Swiss Re Institute

No 5/2020

sigma

Machine intelligence in insurance: insights for end-to-end enterprise transformation

Executive summary

MI offers significant revenue and cost-saving potential.

Machine intelligence (MI) permeates many industries, bringing significant revenue creation and cost savings potential. To date in insurance, MI has yielded returns in areas such as customer analytics and claims processing, based mostly on machine learning (ML) technology. The scope for industry gain is farther reaching. Among others, MI can help insurers more efficiently process text from contracts, documents, email and other online communications tools, and to analyse the massive data sets becoming available from the digital economy and accumulated from Internet of Things (IoT) devices. Insurers can use this information to better design, price and distribute protection covers, and extend their reach into new markets.

Conventional MI methods are already standard in certain areas of the insurance value chain. They could be superseded by more advanced approaches.

Conventional MI approaches such as generalised linear models have become standard tools in insurance for risk assessment and prediction models. Even so, these tools typically only facilitate fragmented, narrowly-focused productivity. Enterprise-scale transformative benefits could be delivered with more investment in data engineering. This focus on data engineering is also necessary to realise the enterprise-wide potential of more advanced ML and artificial intelligence (AI). Early adopters of such approaches are seeing positive results in select areas like faster claims settlement, more targeted cross- and up-selling, and better risk scoring. The foundational technologies necessary to perform MI tasks continue to develop rapidly as algorithms become easier to use and cheaper. As a consequence, we expect that some processes currently profitable with conventional MI may well be supplanted by new ML and AI approaches unlocking new growth trajectories.

Enterprise-scale deployment of MI-enabled systems remains elusive, however.

Beyond such progress, however, enterprise-wide deployment of MI-enabled systems in insurance remains a long way off. In an overview of survey data, we found that less than 10% of firms in all sectors have managed to scale MI pilots for roll out across multiple processes. Primary reasons include data availability and quality. Many ML and AI approaches require large amounts of high-quality data to train algorithms. Even conventional MI is hampered by data quality. Today, many areas of interest in MI are working with data sets that are not complete, clean or timely. This further affirms the importance of data engineering. Without said capabilities, the performance of models/algorithms has proven to be slow and expensive relative to existing human-centric processes. If deployed correctly, the models/algorithms can deliver substantial return on investment but to date they have not been in a ready-state for enterprise-scale rollout. COVID-19 has forced consumers and businesses to become more digitally active. This has accelerated the need to shift to more digitally-oriented business models, further affirming the value of transformative MI.

Issues around data quality and curation are generating interest in new areas of MI potential.

In recent years, the issues around data quality and curation have led to development of new approaches, such as reinforcement learning and ensemble modelling. So-called hybrid models/algorithms based, for example, on a combination of knowledge from physics and ML, and causal-inference approaches, are less sensitive to data quality and compute power inadequacies. These are just two areas of innovative research seeking to address specific performance and model-interpretability issues, yielding solutions that could be a new part of future MI applications in insurance.

It's not just models that matter. A range of issues influence success in roll out of MI-enabled systems at enterprise scale.

All told, MI viability is typically assessed on the basis of small-scale proof-of-concept pilots of models/algorithms. That's not enough. A more holistic view is required because more often than not, deployment failure can be attributed to organisational constraints, not model problems. The criteria to evaluate a new process should include integration of direct (development and running) and indirect (organisational and opportunity) benefits and costs. While chief data officers and scientists have become common-place at insurers, inchoate firm-wide data strategies and inadequate underlying technology hinder their effectiveness. System design, deployment plans and success criteria should focus on business workflow context, decision support and enterprise productivity. Regulatory risks regarding tech-linked innovation in insurance, in particular around data privacy and use, also need to be considered. Importantly, an MI project also needs clear and understandable communications across all facets, to secure senior management buy-in and funding.

