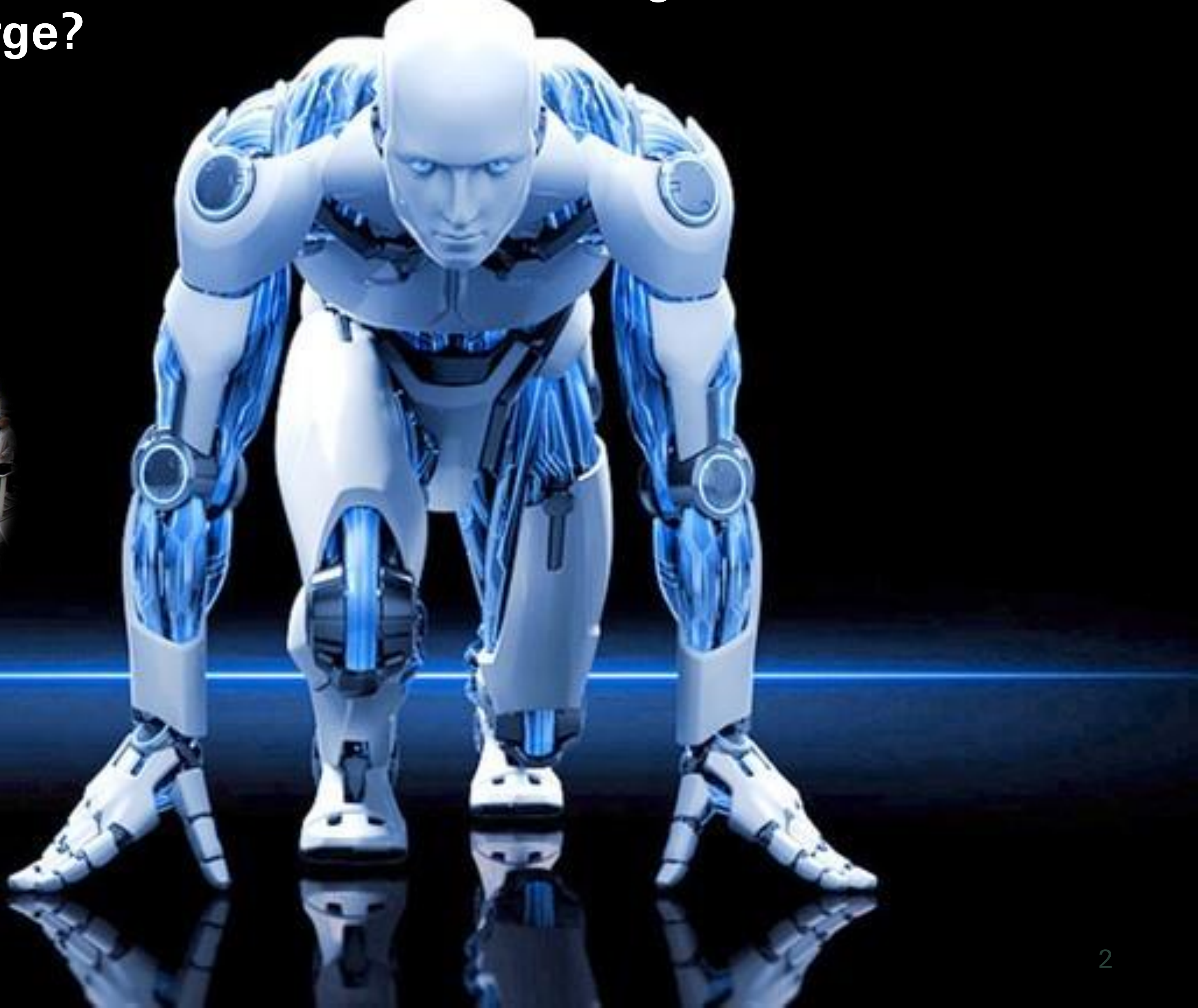
Swiss Re Institute

Evolving machine intelligence and its influence on risk landscapes & analyses

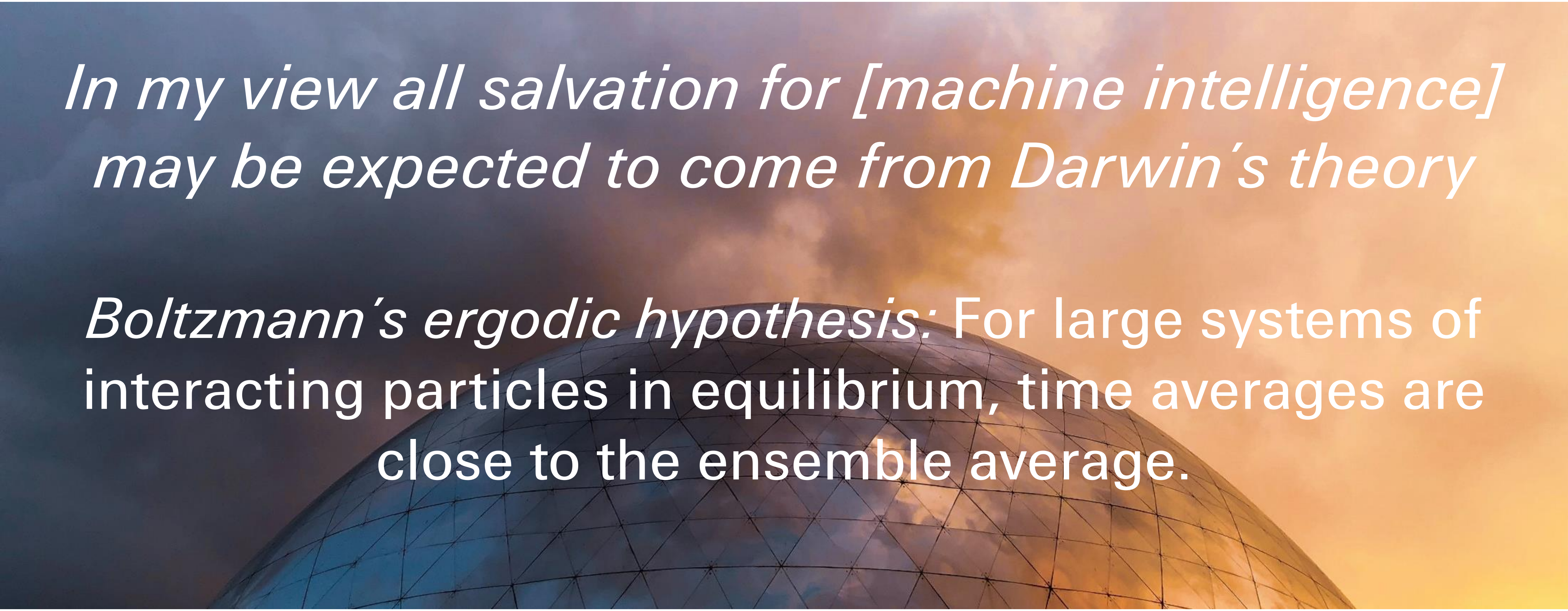
Dr. Jeffrey Bohn, Head, Swiss Re Institute

CDAR Symposium, 19th October 2018

Digital Society = People + Data + Networks + Machine Intelligence
What risks emerge?



Inspiration from Ludwig Boltzmann (Austrian physicist & philosopher who developed statistical mechanics; coined the term *ergodic*)



In my view all salvation for [machine intelligence] may be expected to come from Darwin´s theory

Boltzmann´s ergodic hypothesis: For large systems of interacting particles in equilibrium, time averages are close to the ensemble average.

Outline for presentation

- Digital society
- Changing risk landscapes
- Reinsurance
- Evolution of risk modeling
- Machine intelligence
- Machine intelligence's fragility
- Matching the algorithm/system/process/data to risk-related use cases
- Final remarks

Digital society

The Digital Society is the result of technology abundance

Semiconductors enabled cheap and abundant **computation**

The internet enabled cheap and abundant **connectivity**

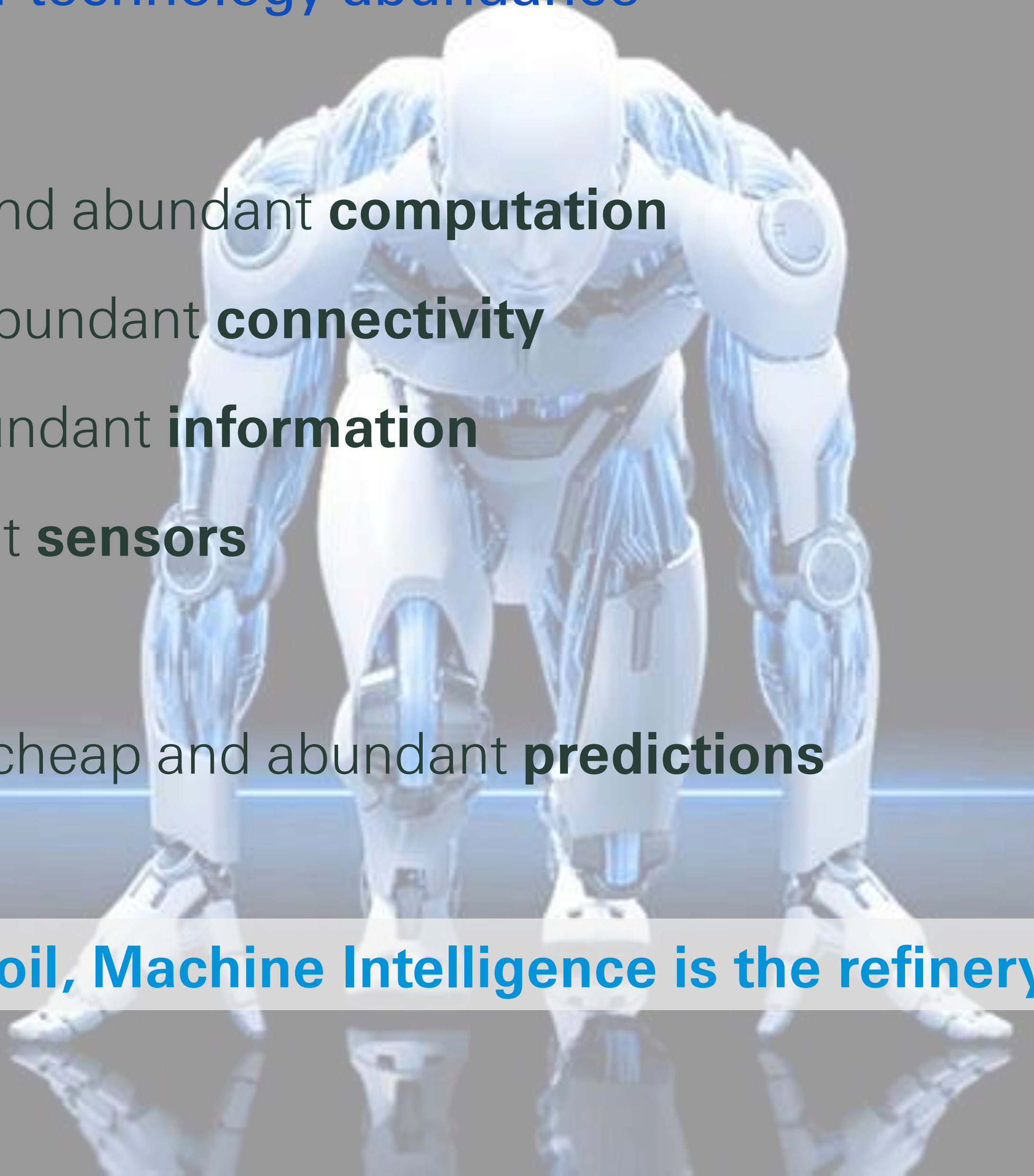
Big data is creating cheap and abundant **information**

IoT is enabling cheap and abundant **sensors**

...

