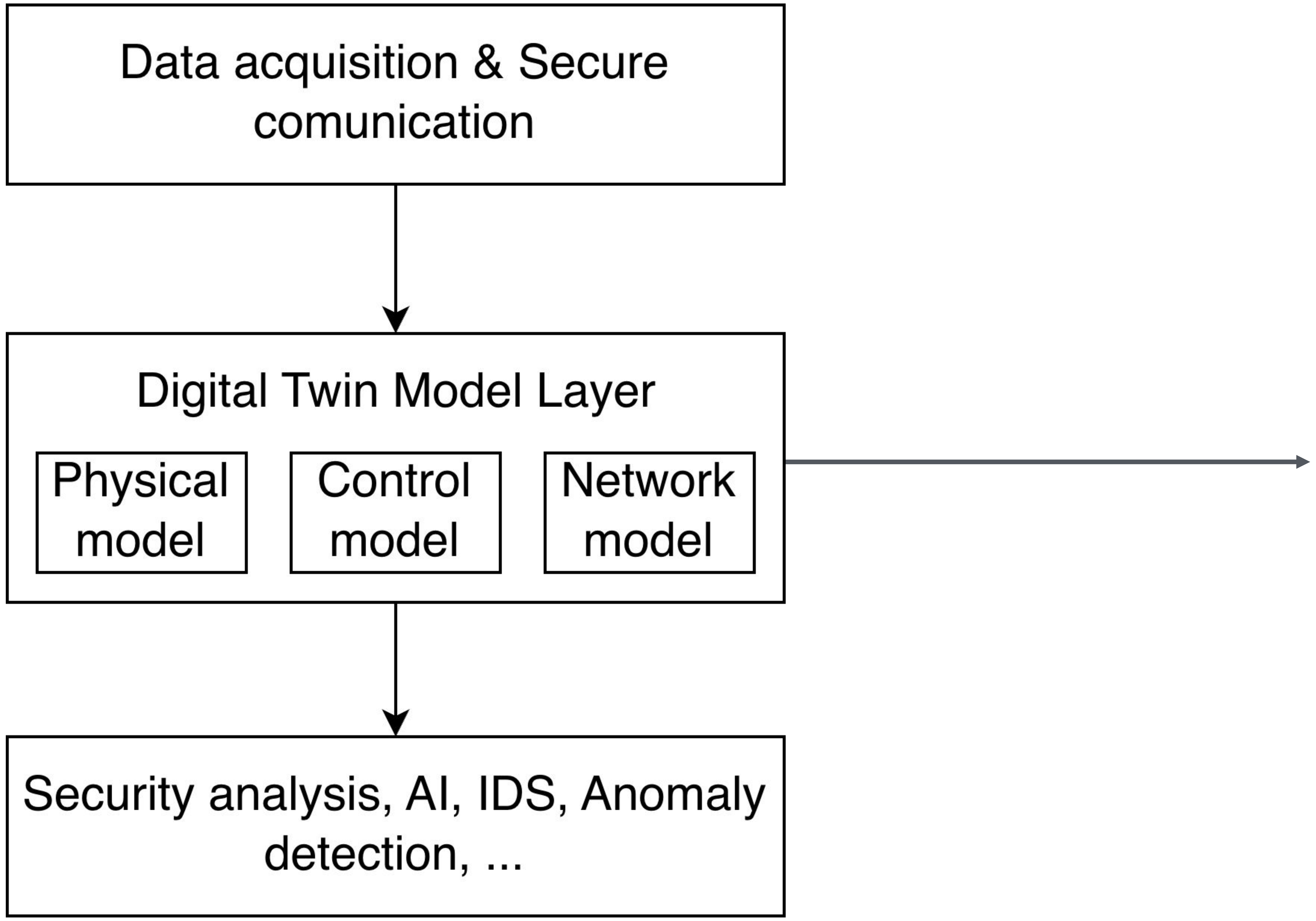
From observation → to estimation → to anticipation.



<u>Fig 8:</u> Attack estimation

From Modelization to Learning Pipelines

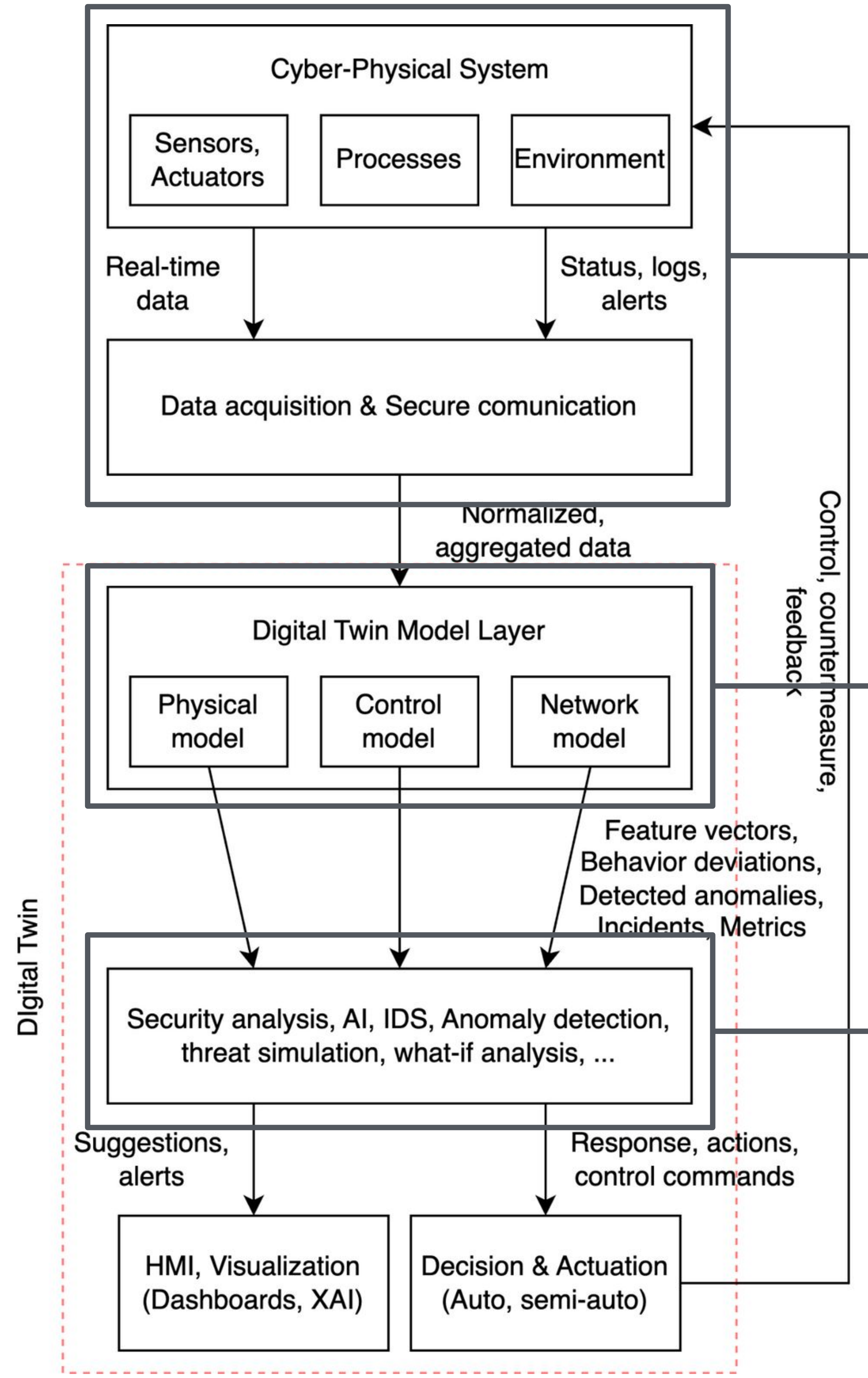
System modeling provides the foundation for safe ML training and estimation.