Glossary: key machine-learning terminology

Term	Description
Algorithms	A list of computer-implementable instructions.
Machine intelligence	A collection of programs and processes that enable a machine (eg, computers) to apply data and information to solve problems.
Conventional curve fitting	A basic form of MI (eg, generalised linear models). These rely on assumptions to understand how variables relate to each other, with the aim of *creating a curve that best-fits* the relationship between data points. Conventional curve fitting can capture some types of non-linear relationships, also.
Machine learning (ML)	Algorithms that learn from data and analyse more complex, inter-related and non-linear relationships among variables. Commonly used in classification, regression and pattern recognition.
Artificial intelligence	AI goes beyond ML by facilitating *adaptive application of understanding*. With these algorithms, machines can store and apply learning flexibly, including to contexts not originally intended.
Supervised learning	To train a machine *using data which are labelled*, ie, already tagged with the correct answer. These labelled data act as supervisor. The machine infers relationships from this sample, which are used to map new examples.
Unsupervised learning	Unsupervised learning is used when *labelled data are not available*. With no teacher to train the machine, it has to discover hidden structures in the unlabelled data on its own. Used for clustering and association.
Clustering and association	A clustering algorithm seeks to discover *inherent groupings* in the data, eg, grouping policyholders by purchasing behaviour. An association problem is when an insurer looks to find out *rules that describe* the data, eg, policyholders that buy X policy also tend to buy Y policy.
Reinforcement learning	Goal-oriented algorithms (agents), which answer the question, *how can this be optimised*? Eg, how can marketing investment be optimised to extract maximum ROI? Learns by interacting with its environment.
Ensemble learning	Uses *multiple algorithms in combination* to obtain better predictive performance than could be obtained from any one of the algorithms alone.
Data engineering	Data engineering is the process of collecting, curating, storing, and transforming data for analytical purposes.
Deep learning	Imitates the human brain to *learn without human supervision*, with data that is unstructured and unlabelled.
False positive	A prediction which wrongly indicates that a particular condition or attribute is *present*.
False negative	A prediction which wrongly indicates that a particular condition or attribute is *absent*.
Physics-based ML	Machine learning that *incorporates a model (eg, hydrodynamic) built using a valid scientific theory based on physical systems understanding* into an ML algorithm/process to provide more structure to the model than would be the case for a less constrained ML model (eg, supervised or unsupervised learning.) This hybrid approach is often easier to interpret and diagnose.
Generative adversarial networks	Generative adversarial networks (GANs) involve learning patterns in data so that the *model can generate new examples* that seem credible enough to belong to the original dataset. The original data and the generated data can then be played off each other in the context of competing neural networks to develop better models.
Causal inference	Causal inference in ML refers to approaches that provide more structure for control and prediction by building capabilities that identify actual drivers of outcomes *to make an ML process more robust to changing circumstances*, eg, attempting to sort out causal drivers of obesity to distinguish what can be controlled across different sub-populations or analysing what design choices lead to more clicks on a website.

Source: Swiss Re Institute

Machine intelligence: establishing a common understanding

Machine intelligence (MI) is an umbrella term covering a range of data processing and manipulation techniques, from conventional logistic regression to sophisticated deep learning. The more advanced MI are generally categorised as machine learning (ML) and artificial intelligence (AI). Today, conventional techniques can be more easily scaled up to augment existing processes. However, in recent years there has been exponential growth in the ease of use and effectiveness of algorithms, and more sophisticated ML and AI could eventually supplant conventional approaches.

Information and data processing

MI encompasses programs and processes that use data to enable a machine to solve problems.

Machine intelligence (MI) is often used as a synonym for artificial intelligence (AI), a term for which common understanding is equally variable and/or vague. To establish a consistent reference understanding for the purposes of this report and public debate, we define MI as a collection of programs and processes that enable a machine (most often computers) to apply data and information to solve problems. In most cases, human intervention is required to make the MI-enabled process useful. We include the following categories within this umbrella definition of MI:

- **Conventional curve-fitting or traditional statistical approaches,** such as generalised linear models (eg, linear or logistic regression). These approaches rely on assumptions to understand how variables relate to each other, with the aim of constructing a curve or mathematical function that best fits the relationship between data points. Note that conventional curve fitting can capture some types of non-linear relationships, also.

- **Machine learning:** Algorithms that learn from data and analyse more complex, inter-related and non-linear relationships among variables. Commonly used in classification, regression and pattern recognition.

- **Artificial intelligence** goes beyond ML by facilitating adaptive application of understanding. In AI, algorithms mirror human-like qualities such as the ability to respond to contextually ambiguous situations. With these algorithms, machines can store and apply learnings flexibly, including to contexts not originally intended. In this vein, more advanced AI is sometimes called neuromorphic or cognitive computing. Some newer AI appear to reflect new kinds of networked intelligence called hive intelligence.