Machine intelligence is enabling cheap and abundant **predictions**

In digital societies, if data is the oil, Machine Intelligence is the refinery



Rise of the machines

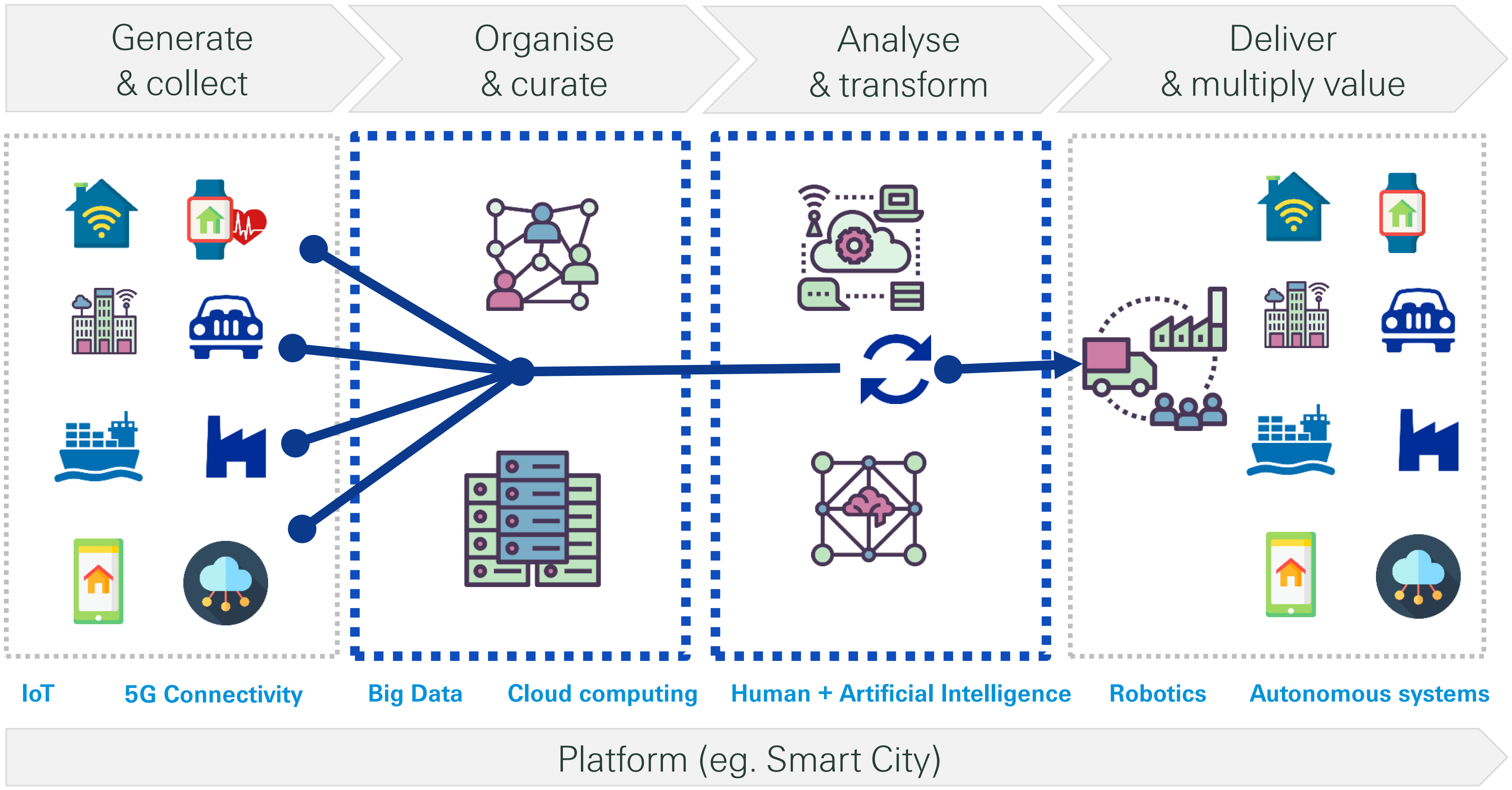
	Early 1980's	Early 1990's	Early 2000's	Late 2000's	Now – Future
	Overcrowding	AI Winter	Machine Learning	Deep Learning	Meta-learning
Description	Primitive AI that required many sets of rules to perform basic functions	No substantial innovation; caused by weak computer hardware and lack of data	Exploration into learning systems that are fed large amounts of data	Innovating machine learning systems to work unsupervised and faster	Current/ongoing development of learning systems aim to mimic humans
Models	Rules based Pattern matching Linear models	Clustering	Regressions Small ANNs Non-linear models	Sigmoid/Back prop. Deep RNNs Boosting	Boosting, Boltzmann machines, GANs Projective simulation Quantum-inspired Evolutionary
Objectives	Perform basic tasks normally done by humans (CPU Chess Players)	Stalled while awaiting more advanced computer systems and data	Create self-learning systems that feed on data (Stock trading algorithms)	Self-learning systems that improve unsupervised (Self-driving Cars)	Systems that accurately mimic human behavior (Chatbots)

Note: AI: Artificial intelligence; ANN: Artificial neural network; RNN: Recurrent neural network; GAN: Generative adversarial networks

Rise of the networks

	Early 1980's	Early 1990's	Early 2000's	Late 2000's	Now – Future
	Government	Worldwide Web & Client-servers	Ubiquitous Internet/ IoT	Cloud	Edge & Fog
Description	Early networks were funded by a few governments such as the U.S.	Migration to dial-up internet and proliferation of client-server networks	Internet penetration in developed world exceeds 50%	Change economics of networked computing in terms of fixed investment	Further change of networked computing to further reduce investment, add IoT
Examples	Arpanet	Top-level domains e.g., .com take hold	Appliances shipped with IP addresses	AWS, GCP, and Azure	Make appliances "smart" e.g., Tesla
Objectives	Connect researchers & military	Connect universities & companies	Connect companies, governments, universities & consumers	Create more elastic networks that extends networked computing	Move machine intelligence out to network edges to increase speed & create new "intelligence"

Data is the new oil: Crude product multiplies in value after refinement



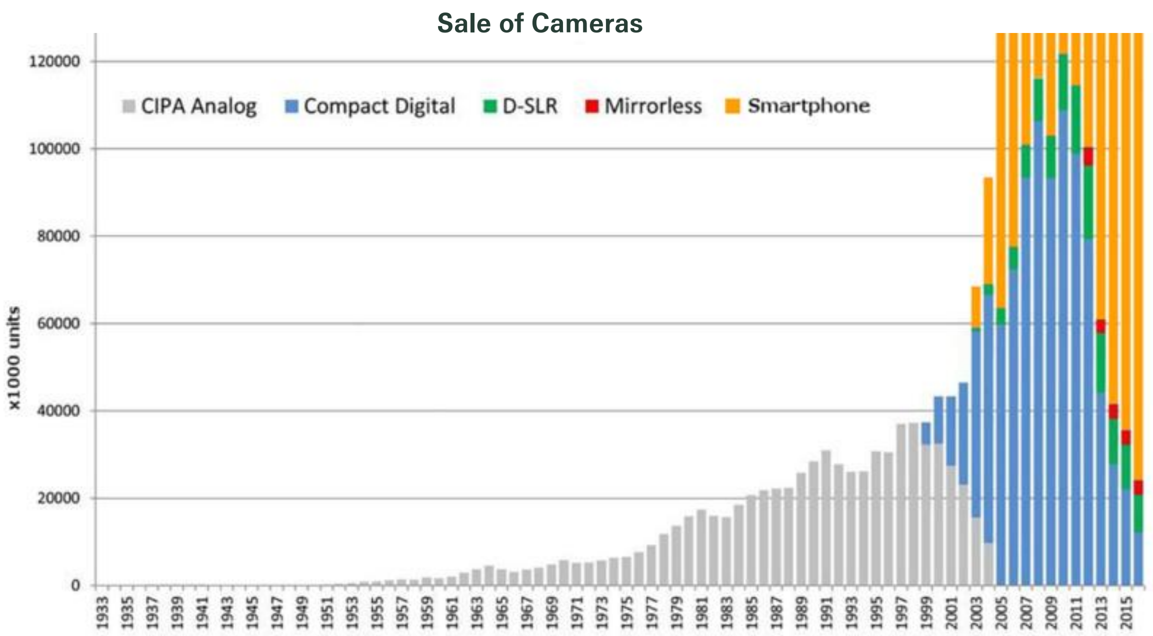
Changing risk landscapes

Digital society and new risk landscapes

- Insufficient investment in data collection & curation systems
- Digital "exhaust" and "hygiene"
- New networks overlayed on legacy networks create gaps and vulnerabilities
- Differing value/regulatory systems underlying society's digitization create disconnects
- Over-dependence on automated/machine-intelligence-enabled systems leads to blind spots and unnecessary risk
- Vulnerabilities arise from IoT-enabled devices with outdated security features
- Mismatching digital tools with a given use case leads to poor implementations & vulnerabilities

Cyber risk and ***algorithmic malpractice*** increasingly becoming most important risk categories

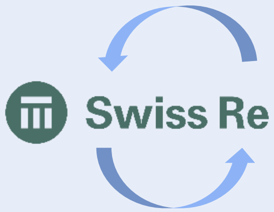
Lessons from technological change in a different industry



Source: Gartner Inc. (2015)

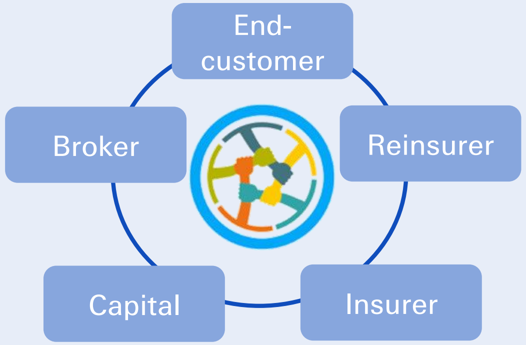
Changes arising from distributed ledger technology (DLT)

Internal improvements



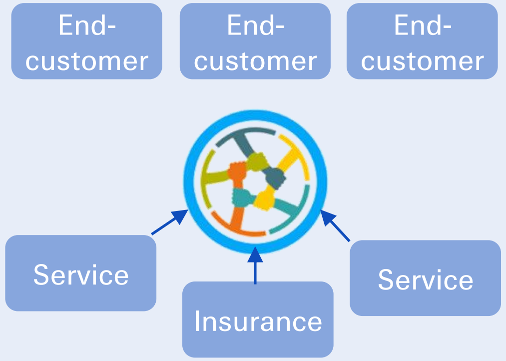
Enhancing **internal business process efficiencies** by hosting shared services on internal DLT platforms

Enhancement of insurance value-chain



Organizing members of **existing insurance value-chain** into a DLT cooperative for driving data standardization, aggregated data access and shared processes

End-customer driven initiatives



End-customers in **non-insurance verticals** organizing DLT cooperatives for aggregation of shared services including insurance

New value-chains

New players creating a **DLT-based alternative value-chain** or market for distribution of re/insurance services

Incremental efficiency gains ⟷ *Business model disruption*

What is Cyber Risk?

