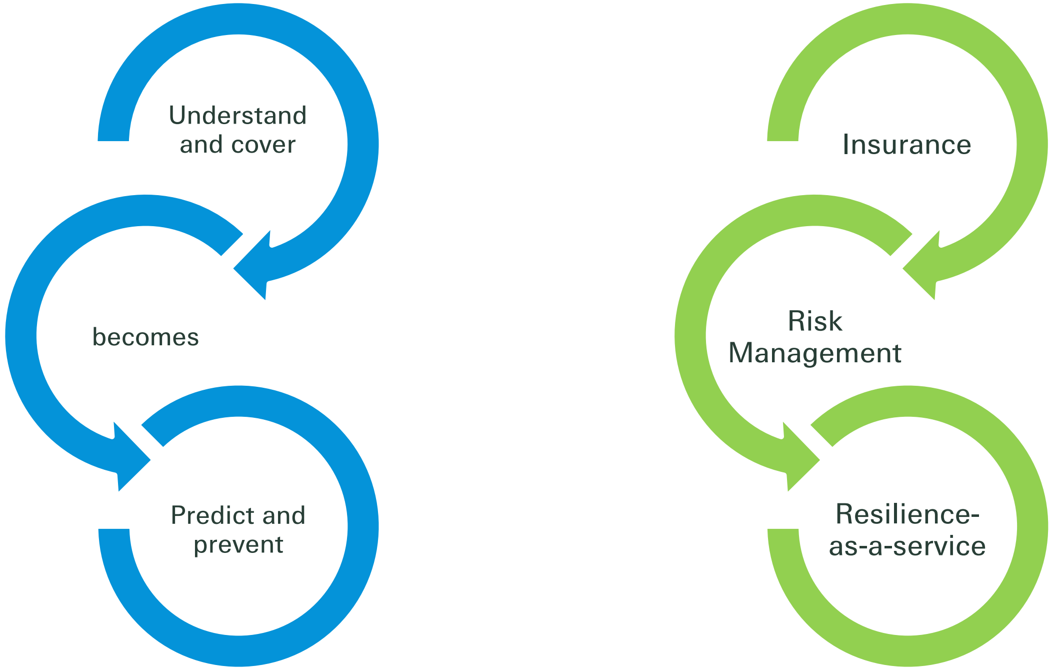
Swiss Re Institute

Physics-based/machine learning (ML) hybridized modeling

Dr. Jeffrey Bohn, Senior Advisor, Swiss Re Institute
ICMIF webinar, 9 June 2021

Shifts in the insurance industry



We are drowning in information, while starving for wisdom. The world henceforth will be run by synthesizers, people able to put together the right information at the right time, think critically about it, and make important choices wisely.

E.O. Wilson

Data deluges, advanced algorithms, and powerful computational tools enable physical and natural system modeling like never before.

Data

IDC forecasts worldwide data to grow by CAGR of 23% to 181 ZB till 2025. A third of these data will be real-time.

Advanced algorithms

Hybrid algorithms can lead to better data curation by addressing issues related to data quality and lack of compute power.

Better processing

Synchronized edge and cloud computing can ease data processing by on-demand access to computing resources.

Computational tools

Modern computational tools' ability to study complex systems enable extreme events analysis at multiple levels.

Swiss Re Institute

Source – Swiss Re Institute

Simulating physical phenomena is evolving from component design to systems assembly to developing *digital twins* of physical assets

Early steps in modelling

Timeline

- Development of basic models
- Better understanding of phenomena

3D component design

~ 1985

- Rapid advances in 3D modelling from computers
- Use of computer aided technology (CAD) in product component design

Holistic systems assembly

~ 2000

- Advances in model-based systems engineering
- Holistic approach to systems assembly

Physics aware digital twin

~ 2015 onwards

- Hybridization of ML by combining the virtual and physical world
- Creation of reduced order models (ROM) to bridge value chains

Key challenges in physics aware ML implementation

- Parametrizations of complex real-world processes
- Keeping physical and digital worlds 'in sync' easily
- Closing the data loop from operations back to design
- Generating knowledge from distributed models
- Overcoming expertise-limited scalability of use
- Applying novel simulation technologies and convergence with data analytics and IoT

Swiss Re Institute

Source – 1) Connecting physics based and data driven models: The best of two worlds, Siemens AG, 2018
2) Swiss Re Institute

Why aren't generalized linear models good enough?