ML and some kinds of AI are more complex analytical approaches to data processing.

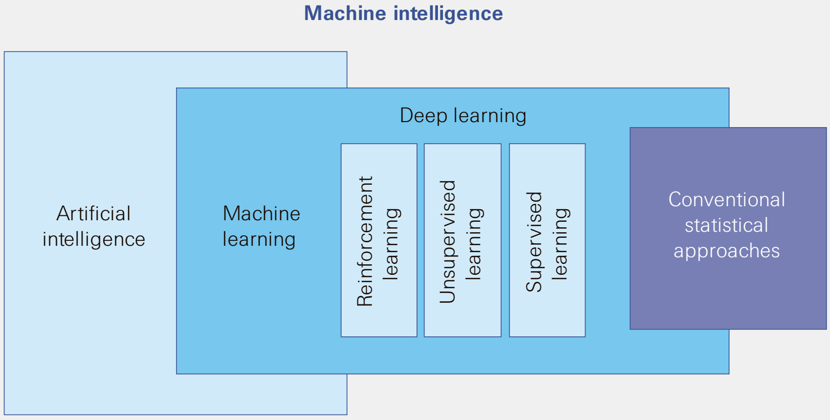
Conventional curve-fitting, ML and AI can be independent and interdependent. As Figure 1 shows, AI typically incorporates both ML and conventional curve-fitting methods. Within ML, supervised learning has seen widespread adoption. Here human intelligence is used to embed each piece of sample data with meaningful tags that help an algorithm understand the data. Unsupervised learning is the method used when labelled data are not available, such as to detect data clusters (eg, in insurance to group policyholders by purchase behaviour) or anomalies (eg, fraud detection, in partnership with human claims experts).[1]

Newer types of ML are still in early stages of development.

Reinforcement learning (RL) exhibits more adaptivity and has been successfully applied in augmented reality tools such as in gaming.[2] However, the field is still in early stages of development in other sectors. RL algorithms are not restricted to existing data, but search out optimised solutions based on rewards or penalties related to each action taken. RL can be combined with simulations and data augmentation to compensate for incomplete, messy, non-stationary or biased data.

1 A key task is to detect any specific grouping or clustered behaviour in the observed data.
2 Reinforcement learning is an area of machine learning concerned with how software agents ought to take actions in an environment in order to maximise the notion of cumulative reward.

Figure 1
Schematic showing overlapping areas within machine intelligence



Source: FSB, Swiss Re Institute

Conventional MI models can be more easily scaled-up to augment existing processes in a firm.

The objective of MI, particularly when deployed in enterprises, is to supplement or emulate human deduction, reasoning and problem solving. An MI-enabled solution can be more successful and transformative when based on a conventional approach like logistic regression, even if ML or AI models, given the greater number of variables they can analyse, perform better in predicting outcomes. This is because the conventional model can be more easily scaled-up to augment existing processes in an organisation. Figure 2 ranks different MI techniques in terms of complexity criteria: interpretability of model results, ease of implementation, stability of models to changes in data, and execution speed. These factors determine success of MI solution deployment. We expect that as algorithms become faster and cheaper, more sophisticated ML and AI models could supplant conventional methods.

Figure 2
Complexity spectrum of different categories of MI

Complexity	Approaches/Algorithms*	Interpret-ation	Implement-ation	Stability	Execution speed	Category
Low	Generalised linear models					Supervised learning
	Naïve Bayes classifiers					Supervised learning
	Instance-based learning					Supervised learning
	Support vector machines					Supervised learning
	Decision trees					Supervised learning
	Random forest					Supervised learning
	Gradient boosting					Supervised learning
	Deep learning (CNNs, RNNs, NLPs, BERT^ etc)					Supervised/semi-supervised
	Generative adversarial networks (GANs)					Supervised/unsupervised/semi-supervised
	Optimal classification trees					Supervised learning
	Reinforcement learning					Reinforcement learning
High	Ensemble learning					Combination of techniques