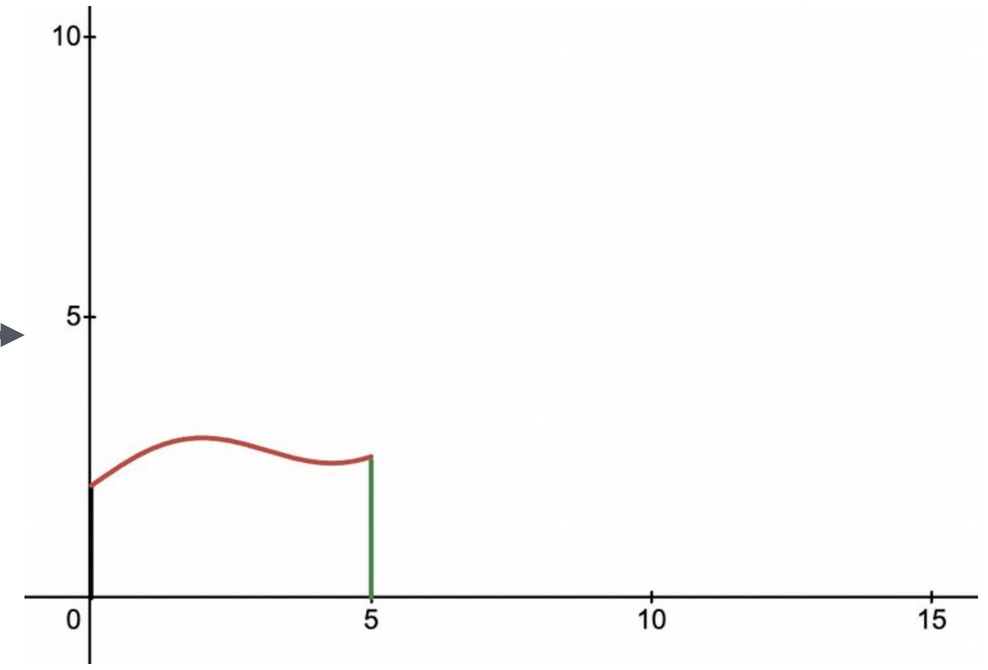


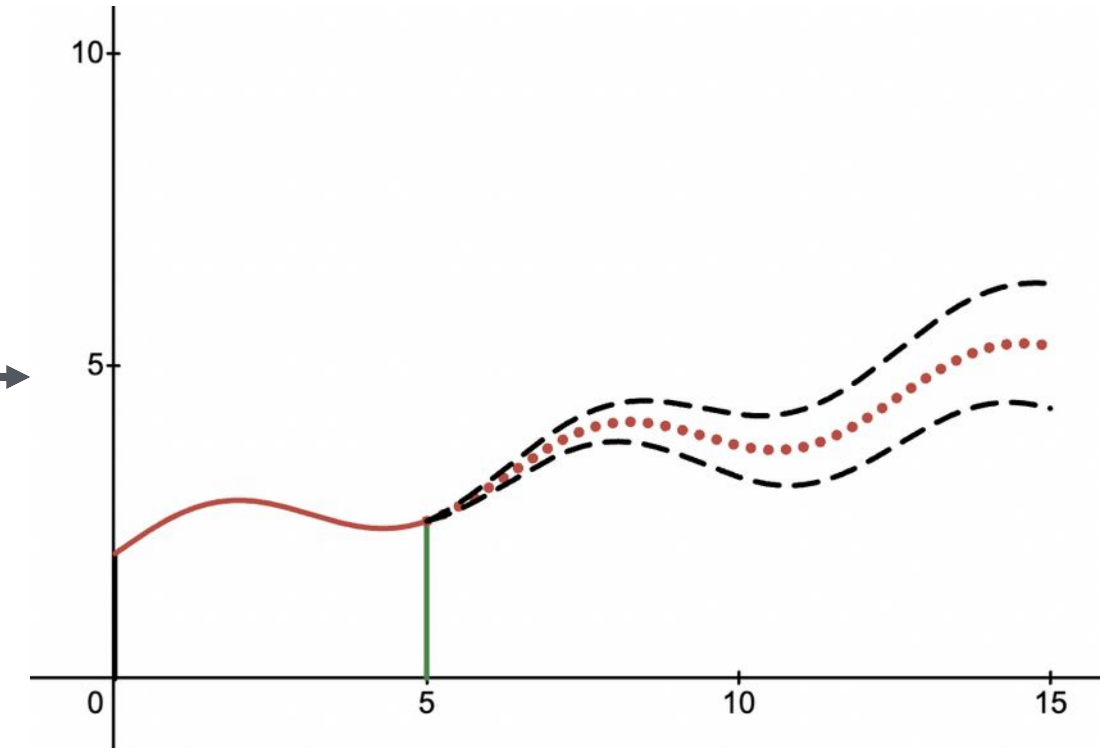
- Proactive detection with event forecasting
- Allow data validation and coherence based on model on the DT before passing it to the ML pipeline
- Can go even further by passing, settings, parameters, comportement based observation to the ML pipeline
- Allow coupling with physical parameters to enhance the detection

Modelizing the Reference System Enables Proactive Estimation

From observation → to estimation → to anticipation.







Metrics to determine probability of attack

Fig 9: Attack estimation lifecycle

Machine Learning Pipeline in the Digital Twin

Training the model safely, applying it reliably.