Challenges

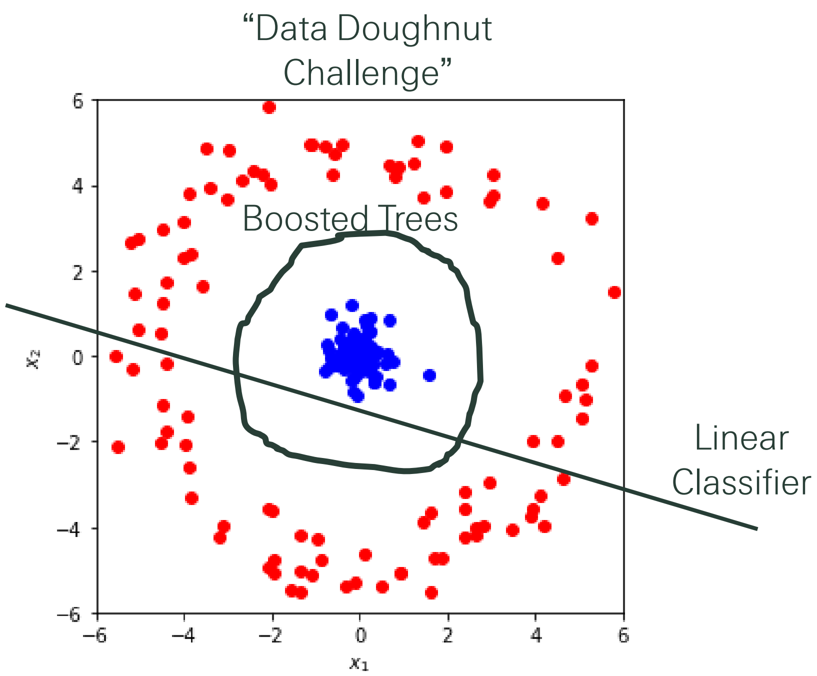
- Data challenges
  - Sparsity
  - Noise
  - Confounders
- Model challenges
  - Non-linear relationships
  - Frequent regime shifts
  - Overfitting risk
  - Complexity

Solutions

- Alternative data
  - Non-standard structured
  - Unstructured
  - Meta
- (More) Data curation
- Regularization
  - Model complexity constraints
  - Incorporate better loss functions
  - Combine "weak learners" i.e., boosting
- ***Model hybridization***