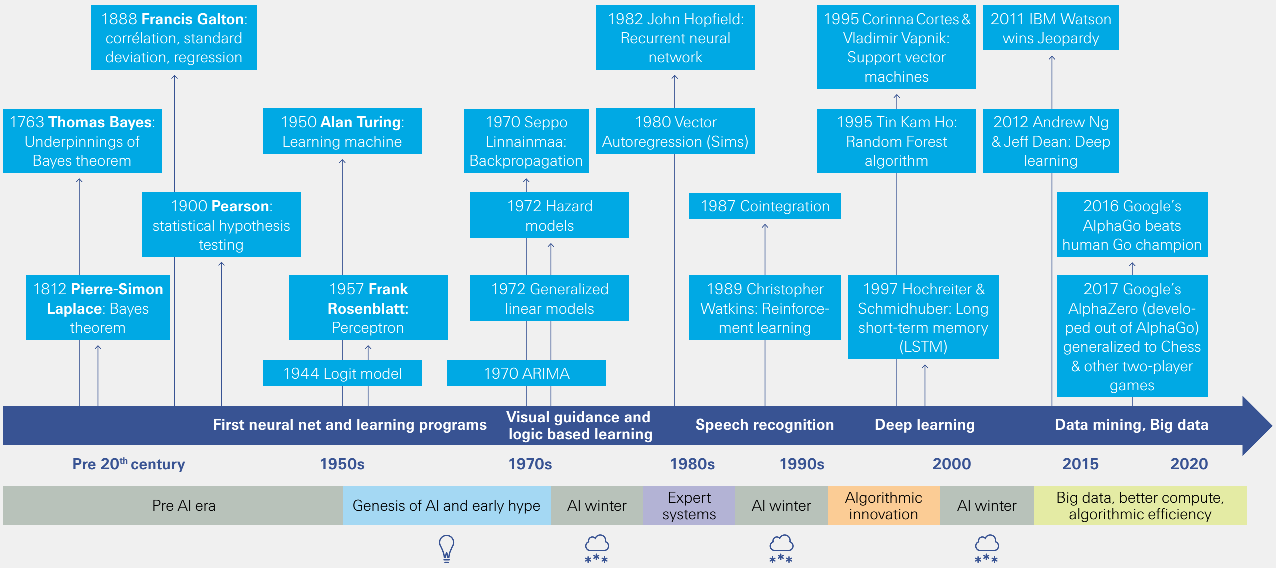
*Certain approaches/algorithms may be better or less suited than others for solving certain problem types
Notes: Interpretation refers to algorithm transparency and explainability. Implementation refers to effort required to develop algorithms and the volume and complexity of data needed to train models. Stability refers to sensitivity of performance to change in data and assumptions. Execution speed refers to time to complete end-to-end execution, starting from data ingestion to final results.
^ Bidirectional encoder representations from transformers (BERT) is a substantial improvement on natural language processing (NLP). CNNs are convolutional neural networks; RNNs are recurrent neural networks; GANs are generative adversarial networks.
Source: Swiss Re Institute

Beware the hype

The evolution of AI has seen periods of very intense and lesser interest.

AI as a term was first coined in the 1950s. This technique has proliferated since then, interspersed with periods of reduced funding and interest ("AI winters"). There was a spike in interest in the early 1990s with the development of neural networks.[3] However, the new potential was not fully realised leading to another period of AI winter in the 2000s, which has only thawed in the last 10 years.

Figure 3
MI development timelines



Source: Swiss Re Institute

The periods of interest can also generate a certain level of hype.

There has been notable progression in AI and ML capabilities in the last years due to improvements in processing power, developments in cloud computing, an explosion in data, and digital transformation. This has also led to, in our view, an unwarranted level of hype. For example, a 2019 study found that two-fifths of Europe's labelled AI start-ups that claim to use AI actually do not. And when they do, the AI use cases are often quite basic.[4] In some instances, mention of AI and ML have been included in business pitches to improve chances of securing financing, although the level of sophistication of the AI being sold is still rudimentary.

Even so, indications are that spending on AI will grow strongly over the coming years.