Accidental breaches of security / Human error.

Unauthorized and deliberate breach of computer security to access systems.

IT operational issues resulting from poor capacity planning, integration issues, system integrity etc.

How severe is this risk?

1.9B records stolen in first 3 quarters of 2017 compared to 600M in entire 2016

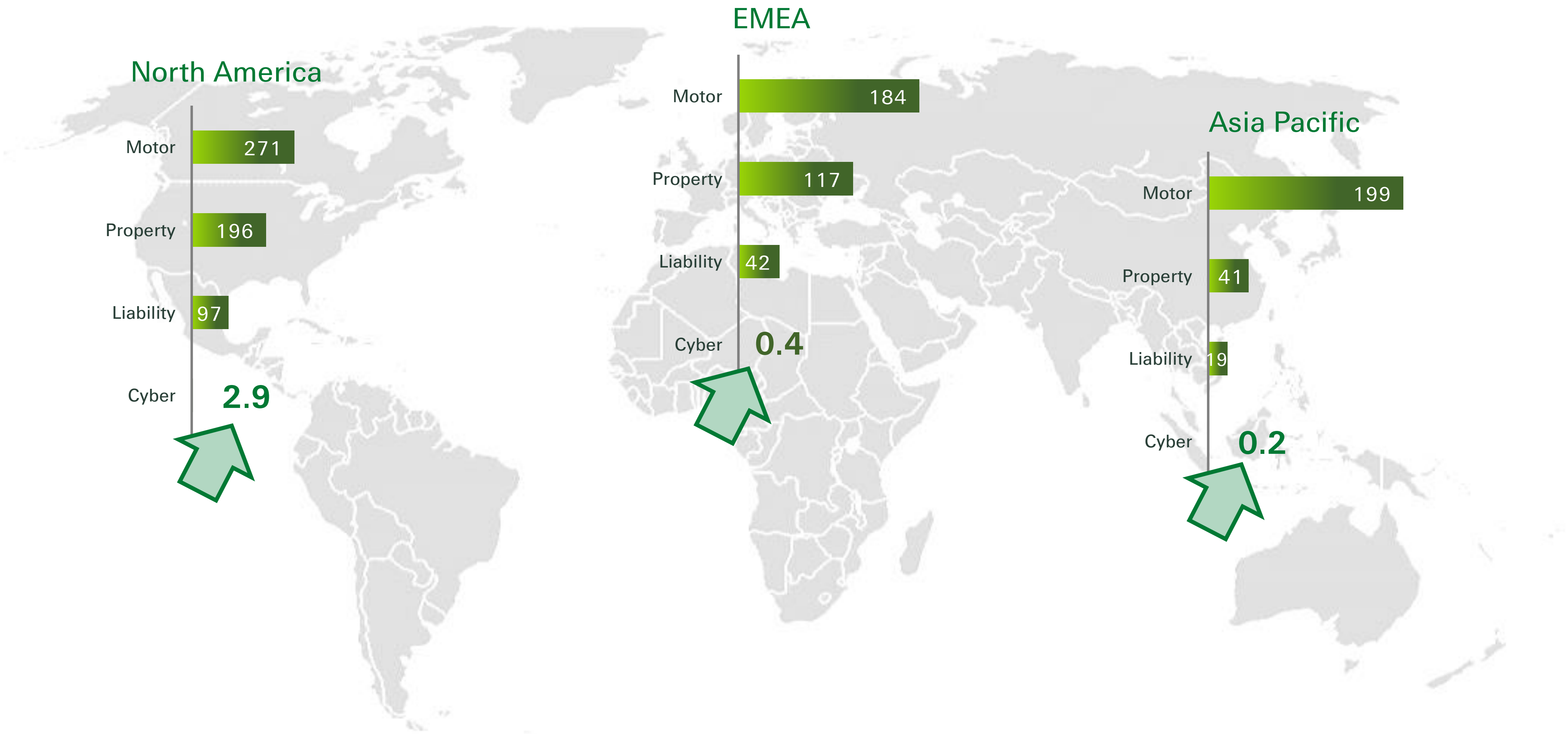
500M forms of malware (400k created everyday)

1M identities breached per breach

Probability of breach in a year is close to 100%

The global cyber insurance market is still small...

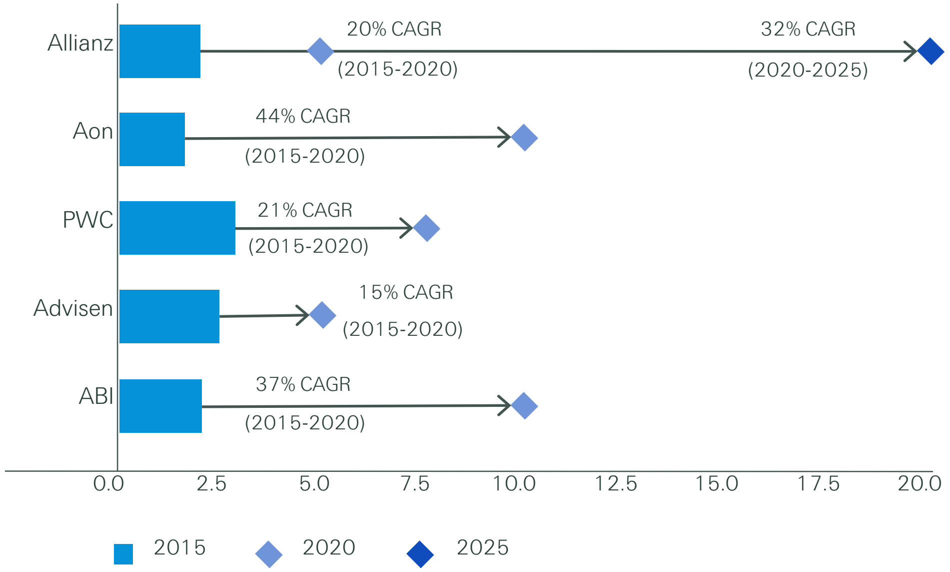
Insurance premium in 2017 per LoB, in USD bn



Source: Swiss Re Institute

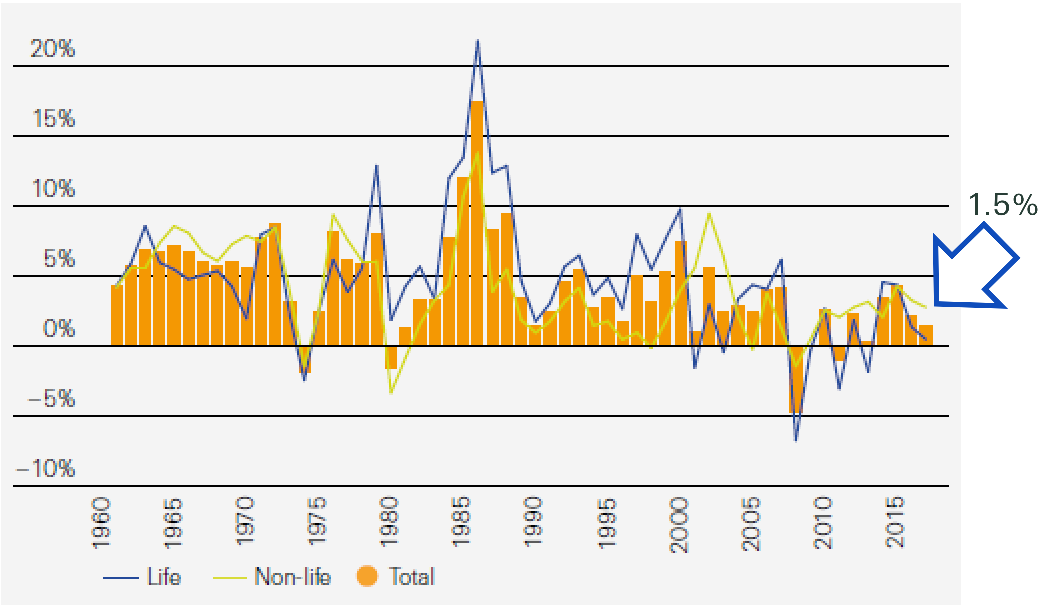
...but expected to grow rapidly

Expected global cyber insurance premium volume (2015 – 2025), in USD bn



Source: Swiss Re Sigma No 1/2017, Cyber: getting to grips with a complex risk

Global insurance premium growth (1960 – 2017)



Source: Swiss Re Sigma No 3/2018, World insurance in 2017

Cyber-insurance market is expected to grow more rapidly than other lines of business. Currently, market focus is on the SME segment. However, first cyber insurance solutions for individuals are also entering the market.