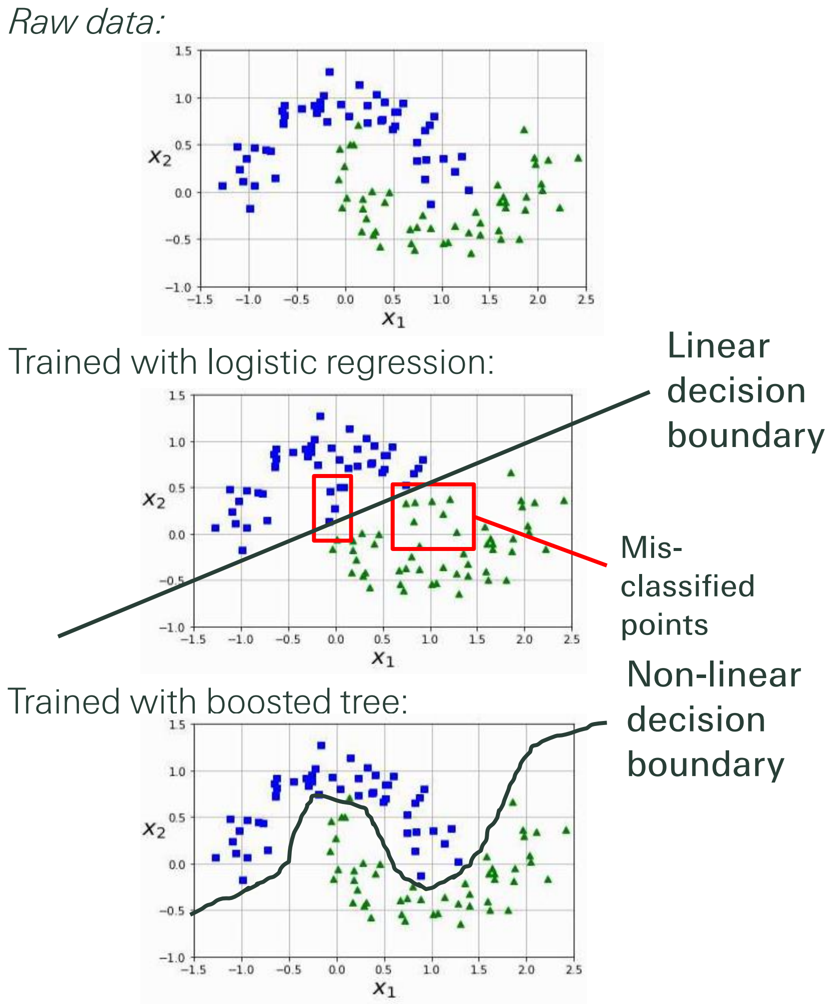
Data limitations & imperfections drive the challenges

“Data Doughnut Challenge”: Capturing non-linear relationships



- When addressing non-linear data relationships, more complex algorithms ensure higher accuracy than simple algorithms.
- Looking at the “Data Doughnut Challenge” graphically illustrates how complex algorithms can solve non-linear problems. Challenge lies in how to classify data in light of non-linear clustering.

Self-trained example – with “make moons” dataset



Data-value-chain process as part of an enterprise data fabric (part 1)
New data sources becoming more important: ***Meta, Unstructured, Privacy-preserved, and Synthesized/Simulated***