Training Phase (Offline)

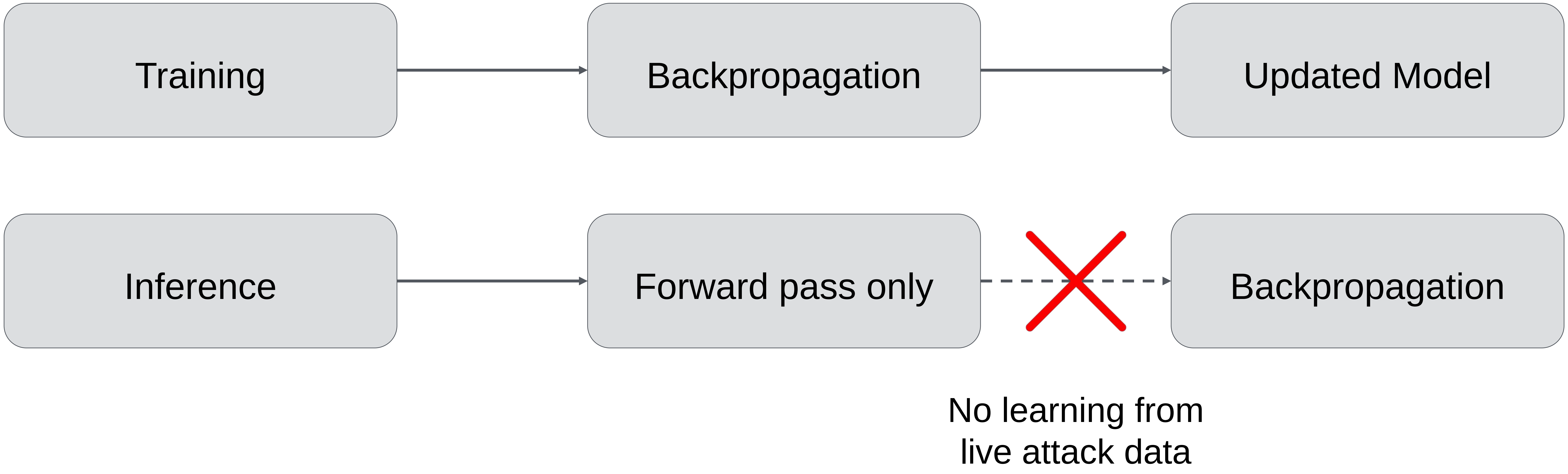
- Data: validated logs, attack simulations, side-channel measurements
- **Goal:** learn decision boundaries via *backpropagation*
- Outputs: trained weights → *frozen model*

Inference Phase (Online)

- Data: live DT data stream
- **Goal:** detect & estimate attacks using forward pass only
- No backpropagation, model is read-only

Preventing Model Contamination

Only validated data feeds training, never online events



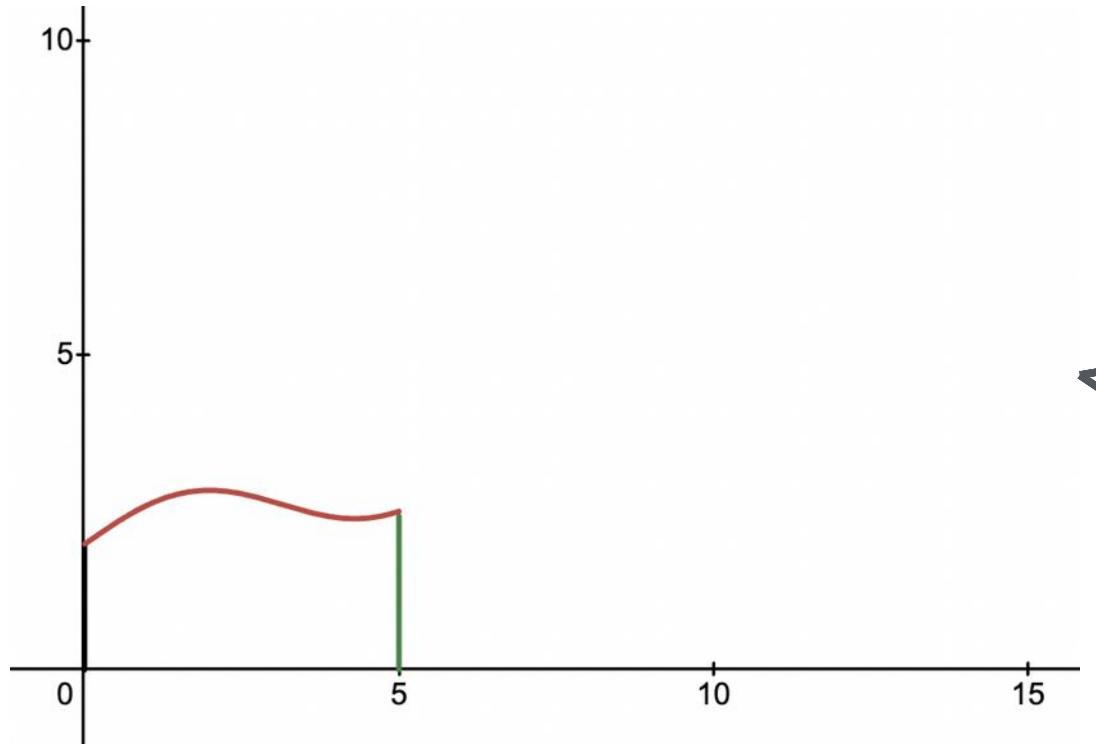
Preventing Model Contamination

Only validated data feeds training, never online events

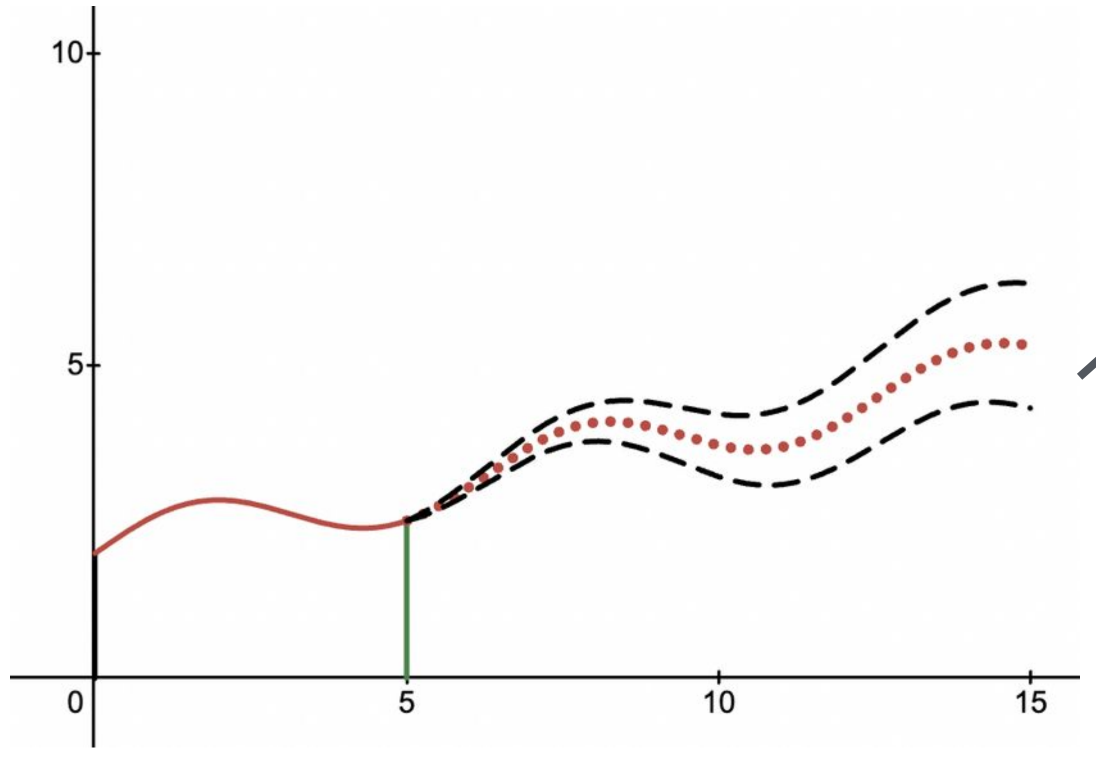
No online backpropagation, good to avoid poisoning, but no update over time?