Swiss Re Institute

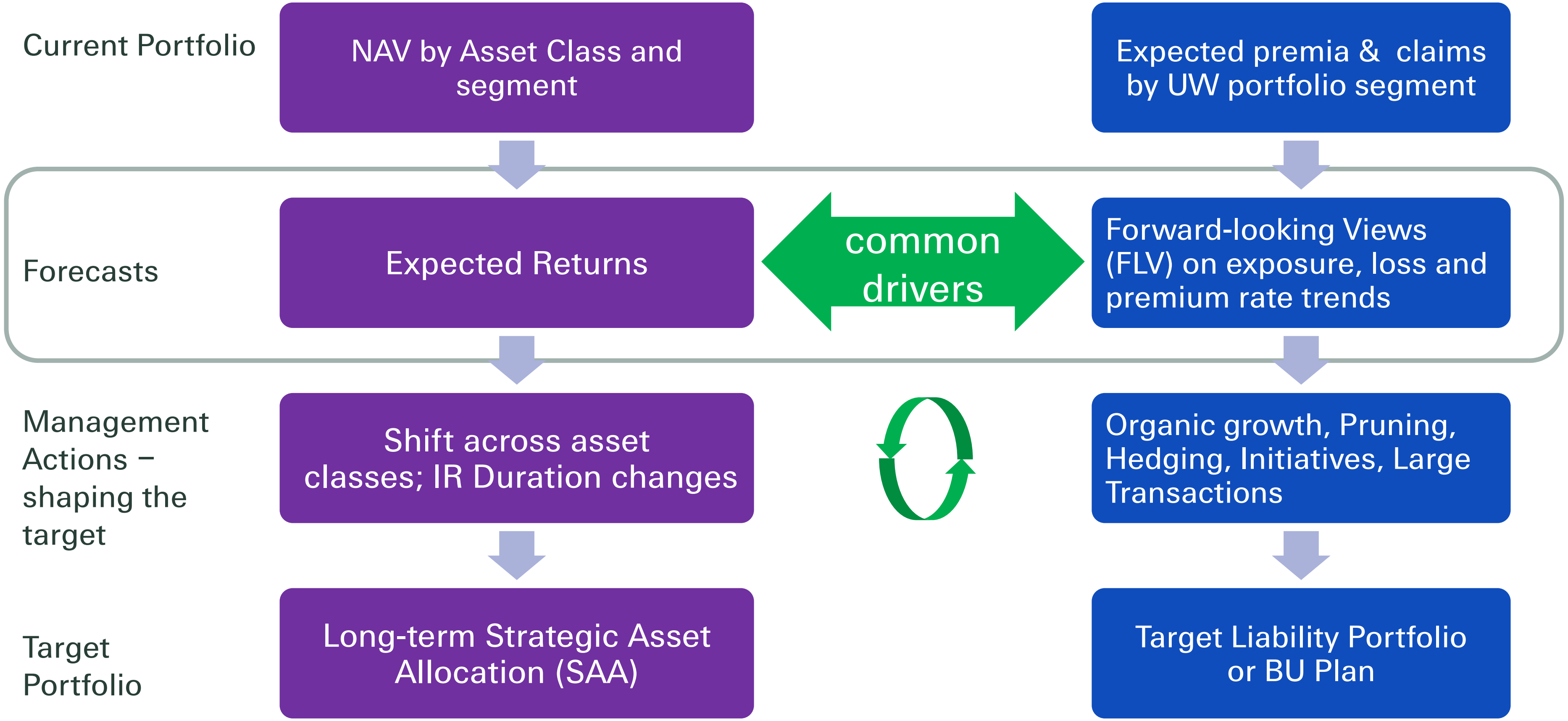
Reinsurance

Reinsurance is a catalyst for economic growth.

	Activity	Benefits
Risk transfer function	Diversify risks on a global basis	Make insurance more broadly available and less expensive
Capital market function	Invest premium income according to expected pay-out	Provide long-term capital to the economy on a continuous basis
Information function and knowledge	Price risks	Set incentives for risk adequate behavior

Reinsurers absorb shocks, support risk prevention and provide capital for the real economy

Assets and Underwriting steering have similar basic building blocks and should be done jointly



Forward-looking modeling is key to improving a reinsurer's performance

- Capital allocation: Choose/avoid outperforming/underperforming insurance portfolio segments (IPS)
- Risk selection: Select risks within each IPS
- Strategic asset allocation: Choose/avoid outperforming/underperforming asset classes

More examples where machine intelligence changes insurance

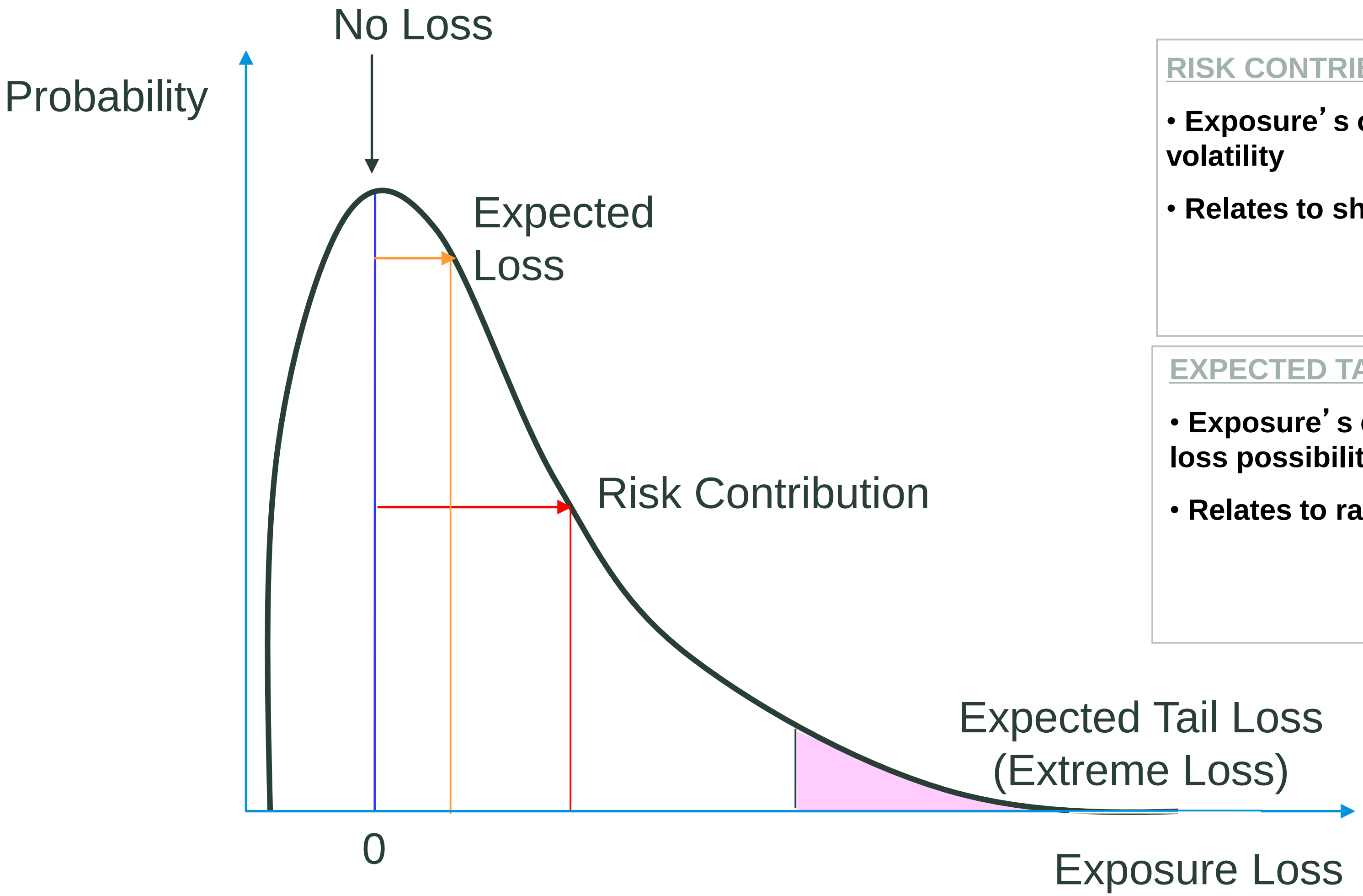
- Forward-looking modeling of risk pools
- Incorporating unstructured data into business and capital steering
- Tracking natural catastrophe damage in real time
- Assessing damage
- Automated underwriting
- Improving customer targeting
- Parametric insurance contract implementation
- Intelligent automation & robotic process automation (RPA) for underwriting and claims processing
- Chatbots for customer support
- Natural language processing applied to contract review

Evolution of risk modeling

“Risk” definitions matter

- Symmetric return/loss distributions with parameterizable distributions based on existing data– “known unknowns”
  - Volatility measures may be sufficient
  - “Beta” risk should be the focus i.e., allocation more important than individual exposure selection
- Asymmetric return/loss distributions with changing distribution parameters with sparse data– “partially known unknowns”
  - Focus shifts to tail risk/expected tail loss
  - Diversification opportunities continue even as portfolio becomes quite large
  - Active management may be better compensated given huge savings to avoiding tail events
- Emerging risk means data availability may be so sparse as to make any parameter estimation infeasible– “unknown unknowns”
  - Ambiguity creates risk that cannot be managed using traditional approaches
  - “Structural” model based on subject matter experts can provide some guidance, which cast some light on unknown

RISK CONTRIBUTION (RC) & EXPECTED-TAIL-LOSS CONTRIBUTION (TLC)



RISK CONTRIBUTION

- **Exposure's contribution to portfolio's volatility**
- **Relates to shorter-term risk**

EXPECTED TAIL-LOSSCONTRIBUTION

- **Exposure's contribution to extreme loss possibility (i.e. the "tail")**
- **Relates to rare, but severe losses**

Simple
* Volatility Multiple
* Analytical solutions

Simulated
* Flexible
* ETL contribution

Stratified
* Dynamic
* Individual exposure differentiation

Swiss Re Institute

Black swans, gray rhinos, and perfect storms

- Defining extreme-downside, scenario categories:
  - Black swans: Unknowable given current information set and virtually impossible to predict
  - Gray rhinos: Highly probable and straightforwardly predictable given current information set, but neglected
  - Perfect storms: Low probability and not straightforwardly predictable given the outcome results from interaction of infrequent events
- Scenario-based analyses vs. forecasts
- Deeper analyses of underlying assumptions, relationships, and data
- More focus on tools/processes to manage multiple sets of scenarios and analyses across time
- Renewed efforts to enforce *preproducibility, reproducibility,* and *out-of-sample testing*
- Process management systems with robust audit logs are more important than ever

Challenges specific to financial market and insurance risk modeling

- Sources of non-stationarity/ non-ergodicity
  - Structural changes in the real economy
    1. Digital ecosystems
    2. Increased concentration within sectors
    3. Information & communications technology
  - Structural changes in the financial economy
    1. Near-zero interest rates
    2. Inflation
    3. Globalization of capital markets

- Ergodic systems
  - Closed
  - Low dimensional
  - Not evolving, but can be cyclical
- Non-ergodic systems
  - Open
  - Generate new information all the time
  - Learn and adapt over time (e.g., Darwinian evolution)
  - Past data mostly not useful
  - Purely inductive models are mostly ineffective

It is far better to have absolutely no idea of where one is– and to know– than to believe confidently that one is where one is not.
-- Jean-Dominque Cassini, astronomer, 1770

Marrying qualitative and quantitative data

- Roughly 90% of available data are qualitative and unstructured e.g., articles, blogs, e-mail, regulatory filings, slide presentations, social media, etc.
- Quantitative data may not reflect all forward-looking risks (e.g., Environment, Social, Governance-- ESG)
- Transforming qualitative data into indicators and combining in some way (e.g., shading, weighted combinations, etc.) with quantitative data may be a path to improving existing models

Misusing Occam's razor

William of Ockham's actual quote (most likely): *Numquam ponenda est pluralitas sine necessitate*

[Plurality must never be posited without necessity]

Simplicity in concept may belie complexity in reality– e.g., biological evolution, construct an optimal portfolio, optimally allocate capital to insurance portfolio segments, etc.