Identify
- Compare data catalogue
- Explore data ontologies
- Review compliance

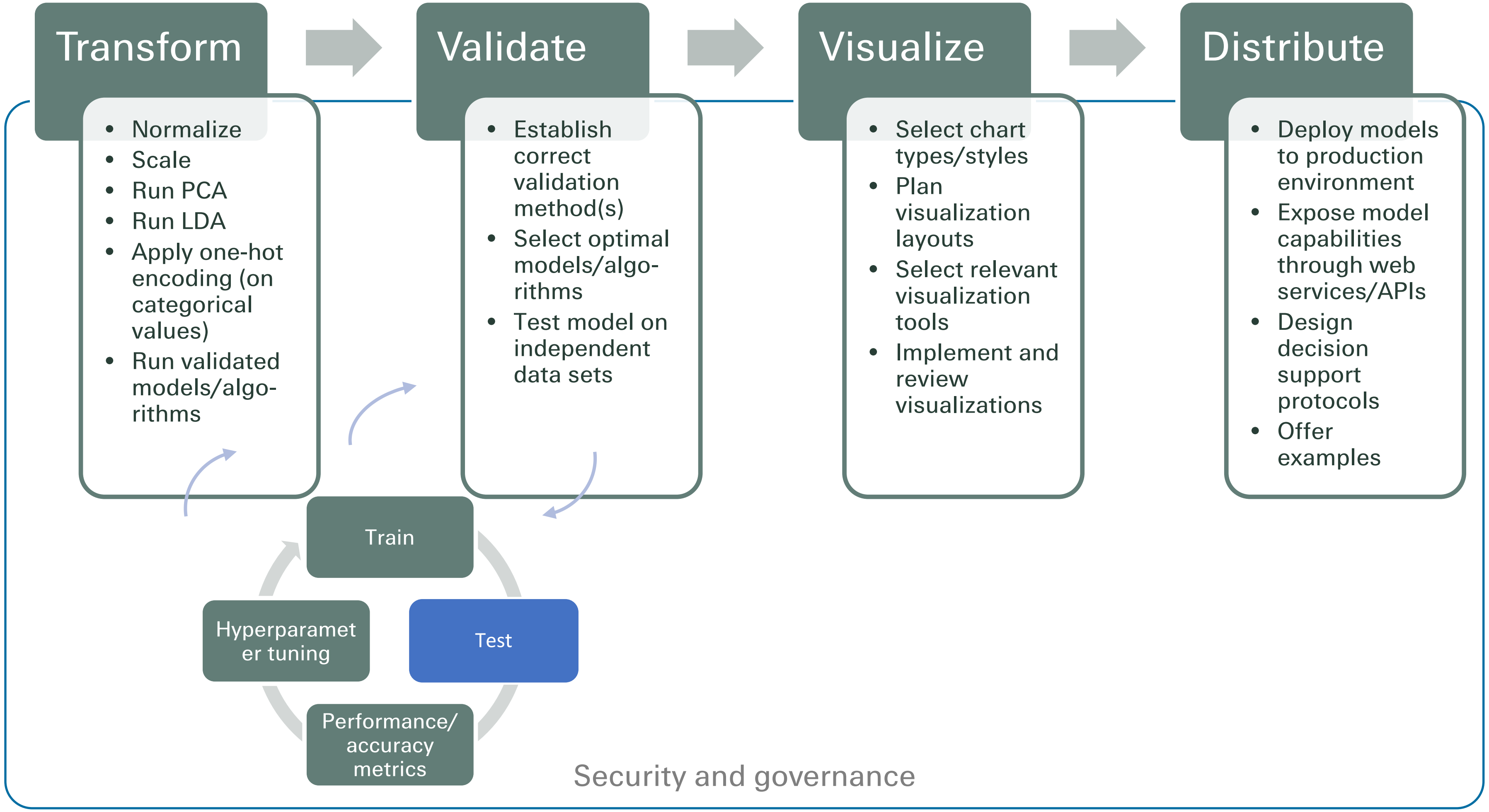
Ingest
- Retrieve and store data from source systems, file feeds, web services, etc.

Curate
- Clean
- Combine
- Standardize
- Conform
- Interpolate missing data *(e.g., EM algo)*
- Test data items against criteria: i) Possible ranges; ii) Plausible ranges; iii) Plausible relationships

Collaborate
- Provide access credentials
- Create and share notebooks
- Visualize data
- Run exploratory analytics
- Develop models