Forecasting Mode: Testing Without Learning

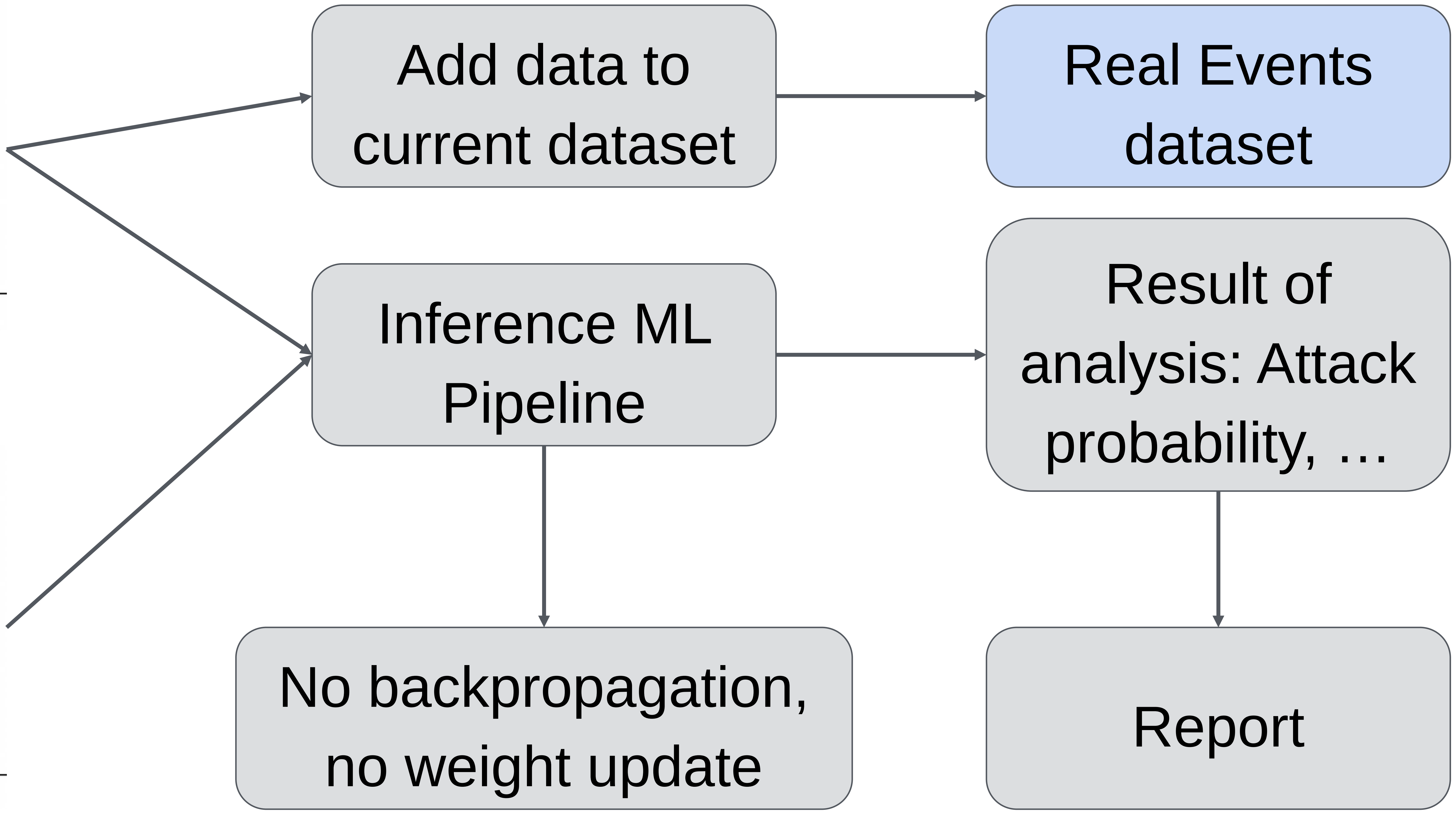
Use the trained model to simulate and estimate risks.



Occurring events

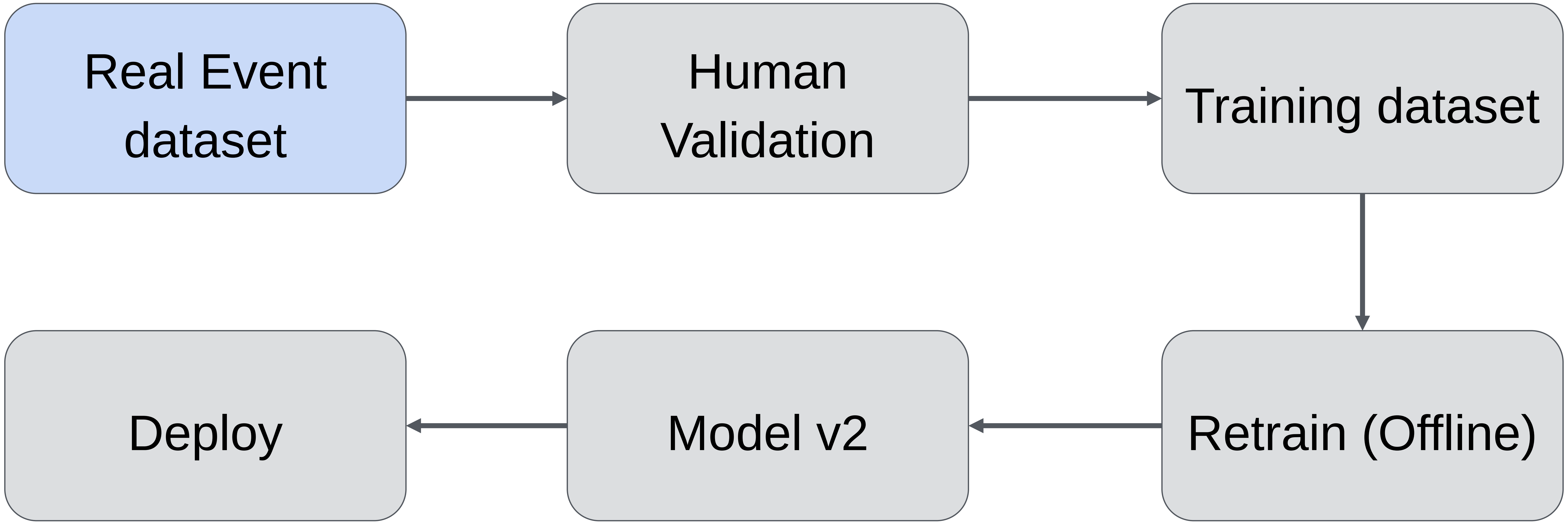


Estimations



Continuous Learning from Validated Events

Use confirmed incidents to update the model offline.