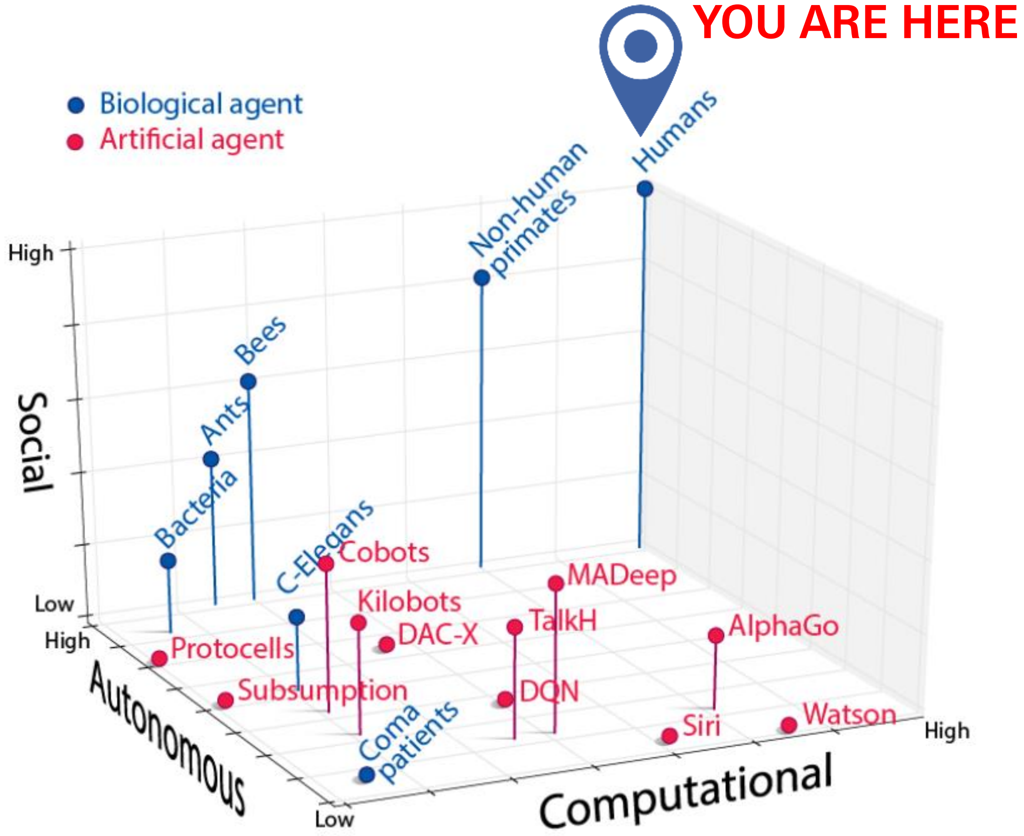
Machine intelligence

How we define "artificial" and "intelligence" will influence research and development in machine intelligence

- Artificial: Human-made, contrived, not natural, not real
- Intelligence: Learn and **apply** knowledge or skills, solve problems, ability to reason/ plan/ **adapt**/ respond/ think abstractly
- Artificial intelligence: Ability to simulate human intelligence– is this all?

Are these definitions adequate? How we define machine intelligence materially influences the tools we create

"Viewed narrowly, there seem to be almost as many definitions of intelligence as there were experts asked to define it." Sternberg



The future will be Human Intelligence augmented by some kind of Machine Intelligence

Machine intelligence taxonomy

- **Artificial intelligence:** Mimic human intelligence and possibly go beyond
- **Artificial general intelligence:** Possibly self-aware intelligence
- **Machine learning:** Data-dependent calibration
- **Deep learning:** Model-free, data-dependent calibration
- **Meta-learning:** Learning how to learn
- **Cognitive computing:** Simulate human-brain processes
- **Augmented intelligence:** Human assistants
- **Expert systems:** Advice systems using knowledge databases
- **Robotic process automation:** Roboticized systems that replicate repetitive processes
- **Intelligent automation**: Hybridized processes that use automation to better leverage human productivity

Six major categories for applying machine intelligence in business applications

- Machine Intelligence for Business Applications
  - Machine Learning Platform (ML)
    - Supervised Learning: Regression, Classification, Dimension Reduction..
    - Unsupervised Learning: Association rule learning, clustering...
    - Reinforcement Learning
  - Deep Learning Platform (DL)
    - Neural Network/ Convolutional Neural Network
  - Natural Language Processing (NLP)
    - Speech Recognition/Question Answering
    - Content Creation
    - Natural Language Generation
  - Internet of Things (IoT)
    - Digital Twin/AI Modeling
    - Peer-to-Peer Networks/ Edge Computing
  - Recognition Technology
    - Virtual Agents
    - Emotion Recognition
    - Image Recognition
  - Decision Management
    - Decision Support System
    - Infotainment System

Boltzmann machines

- **Boltzmann machine**: Stochastic recurrent neural network that can be seen as the stochastic, generative counterpart of Hopfield nets. Maximizes the log likelihood of given patterns exhibiting Hebbian engrams. Threshold is binary. Finds hidden relationship layers.
- **Hopfield networks**: Form of recurrent artificial neural network that is content-addressable with binary threshold nodes.
- **Hebbian engram**: Two cells (or systems of cells) that are repeatedly active at the same time tend to become "associated."
- **Dynamic Boltzmann machines (DyBM)**: Maximize log likelihood of time series with property of spike-timing dependent plasticity (STDP) [amount of synaptic strength change depends on when two neurons fired.] Threshold is binary.
- **Gaussian Dynamic Boltzmann machines**: Extend DyBM to deal with real values capturing long-term dependencies (similar to VAR, which is a special case). Can add an RNN layer to induce hidden, high-dimensional non-linear relationship features.

Ackley, David H., Geoffrey E. Hinton, and Terrence J. Sejnowski, 1985, "A learning algorithm for Boltzmann machines," *Cognitive Science*, 9(1), 147-169.

Dasgupta, Sakyashingha and Takayuki Osogami, 2017, "Nonlinear dynamic Boltzmann machines for time-series prediction," *Proceedings of the thirty-first AAAAI conference on artificial intelligence.*

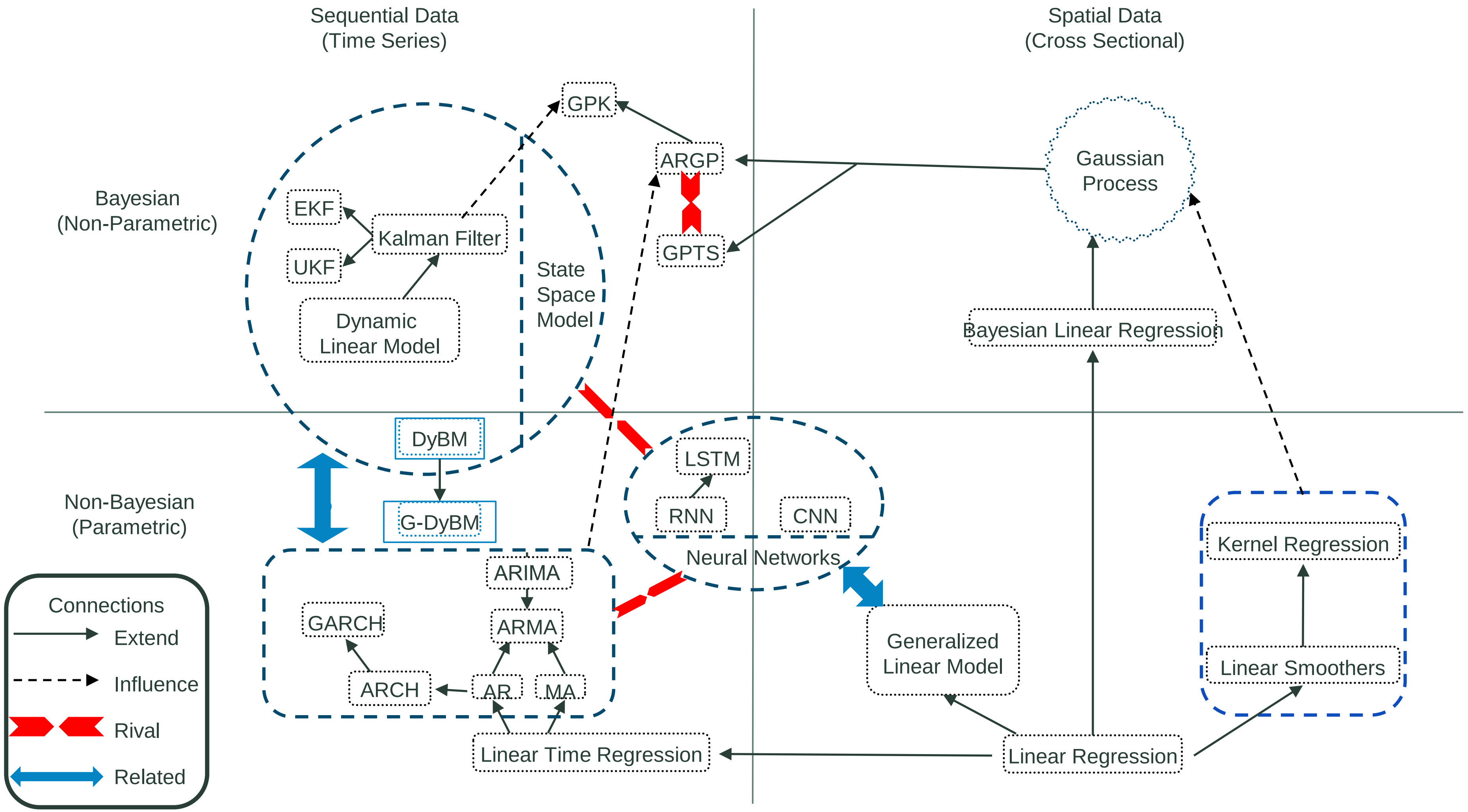
Hebb, Donald O., 1949, *The organization of behavior*, New York: Wiley & Sons.

Hopfield, John J., 1982, "Neural networks and physical systems with emergent collective computational properties, " *Proceedings of the National Academies of Sciences*, 79, 2554-2558.

Marks, T., and J. movellan, 2001, "Diffusion networks , products of experts, and factor anlaysis," *Proceedings of the Third International conference on Independent compennet analysis and blind source separation.*

Osogami, Takayuki and M. Otsuka, 2015, "Learning dynamic Boltzmann machines with spike-timing dependent plasticity," Techicial report RT0967, IBM Research.

“Big Picture” for machine intelligence– four schools of thought



Swiss Re Institute

Machine intelligence´s fragility

What is Meant by “Fragility of Machine Intelligence”?