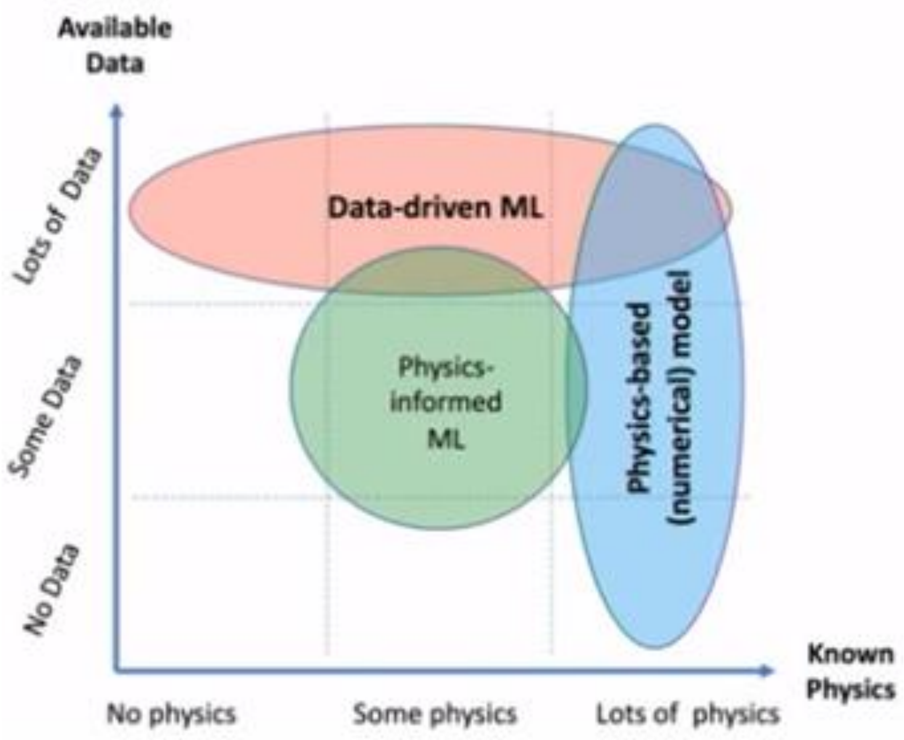
Security and governance

Data-value-chain process as part of an enterprise data fabric (part 2)



* Optimal architecture separates the development environment (i.e., starts with train) from the production environment

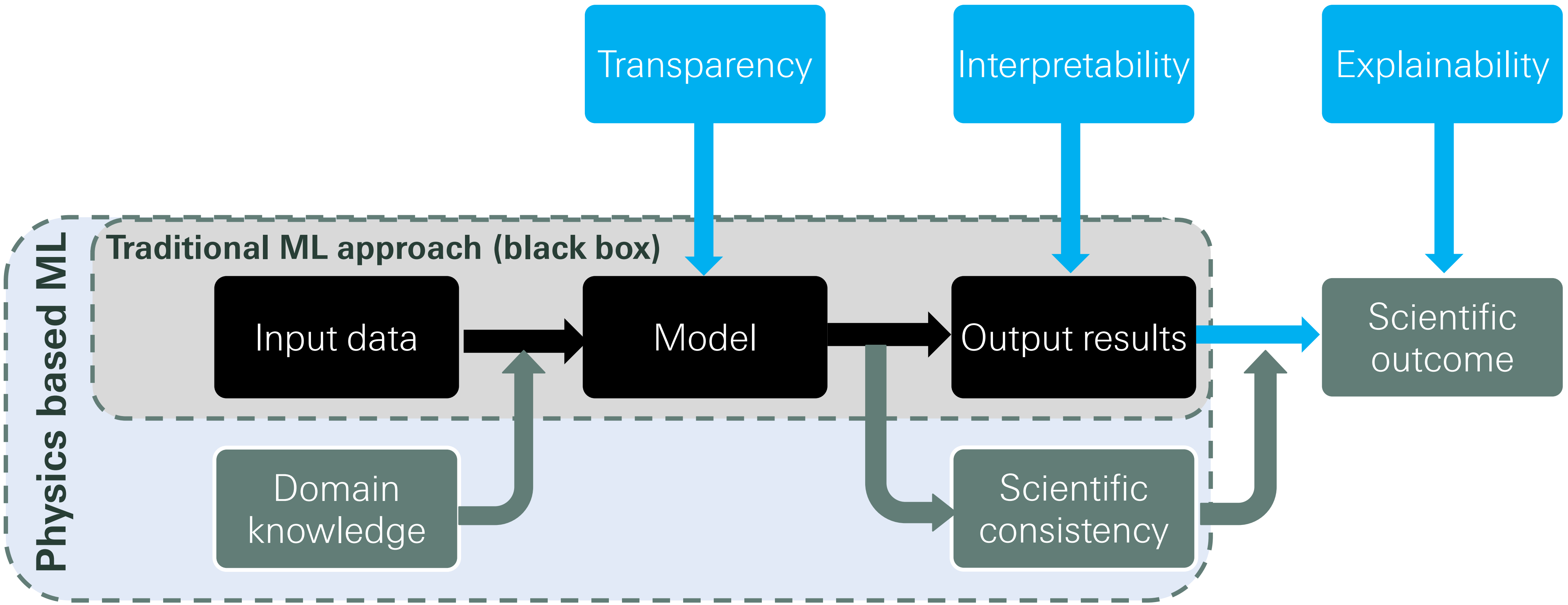
Physics-based, hybridized machine learning approaches can offer the best of data science and mathematical models to develop new hybrid solutions