Dual-Pipeline Learning: Safe Adaptation and Reliable Prediction

Use the trained model to simulate and estimate risks.

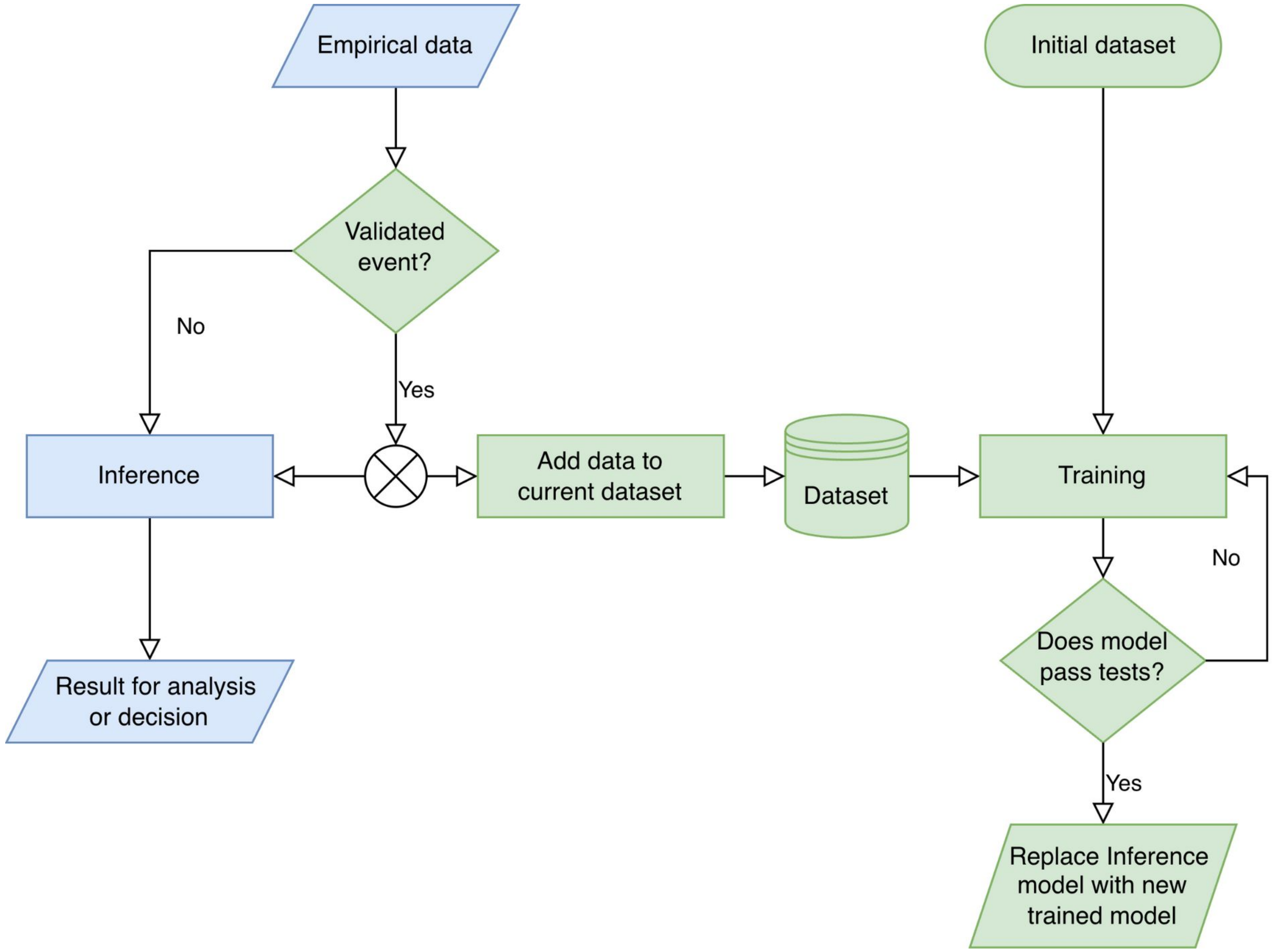


Fig 10: Dual ML Pipeline

- Training (backpropagation) happens offline on validated data.
- Inference (forward pass) happens online, no learning.
- Events can later enrich the training dataset *after validation*.
- This ensures robustness, traceability, and resistance to data poisoning.

3

Future work

Toward Predictive Digital Twins for Cyber-Physical Security

Integrating Physical Parameters for Attack Detection

Expand observability: Incorporate measurable physical indicators into the Digital Twin.

Enhance estimation: Use these parameters to improve attack precision, anticipate consequences, and localize anomalies.

Validate prediction: Evaluate how added physical features affect accuracy and detection latency across datasets and testbeds.

Temperature of a component

Human interactions (Operators)