Fragility of Machine Intelligence: *When a machine intelligence-based solution fails to meet accuracy or performance criteria expected of the model´s application, the model outcome may adversely affect dependent systems. The unexpected output of the machine intelligence is a result of the model´s fragility, or inability to process data in relation to the expected performance criteria.*

Sources of fragility
- Overlaying new systems on existing systems that are not likely to be seamlessly interoperable
- Mismatching systems/algorithms with particular use cases
- Interaction of humans and machines
- Poorly designed system architectures
- Poorly designed algorithms buried in a system architecture– *algorithmic malpractice*
- Overall exposure to cyber attacks

Major vulnerabilities exist across many modern machine intelligence methods

Vulnerabilities	Problems	Typical Algorithms
“Blackbox”	Difficulty in debugging, error tracking	Neural Networks
Computational Cost	1. Lack of computing power 2. High computational complexity (e.g. $O(n^3)$)	Neural Networks, RNN, Search Algorithms
Data	1. Biased/volatile/fat-tailed/arbitrary data 2. Requirement of well-labeled data 3. Insufficient data	Neural Networks, Clustering, Classification Algorithms (RNN, Random Forest, SVM), Assembly Algorithms
Security	1. Cyber attacks 2. Privacy gets threated	Neural Networks, IoT
Algorithm Design	1. Impractical vocabulary bank in NLP 2. Complex hyperparameter selection 3. Not scalable 4. Cannot solve non-linear problems 5. Not interpretable 6. Fail to capture error in learning process	Neural Networks, NLP, RNN, Gaussian Regression, Kalman Filter, Sorting Algorithms, Blockchain Algorithms, Robotics Algorithms
Information Loss	1. Memory loss in training process 2. Feature loss in dimensionality reduction	Neural Networks, PCA, Clustering Algorithms
Regulation	Lack of standardized regulations	IoT

Impacts of Machine Intelligence Fragility

- Regulatory
- Impact to Peripheral Machine Intelligence Modules
- Impact to Transaction Flows
- Financial Impact to Business (resulting from Impact to Transaction Flows)
- Financial Impact to Business (as a "Residual Result")
- Property Damage
- Human Injury or Loss of Life

Matching algorithm/system/process/ data to risk-related use cases

Matching techniques to objectives

- Objectives/use cases:
  - Identify trends (first moment estimation e.g., expected return)
  - Identify short-term risk (second moment estimation e.g., volatility estimation)
  - Identify long-term "risk" (higher moments, e.g., expected tail loss or expected shortfall from a skewed, non-normal [likely non-ergodic] distribution)

- OLS/Factor modeling (Explicit & implicit)
- Better regression techniques: Random forests, lasso regression
- ARIMA/GARCH/Vector Auto Regression (VAR)
- Recursive Neural Networks (RNNs)
- Boltzmann machines
  - Restricted Boltzmann machines
  - Dynamic Boltzmann machines
  - Gaussian dynamic Boltzmann machines
- Gradient boosting
  - XGBoost
  - LightGBM

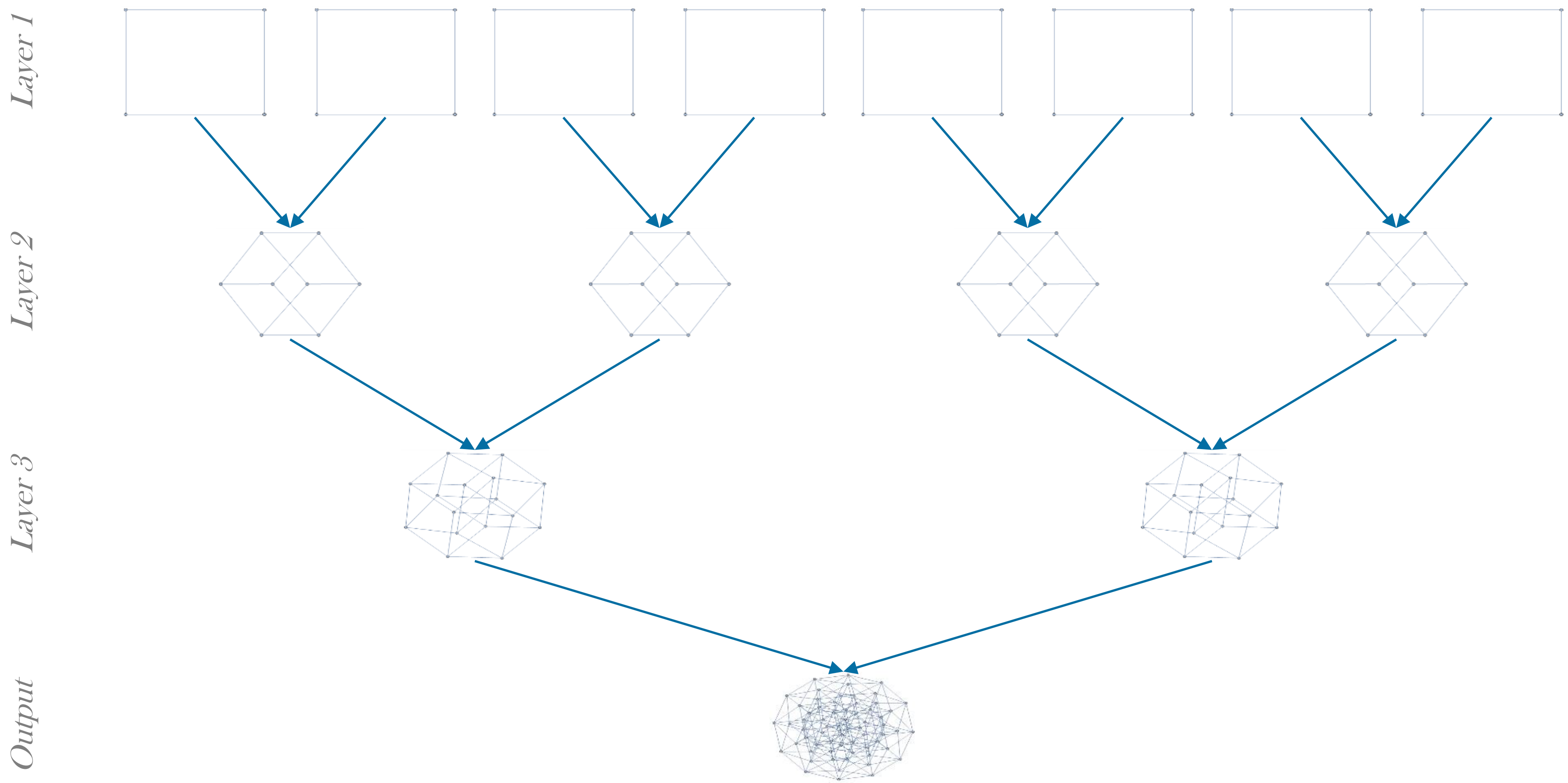
Modeling challenges

- Variable data availability & quality
- Changing underlying data generation processes
- Moving up the ladder of causation (Pearl, 2018, p. 28)
  - Level 1: Association (Seeing, observing)
  - Level 2: Intervention (Doing, intervening)
  - Level 3: Counterfactuals (Imagining, retrospection, understanding)
- Macroeconomic interaction
- Insurance market interaction (Game theory)
- Behavioral economics (Changing product design, marketing, distribution, strategy can change opportunities within a given segment)
- Digitizing society impact

How to persuade an individual to make a decision (Inspired by Koomey (2001) figure 19.1 p. 88)

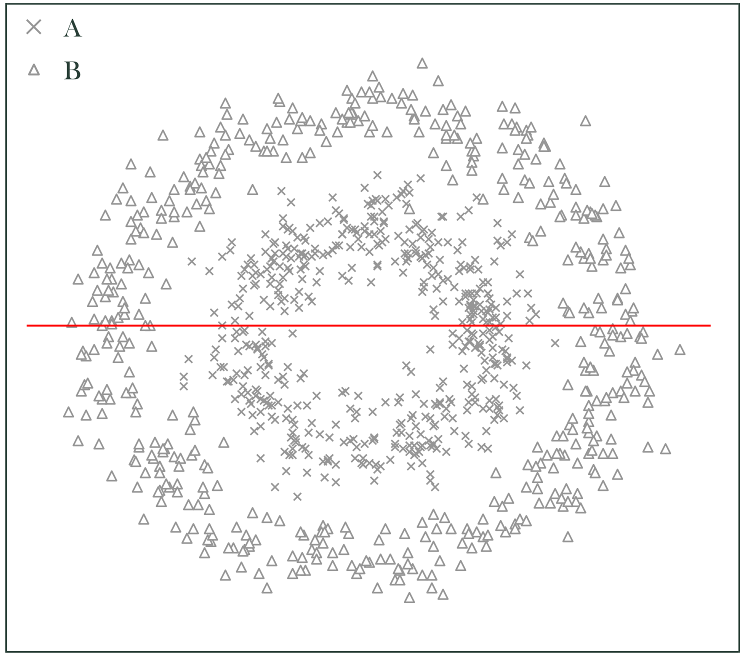
	Agree on objectives and criteria for acceptance	Disagree on objectives and/or criteria for acceptance
Agree on data	Computational decision	Negotiate
Disagree on data	Experiment	Paralysis or chaos

Ensemble modeling is the process by which multiple weaker systems are combined to create more complex systems



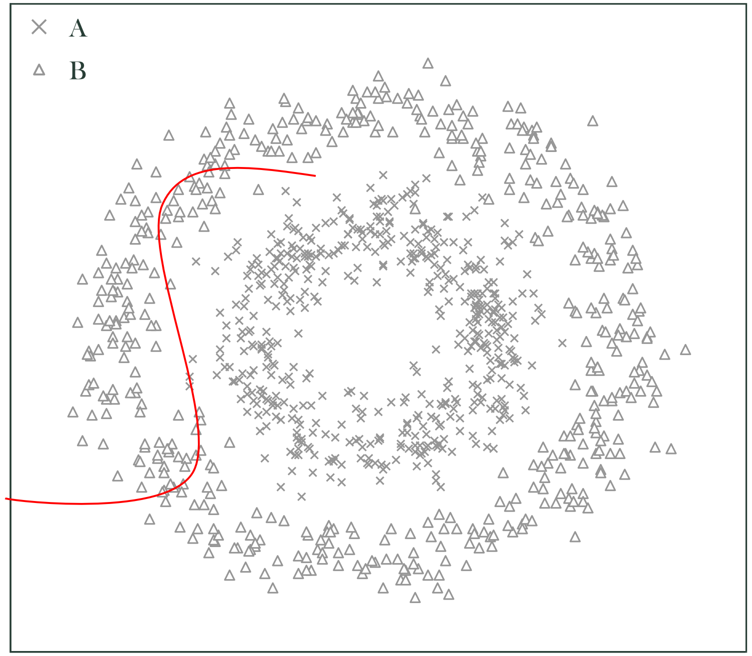
Example: highly complex datasets, such as "donut problem" can't be learned without ensemble modeling (Source: XLP)

Linear Model



When splitting the donut into binary classifications, a linear model would simply cut the data in the middle. This leads to a very high error rate of 50%, and is no better than random guessing