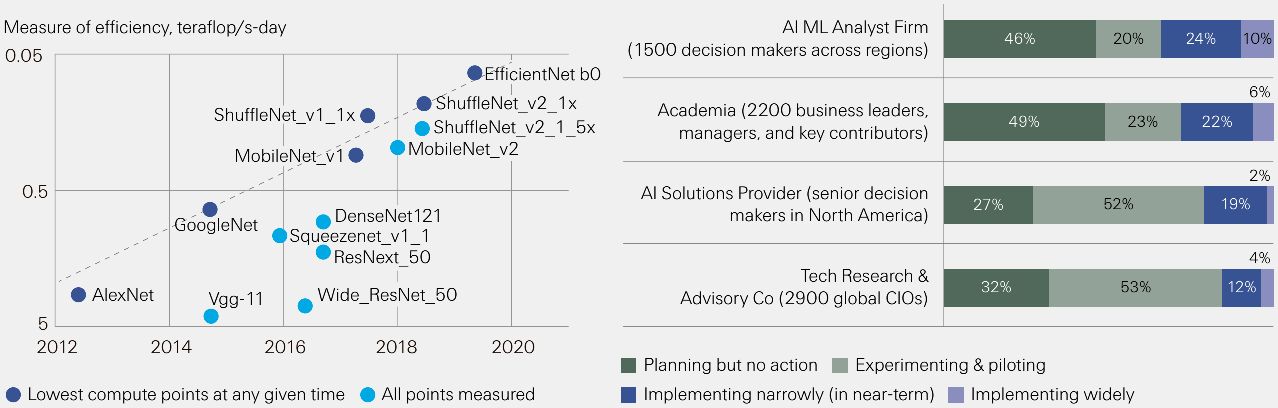
Often technology spending tracks hype. According to estimates from International Data Corporation (IDC), spending on AI systems will reach USD 98 billion in 2023 (a compound annual growth rate (CAGR) of 27% from 2019).[5] These surveys do not necessarily reflect all MI spending, some of which gets lumped in with general IT budgets, so it is difficult to find a comprehensive estimate. Of late, many categories of MI have passed through the peak of inflated expectations (the hype), and some are falling into the trough of disillusionment.[6] With respect to insurance specifically, not much has changed over the last 30 years. If anything, the industry has yet to enter the "slope of enlightenment".

[3] A network of artificial neurons mimics the connections and associated responses of the human brain.
[4] *The State of AI 2019: Divergence.* MMC Ventures, March 2019.
[5] *Worldwide Spending on Artificial Intelligence Systems Will Be Nearly USD 98 Billion in 2023*, IDC, September 2019.
[6] *Hype Cycle for Artificial Intelligence*, Gartner, 25 July 2019.

Exponential improvements in algorithm efficiency could make deployment of more advanced MI techniques more profitable.

Nevertheless, the foundational technologies necessary to facilitate successful enterprise-scale MI deployments continue to develop rapidly, as algorithms become faster, leaner and cheaper.[7] The consequence will be that some enterprise processes that are profitable with conventional MI could suddenly become more so with newer MI (eg, ML and AI). Figure 4 (LHS) shows what has been an exponential increase (see Y-axis) in efficiency of algorithms (eg, AlexNet, ShuffleNet) over recent years, far outpacing the rate of capability improvement according to Moore's law.[8] For example, the processing performance of newer algorithms such as EfficientNet in 2019 on a computer vision task far exceeds that of AlexNet in 2012.[9]

Figure 4
Increase in algorithmic efficiency (LHS) and progress in AI implementation (RHS)



Note: a teraflop refers to the capability of a processor to calculate 1 trillion floating-point operations per second.
Source: D. Hernandez, et.al. *Measuring the Algorithmic Efficiency of Neural Networks*, 2020

Source: Swiss Re Institute (based on multiple surveys, data anonymised, standardised and rescaled to make them comparable).

To date, digital-native firms have been the most successful in transformative MI deployment.

We looked at a collection of surveys (see Figure 4 RHS) and found that some enterprises (less than 10%) have managed to build on successful pilots to deploy MI in multiple processes across the organisation. Enterprise-wide transformative MI demands initial investment in digitising firm-wide operations. This will lay the foundation for firms to (1) apply MI to automate processes; and (2) introduce new offerings by cutting across company-wide silos to integrate data. Examples of advanced deployments include MI-native digital-first platforms/firms like Uber and Airbnb. Also included are firms that started off with traditional technology (eg, Microsoft, Amazon) but have since invested heavily in MI capabilities.[10] Firms in more traditional sectors like financial services (including insurance) are playing catch up. To reap large-scale benefits from MI, they first need to digitise operations and break down data silos. We expect to see progress in this direction over the coming decade.