Source – 1) Pacific Northwest National Laboratory
2) Swiss Re Institute

Approaches	Advantages	Disadvantages
Pure science or physics-based approach	• Tried and tested. • Explainable. • Governing equations. • Structure and stability preserving. • Predictive (error estimators).	• Slower. • Many assumptions • May not factor in new data. • May not capture relationships.
Purely data intensive machine learning	• Multidimensional analysis. • Discover hidden structures. • Non-intrusive implementation. • Flexible, accessible & available.	• Not explainable – Blackbox. • Data intensive. • Does not respect physical constrains. • Noisy and incomplete data.

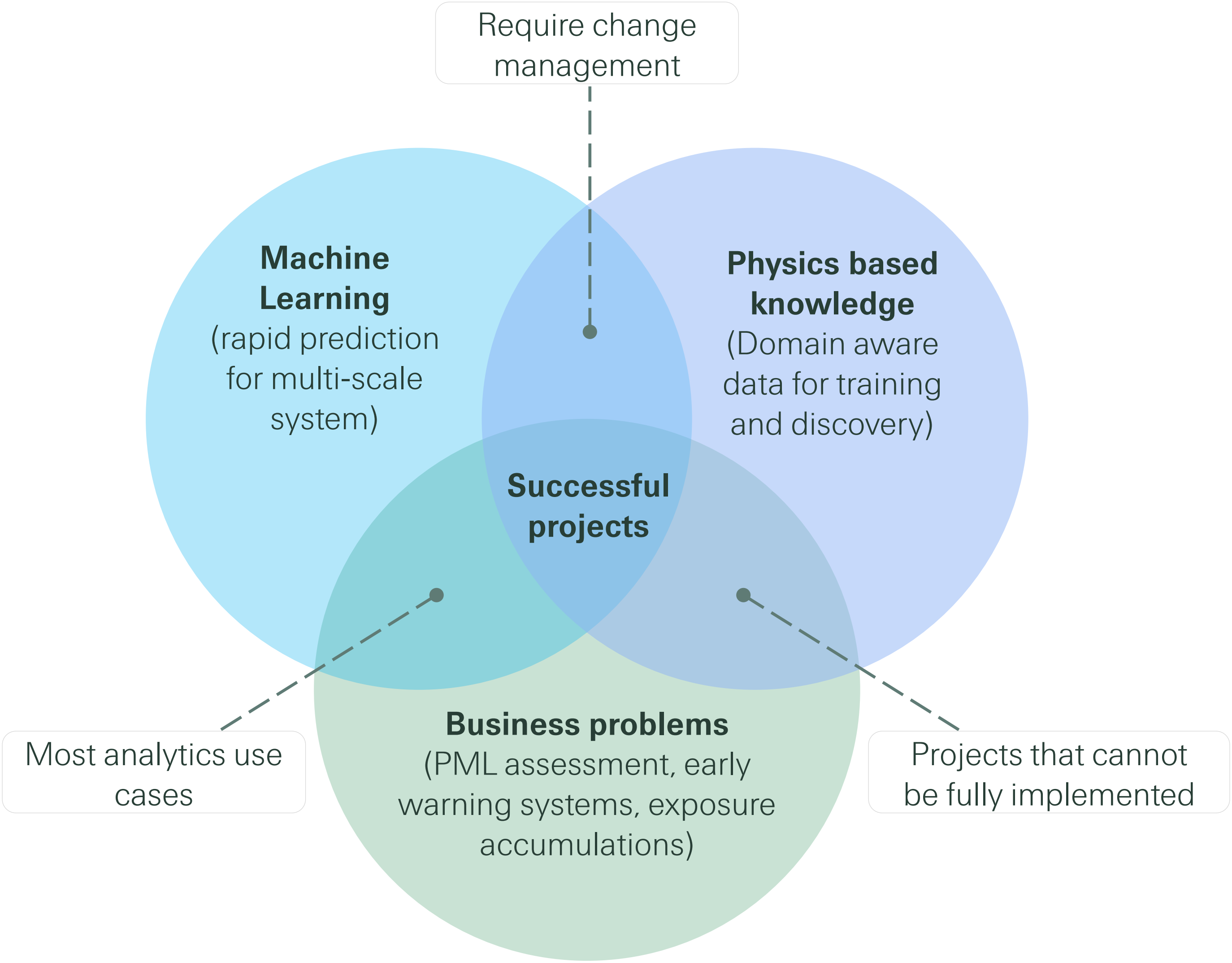
Scientific outcomes can be made more consistent, transparent, and explainable by combining physics-based domain knowledge with ML models