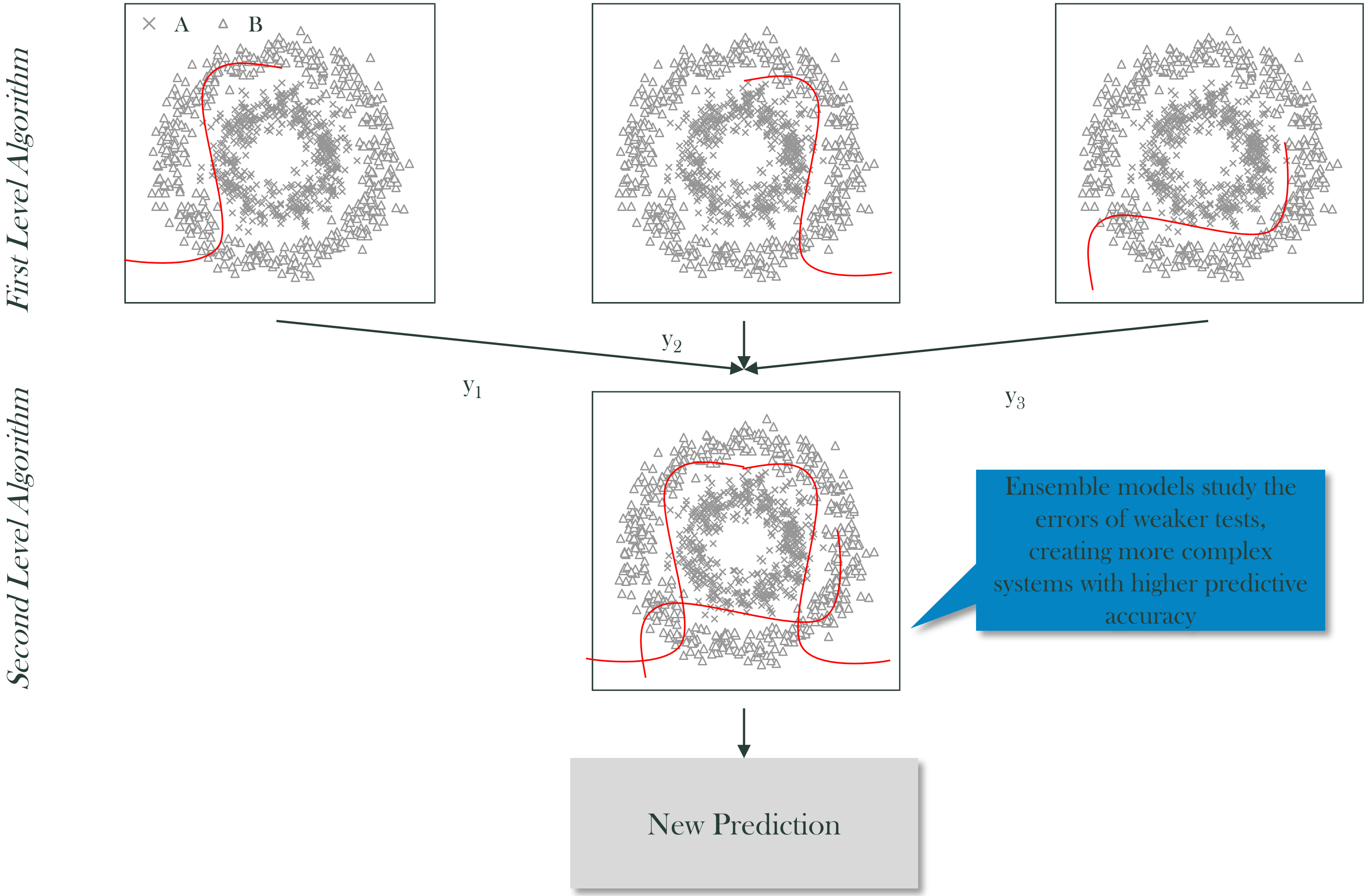
Non-Linear Model



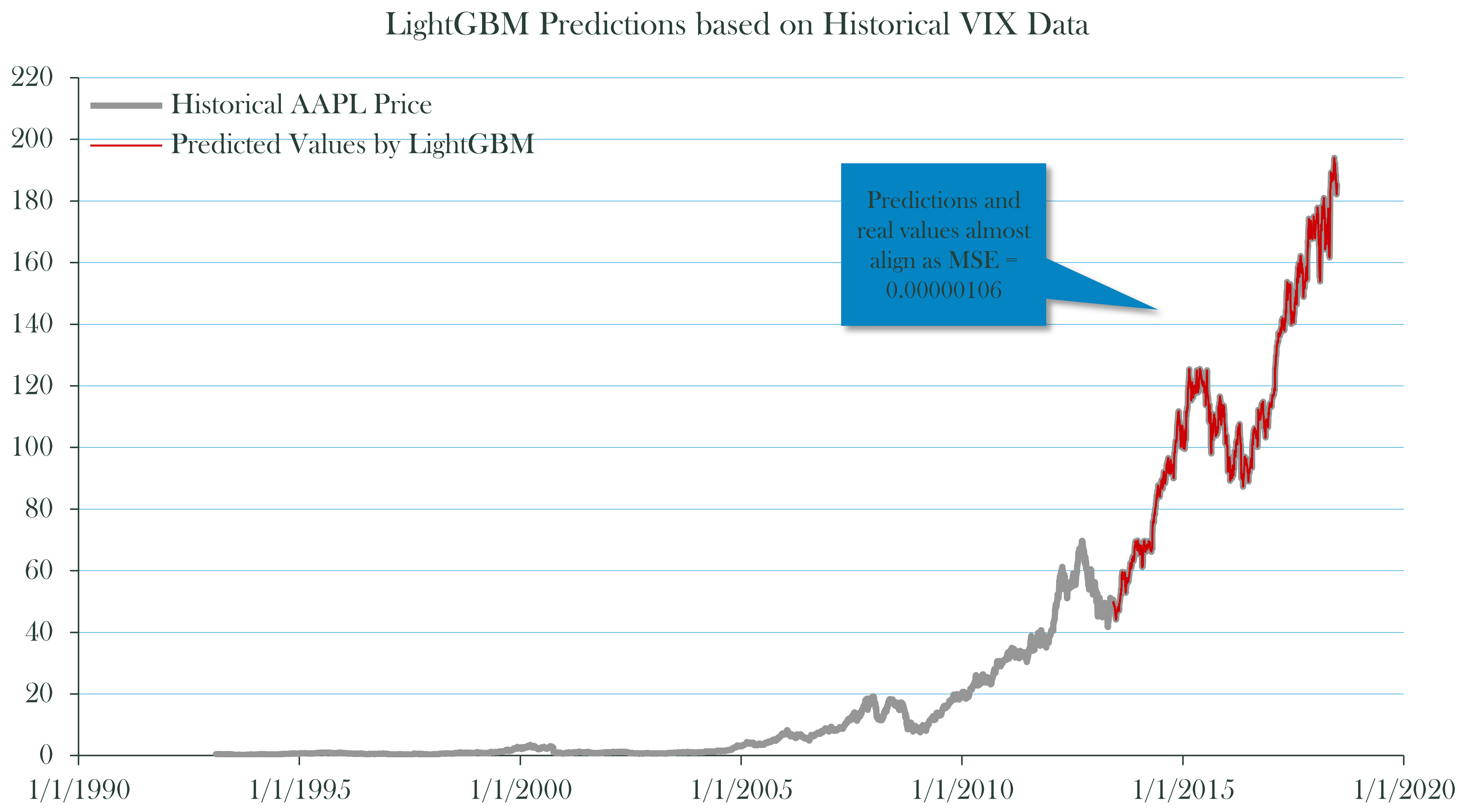
A single, non-linear model would draw a curve that most minimizes the error of binary classification. The error rate using a non-stacked model would still be extremely high

Swiss Re Institute

Stacking to generate more complex system, ensemble models learn from data with greater accuracy (Source: XLP)

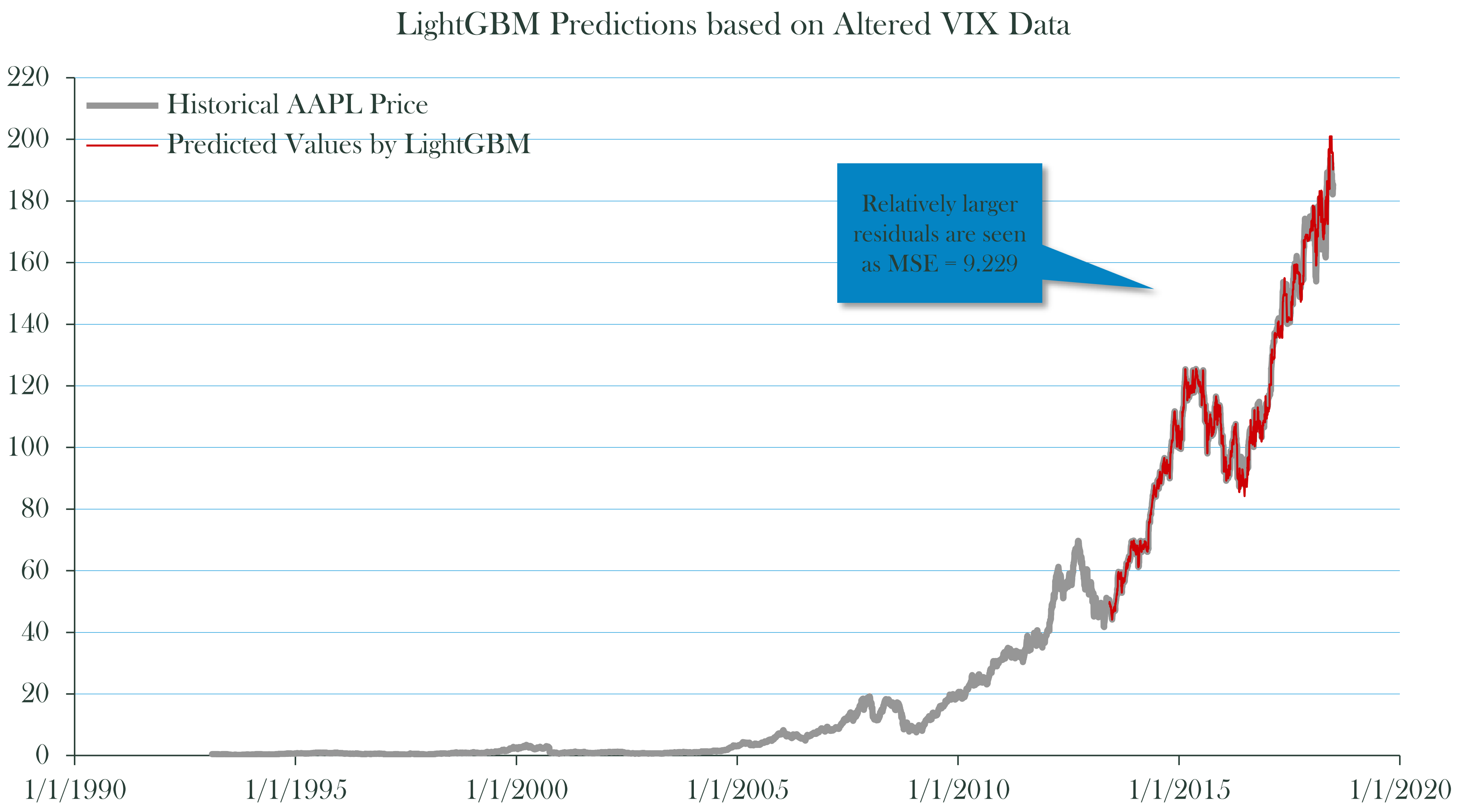


Gradient Boosting algorithm has a high prediction accuracy using the original VIX data