[7] A. Agrawal, J. Gans, et.al, *Prediction Machines: The Simple Economics of Artificial Intelligence*, 2018.
[8] According to Moore's Law, the speed and capability of computers increase every two years.
[9] By 2020, with more efficient algorithms, it took 44x less compute power than in 2012 to train a neural network to the level of AlexNet. Over the same time span, Moore's Law yields 11x improvement.
[10] *AI from exploring to transforming: Introducing the AI Maturity Framework*, Element AI, May 2020.

Machine-intelligence system implementation

MI applicability and implementation are not uniform across industries. Successful implementation is dependent on data availability, interpretability, system complexity and regulation. Implementations require a strong business case, competent system architects and developers, supportive regulations, committed management, and an enterprise-wide, production-ready data strategy. In a post-COVID-19 lower growth environment, ROI will be a key consideration as analytics projects are evaluated. Investment in data engineering capabilities is critical for successful deployments.

Criteria for success

Things to think about include…

When companies think about MI-enabled systems, there are many things to consider: which MI tools to implement; how to address constraints arising from legacy systems; where to find staff to resolve capability deficits; and how to secure more benefits than costs. Table 1 presents four overriding criteria we recommend companies should consider as they prepare for MI-related technology adoption.

Table 1
Criteria for successful implementation of enterprise-wide MI for the long term

Criteria for success	Present inadequacies	What successful projects look like
Long-term positive ROI for end-to-end processes	Underestimation of implementation costs, poorly designed MI-ready workflow processes, scalability difficulties arising from poor planning	Workflow process architecture matches deployed MI strengths, recurring costs detailed in implementation plan, integration questions prioritised, security and privacy built into planning from the beginning
Production-ready data strategy	Underinvestment in data ingestion and curation, lack of detailed data stewardship, insufficient number of data engineers.	Appropriate focus on data strategy, data engineering, data tools, and data modelling.
Use cases that fit business and regulatory contexts	Among others, use cases do not consider data constraints, and organisational issues that prevent MI from adding value.	Deployment leverages MI strengths in the context of existing organisation and workflow processes. Deployment plan includes clear match of MI and pain points, data and technical feasibility evaluated from the start.
Management commitment	Senior management not adequately briefed on proposed MI-enabled deployment, poor implementation by frontline staff, poor coordination and buy-in among business units, and shallow understanding of MI operations.	Regular and detailed updates to management, willingness to change the process to accommodate new findings, and ongoing investment and talent continuity to extract value.

Source: Swiss Re Institute

Long-term return on investment

…return on investment.

To secure positive return on investment (ROI) from spending on MI-transformation of end-to-end processes across a firm over the long-term, factors to consider include:

- **Net benefits from workflow transformation vs recurring costs:** The benefits from the transformation exercise in terms of reduced cost, increased revenue and new business opportunities must be greater than the costs of organisational integration, both direct (development and running) and indirect (organisational and opportunity costs). In a recent survey, 93% of respondents at US insurers going through MI transformation expressed concern around the costs of implementation and ROI.[11] Many ROI calculations look only at the cost of vendor solutions, not the added costs to the business (eg, data curation, training, etc). Further, ROI is an ongoing calculation as new inputs (eg, regulation and changes in costs of key data) can render initial cost/benefit estimates inaccurate. Figure 5 presents key cost/benefit considerations for deployment of such systems.

[11] *State of Artificial Intelligence and Machine Learning in the Insurance Industry Study*, LexisNexis® Risk Solutions, 3 December 2019.

Figure 5
Cost/benefit considerations for deployment of MI-enabled systems

Financial model build: For example, a model with anticipated ROI is available for review	**Potential upside return:** eg, Increased customer retention of 5%. Decreased cost per transaction of 5%	**Potential downside risks:** eg, decreased customer retention (10%). Increased cost per transaction (5%)
Worst-case downside: eg, Automated claims processing crashes or does not filter fraudulent claims	**Liability:** eg, Automated underwriting system does not capture the risk accurately	**Cost of building model:** eg, six months for a team of two data scientists
Cost of maintaining model: eg, 10 hours a month	**Quality of predictions vs expectations:** eg, model assumes 100% correct predictions, but results 85%	**Uncertainty in model predictions:** eg, prediction might be accurate +/–10%