Source – 1) Explainable Machine Learning for Scientific Insights and Discoveries, IEEE, 11 March 2020
2) Swiss Re Institute

Careful selection of physics-based machine learning projects can enable productive enterprise scale transformation at insurers

- **Physics-based** reduced order models of complex assets and processes combined with **machine learning** can allow Re/Insurers to **uncover hidden entanglements** between insured assets and the external world.
- **Solutions can be made available to clients** via scalable SaaS platforms for better monetization. Internally, these can be applied to synthesize exposure data, claims data and physical models to better quantify and monitor risks.
- **Successful** physics aware machine learning **projects need** substantial investment and **cross-industry collaboration** for alignment of interest between insurers, governments and other stakeholders.

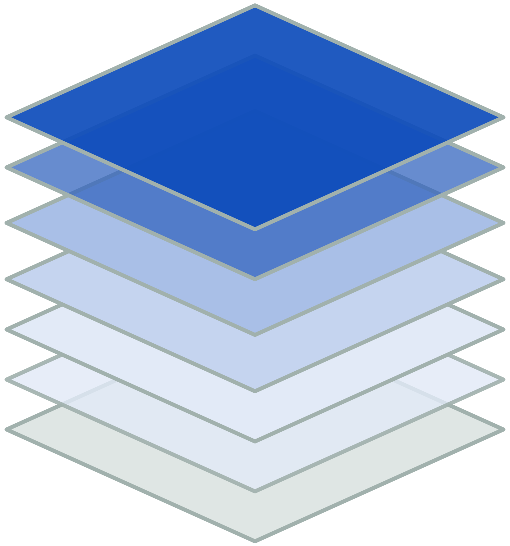


Physics-based modelling of cities could allow insurers to use a systems approach to assess the impact of extreme events on each layer

Physical footprint of a city

- Transit system data
- Water system data
- Utility system data
- Critical infrastructure & hubs (CIH)
- CIH dependencies
- Asset footprint data
- Natural environment data