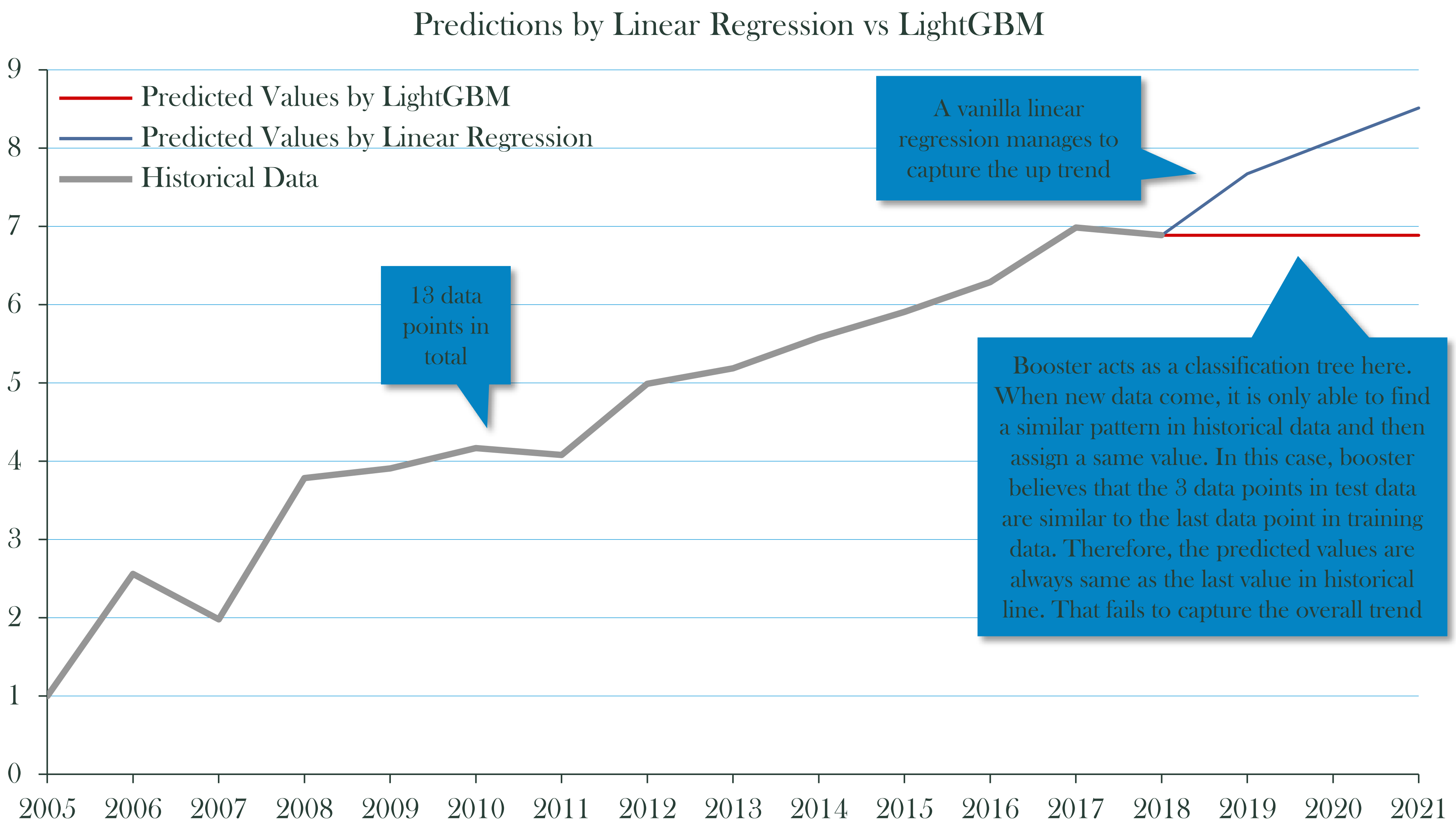
Swiss Re Institute

However, a minor jump in VIX can alter the accuracy of the model predictions significantly

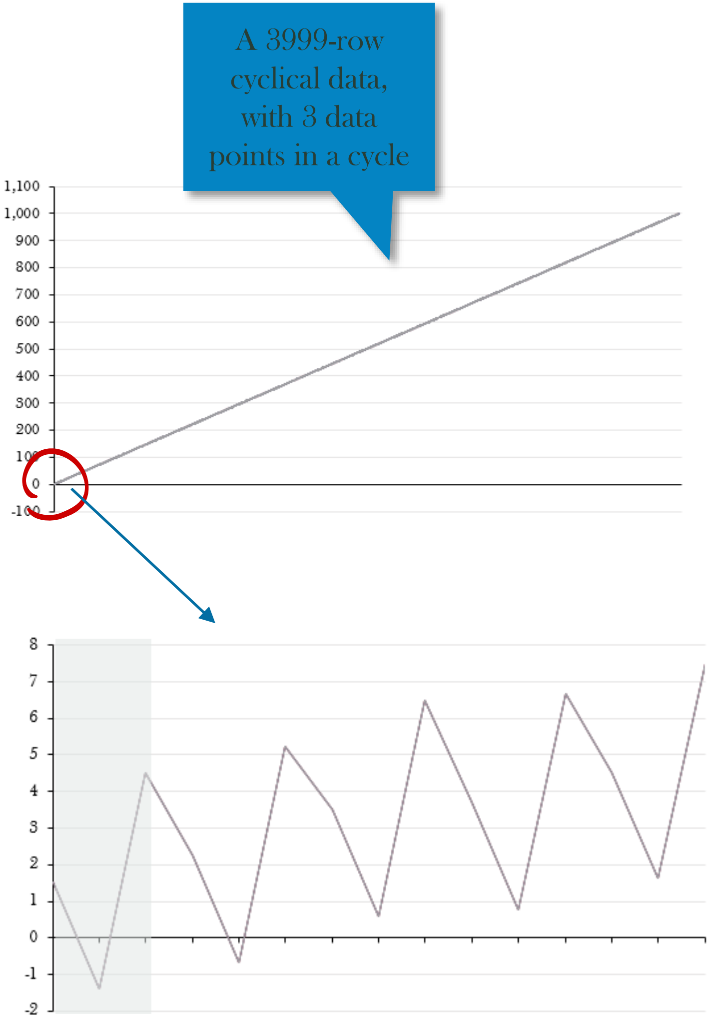


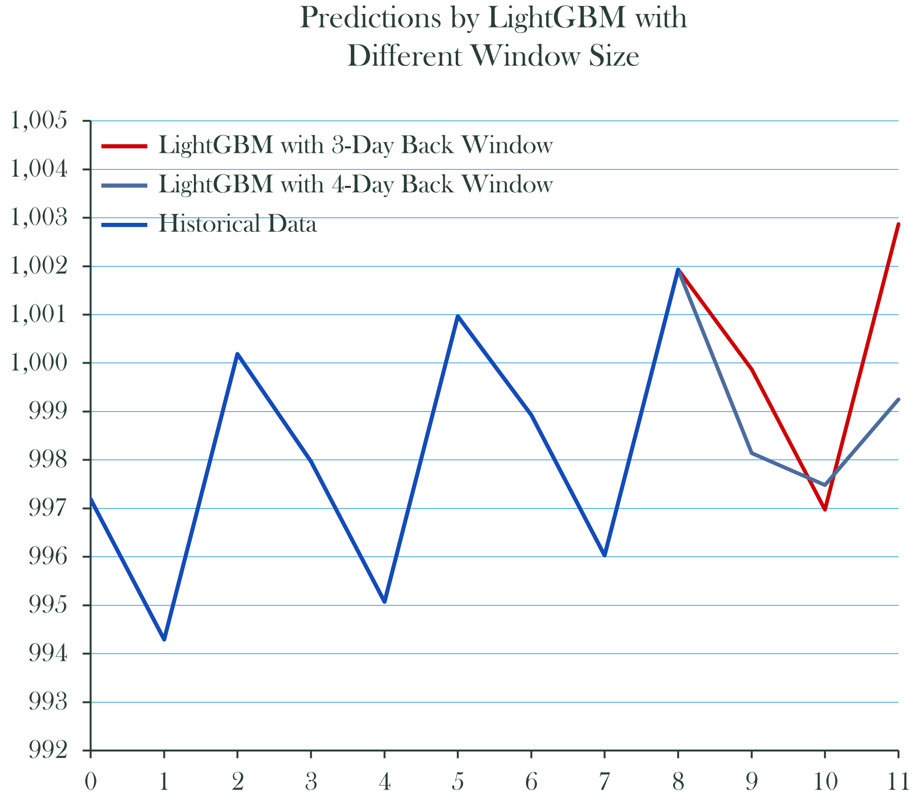
Swiss Re Institute

Gradient Boosting fails when data is limited, as the algorithm merely tries to find the similar pattern in the past data



Gradient Boosting fails when a proper window size is not well-selected for a cyclical time series





Final remarks

Defining a research agenda at the intersection of machine intelligence and digitizing societies

- Focus more on…
  - Addressing the challenge of data curation
  - Communicating actionable insight
  - Testing algorithms on real, useful data at scale
  - Discussing how machine intelligence should integrate into a digitizing society (define algorithmic malpractice; shape regulatory environment regarding data & machine intelligence)
- Focus less on…
  - Developing more sophisticated algorithms without a specific applied use case
  - Testing algorithms on toy datasets

Challenges

- Collecting & curating suitable & sufficient data
- Matching the type of machine intelligence with an objective or use case
- Making algorithms interpretable and diagnosable
- Defining what constitutes *algorithmic malpractice* in the machine intelligence arena
- Dealing with data sparsity and non-stationary data-generating processes
- Architecting and complying with data privacy regulations
- Addressing conflicts arising from human notions of law, fairness, and justice and machine-intelligence capabilities that can circumvent protections
- Addressing system fragility as interconnected and complex networks are infused with machine intelligence
- Addressing growing cyber-risk in digital ecosystems infused with machine intelligence

Inspiration from biology (Theodosius Dobzhansky)

Nothing in a digitizing society makes sense except in the light of the evolution of machine intelligence

Legal notice