CPU Load

Physical access

Identification of physical impact

Electric consumption

Fan speed

Noise of the system

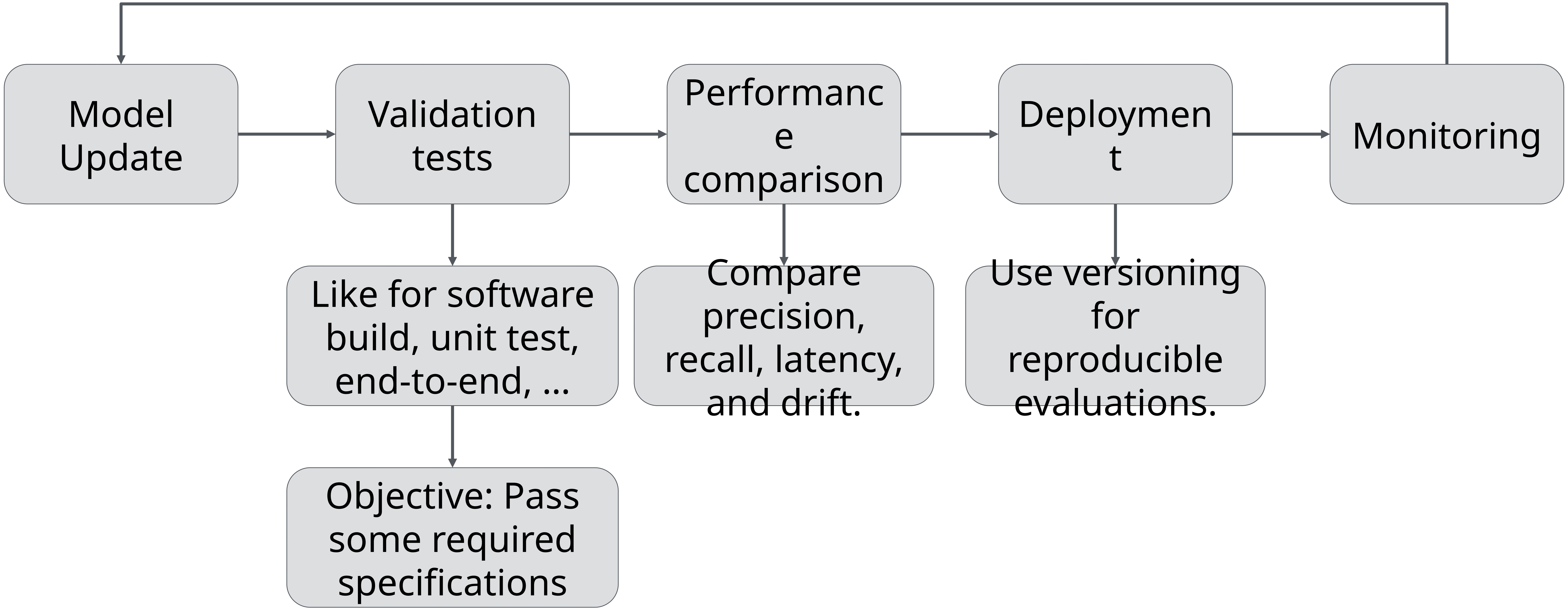
Memory usage

GPU Load

Fig 10: Hardware parameters

Ensuring Continuous Trust in the ML Pipeline

Automated validation and regression testing for DT-based ML.



3

Challenges

Key Challenges: Data and Knowledge

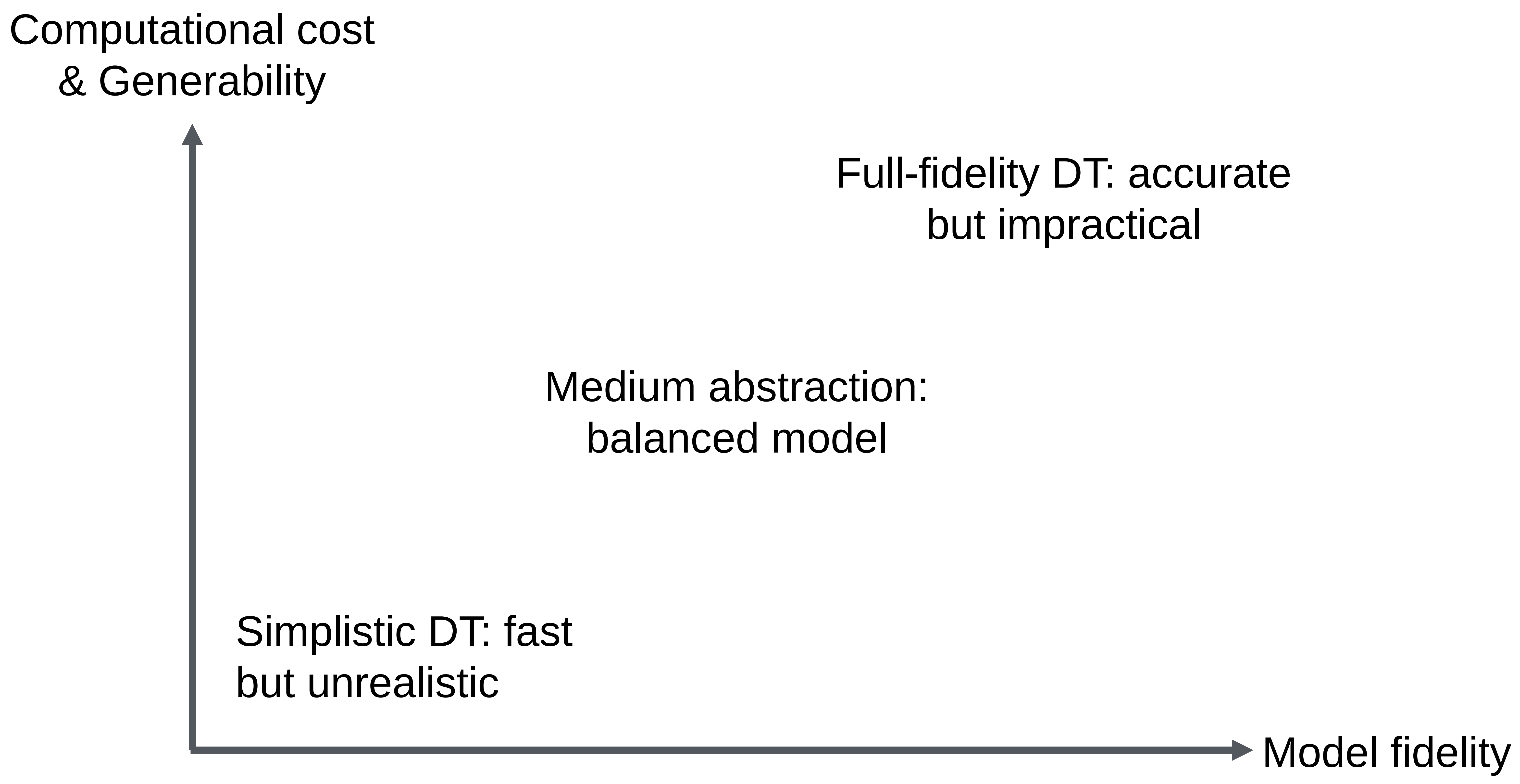
High-quality, multi-domain data remains the main bottleneck.

Aspect	Label
📊 Database	Data realism & quality (CPS + side-channel)
🌍 Sources	Diversity & representativeness
🧩 Correlation	Link physical and logical worlds
📈 Precision	Benchmarking & validation

Multi-domain, synchronized datasets are the foundation but remain scarce.