Source: Adapted from V. M. Megler, *Managing ML Projects, Balance Potential with the Need for Guardrails*, February 2019

- **Performance and efficiency of re-engineered end-to-end processes:** MI implementation must be scrutinised in technological terms and with respect to business drivers and constraints. For example: (1) a new MI-enabled process should be at least as accurate and robust as the current process, and more time efficient; and (2) the new process should generate new and sustainable business opportunities. At times, the cost expectations of realigning human resources to accommodate forecast benefits may not be realistic. For example, anecdotal evidence suggests that ML fraud detection systems sometimes flag far more new cases than existing staff can verify. Measurement metrics must be tied to realistic business outcomes rather than mere model performance.

- **Investment needed to maintain system quality and robustness:** A key criterion for success of MI-enabled systems is an effective continuous monitoring framework for model lifecycle management. Model development – even for most advanced MI – is often the more straightforward and less costly aspect of enterprise MI deployment. Integrating a new MI system into an organisation, on the other hand, will likely require workflow process re-engineering and constitute most of the system deployment costs. Further, maintaining the integrity, security and privacy of a new system will require a large budget at first (although for well-architected systems these running costs should decrease over time).

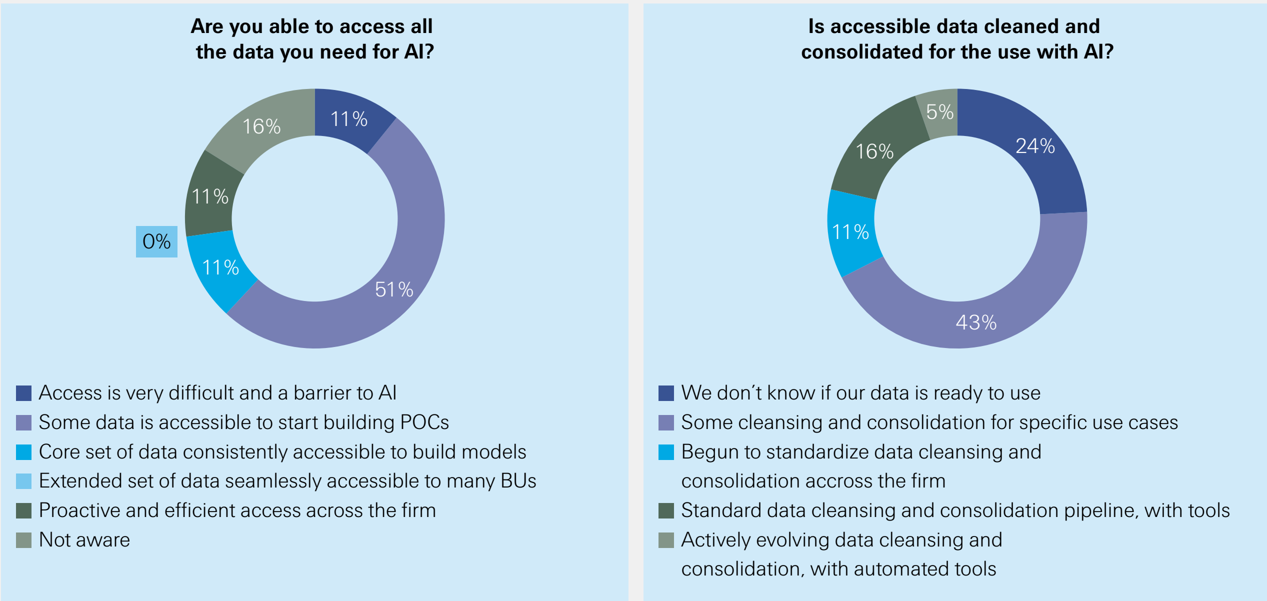
- **Exception handling:** In the post COVID-19 environment, models may give unexpected results due to major shifts in consumer behaviour, data inputs and the way businesses are run. However, the renewed emphasis on digitalisation will create more and diverse data sets to further refine MI-models, and widen the scope of training information available to better exception handling capabilities. On the other hand, companies already investing heavily but facing cost pressures may now prioritise projects more carefully and continue with projects that already deliver positive ROI or are close to doing so. Several areas will continue to see greater attention and investment, including automation and fraud detection capabilities.