Digital footprint of a city



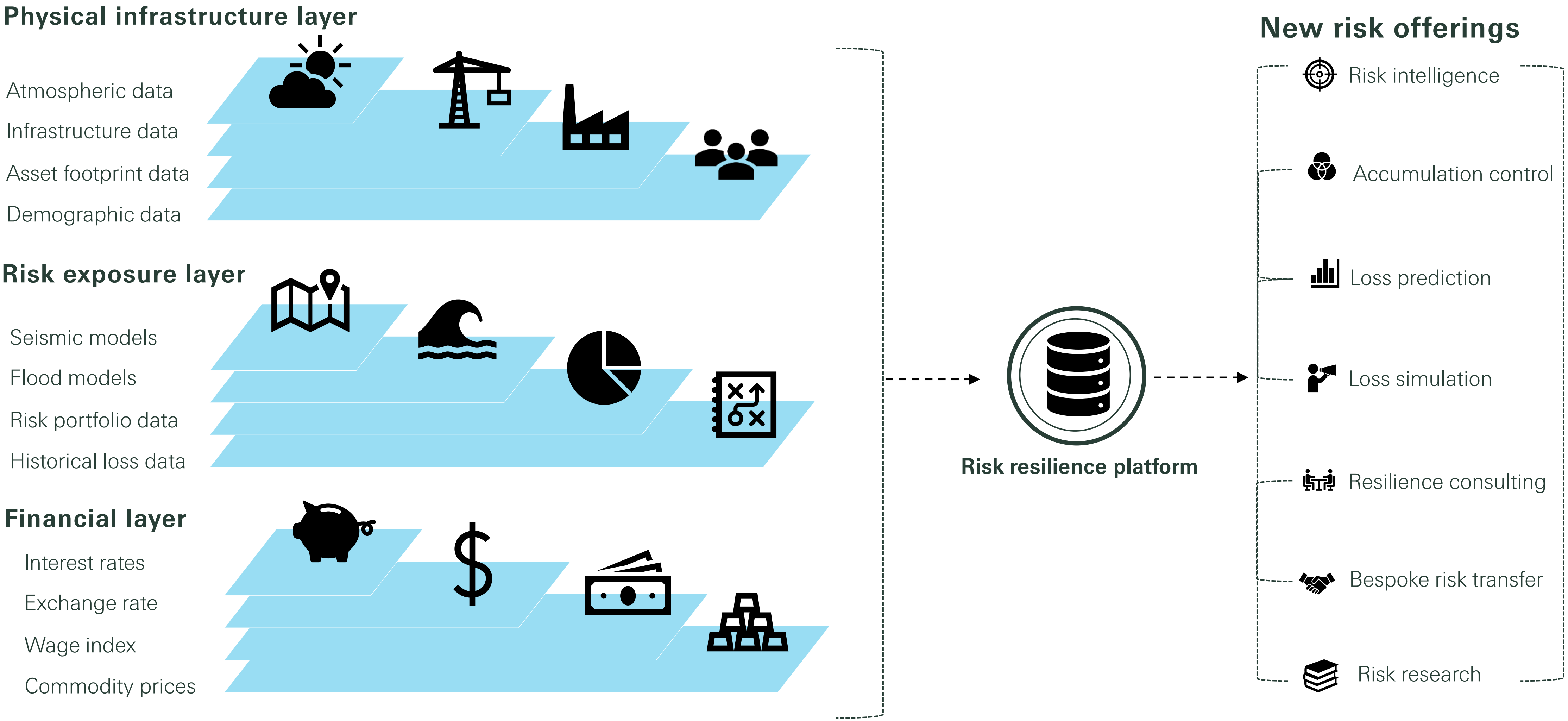
Risk footprint of a city

- Seismic impact analysis
- Flood impact analysis
- Wildfire Impact analysis
- Supply chain vulnerability

Swiss Re Institute

Source – Swiss Re Institute

Physics-based resilience models can help insurers develop new risk offerings and improve their portfolio view for pricing, reserving and large event losses



Case Study

- **One Concern**, a California based start-up, and **Sompo**, one of Japan's leading insurance companies, deployed a hybridized physics-based/machine learning (ML) based disaster prevention and mitigation system for real-time prediction of damage from earthquakes and floods in Kumamoto City, Japan

- The system uses a combination of physics-based models and ML for model development and validations:
  - Physics-based models for simulating earthquake and flood events and their impact (hazard and vulnerability analysis)
  - ML to derive missing building attributes, and to train the system based on a variety of data (damage data from historical events, and live incidents such as detection of river water levels and earthquakes)

Any questions?

Legal notice