Key Challenges: Modele usage

Model validation

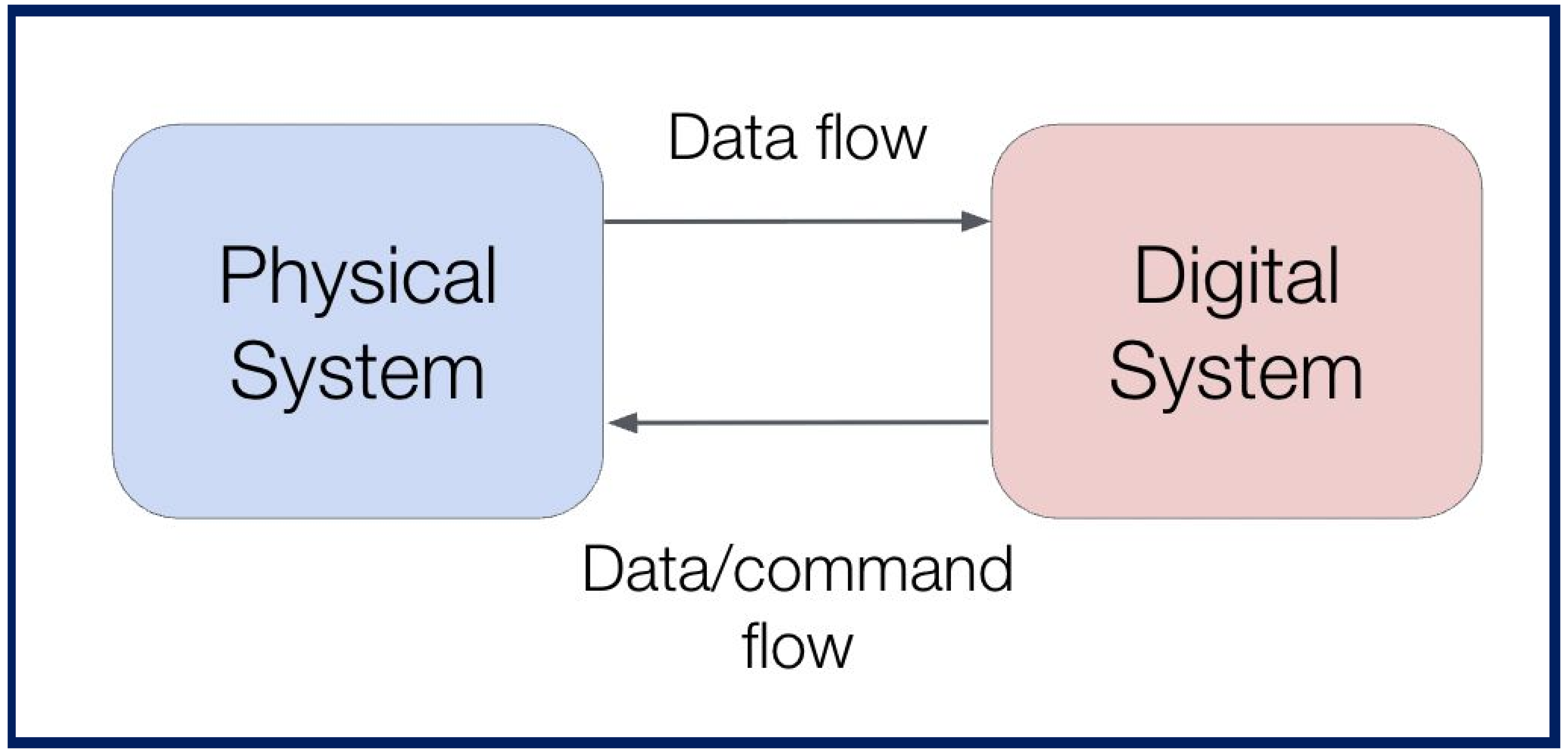


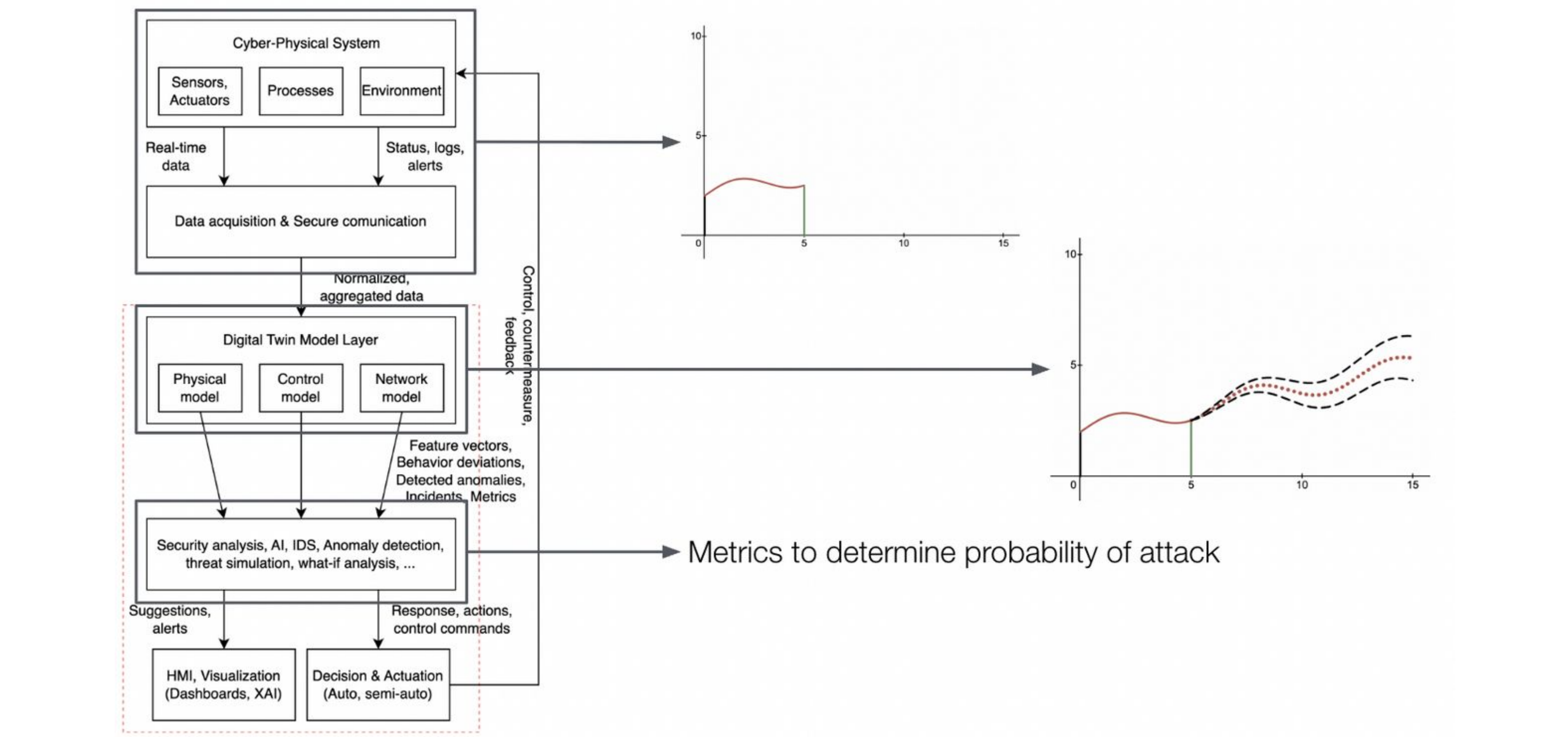
- Validating DT behavior vs. real system is complex. ⇒ How to modelize the comportment of the reference system? Physical equation, correlation, …
- ML trained on one DT may fail in another due to parameter shifts.

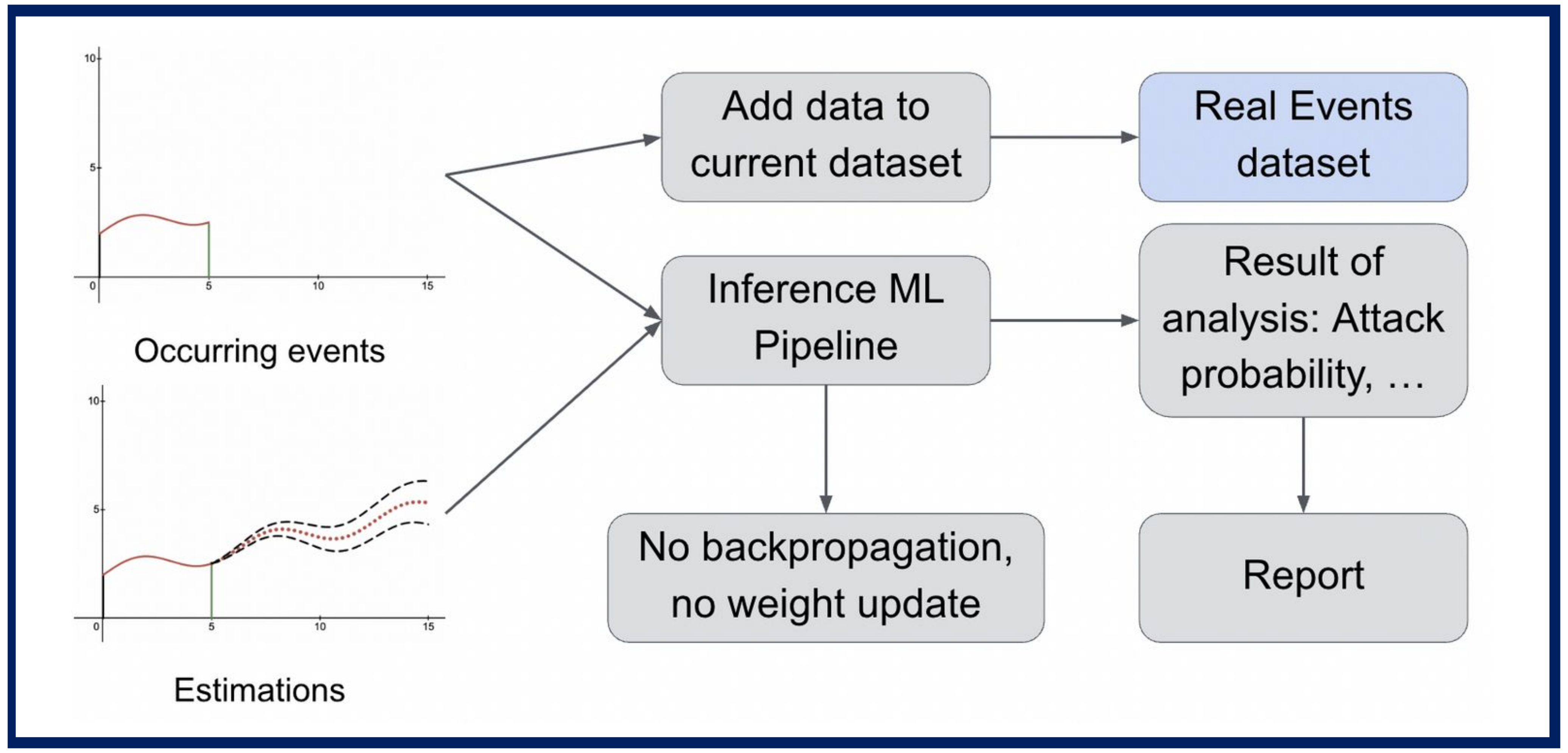
3

Summary

My presentation in a Nutshell







Integrating Physical Parameters for Attack Detection

Expand observability: Incorporate measurable physical indicators into the Digital Twin.

Enhance estimation: Use these parameters to improve attack precision, anticipate consequences, and localize anomalies.

Validate prediction: Evaluate how added physical features affect accuracy and detection latency across datasets and testbeds.

Contact: hugo.bourreau@imt-atlantique.fr

References

[1] Varghese, S. A., Dehlaghi Ghadim, A., Balador, A., Alimadadi, Z., & Papadimitratos, P. (2022). Digital Twin-based Intrusion Detection for Industrial Control Systems. *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, 611-617. https://doi.org/10.1109/PerComWorkshops53856.2022.9767492

[2] Jyothi, R., & Jagadeesha, R. (2024). Next-Gen Threat Detection : Leveraging AI and Cyber Twin Technologies for IoT Security. *2024 First International Conference on Software, Systems and Information Technology (SSITCON)*, 1-6. https://doi.org/10.1109/SSITCON62437.2024.10796384

[3] Pisani, J., Cavone, G., Pascucci, F., & Giarré, L. (2023). Using Digital Twin to Detect Cyber-Attacks in Industrial Control Systems. *IEEE EUROCON 2023 - 20th International Conference on Smart Technologies*, 467-471. https://doi.org/10.1109/EUROCON56442.2023.10198927

[4] Al-Qirim, N., Majdalawieh, M., Bani-hani, A., & Al Hamadi, H. (2025). Cyber threat intelligence for smart grids using knowledge graphs, digital twins, and hybrid machine learning in SCADA networks. *International Journal of Engineering Business Management*, *17*, 18479790251328183. https://doi.org/10.1177/18479790251328183

[5] Sasikala, M., Mahaboob John, Y. M., Jothi, B., S, N., & S, S. K. (2024). Integrating Digital Twins with AI for Real-Time Intrusion Detection in Smart Infrastructure Networks. *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)*, 1-6. https://doi.org/10.1109/IACIS61494.2024.10721892

[6] Mayan, A., Krishanveni, S., & Jothi, B. (2024). AI Enabled Digital Twin Models to Enhance Security in Smart Cities. *2024 International Conference on Intelligent Computing and Sustainable Innovations in Technology (IC-SIT)*, 1-6. https://doi.org/10.1109/IC-SIT63503.2024.10862955

[7] Supriya, K. S., P, J. L. S., Arora, R., Bhatia, R., Yadwad, S., & L, N. (2024). Securing IoT Systems with AI-Infused Software and Virtual Replica Models. *2024 International Conference on Integrated Intelligence and Communication Systems (ICIICS)*, 1-6. https://doi.org/10.1109/ICIICS63763.2024.10860178

[8] Ukwuoma, H. C., Dusserre, G., Coatrieux, G., Vincent, J., & Ahmed, N. B. (2024). Optimising Intrusion Detection in Cyber-Physical Systems. *2024 8th Cyber Security in Networking Conference (CSNet)*, 7-14. https://doi.org/10.1109/CSNet64211.2024.10851766