Production-ready data strategy

MI-enabled system performance materially depends on data-value-chain management

Often, deployment fails because of poor data engineering. In an end-to-end enterprise process, low-quality algorithms with high-quality data engineering will tend to outperform high-quality algorithms with low-quality data engineering. In financial services, firms typically start off with developing an algorithm and then under-invest in data engineering. For transformative impact enterprise wide, they should do the reverse. Figure 6 presents findings from a recent survey on the low maturity levels of insurers in terms of accessing and curating data for AI models.[12]

Figure 6
Data maturity of insurers



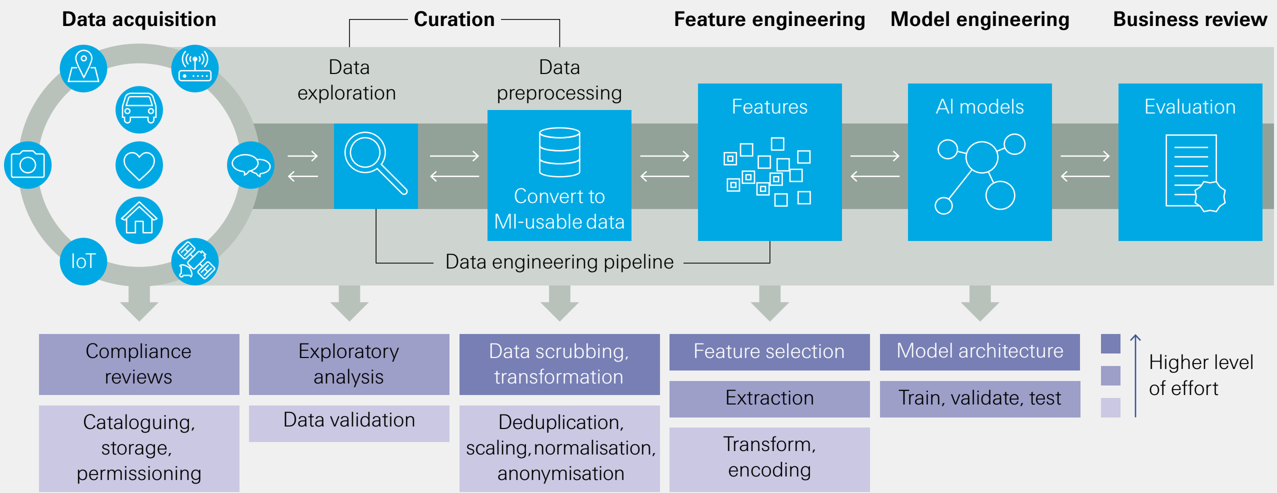
Source: *The Five Dimensions of Enterprise AI, Element AI*, May 2020, Insurance respondents only

Lack of data strategies can detract from the effectiveness of newly-created CDO roles.

Some of the most innovative AI under development (eg, reinforcement learning and ensemble modelling) can be robustly trained without extensive high-quality training data. In these cases, systems use simulations, data augmentation algorithms and synthesise subject-matter expertise.[13] However, these new approaches are not yet ready for enterprise scale and nearly all successful MI-enabled system deployments still depend heavily on data quality and quantity. Here, the capacity and tools available today to process structured and unstructured data open new doors of MI-related opportunity. To optimise the potential, a firm-wide data strategy and system architecture such as represented in Figure 7, is essential. The lack of said strategies and architectures hamper the effectiveness of the also critical newly-created CDO roles. For example, in a recent survey, less than 10% of CDOs across industries said they are able to measure the financial value of their information and data assets.[14]

[12] Survey with senior decision-makers at large organisations in the US and Canada. See Element AI, May 2020, op. cit.
[13] Integrating subject matter expertise in this way reflects a Bayesian updating approach to ML or AI in the sense that subjective priors are incorporated into the algorithm to reduce the scope of evaluated parameter space, which can counteract (assuming the priors are sound) data inadequacies.
[14] *Gartner Survey Finds Chief Data Officers Are Prioritizing the Right Things, But Higher Strategic Focus Is Required*, Gartner, 10 June 2019.

Figure 7
MI-data strategy schematic



Source: Swiss Re Institute

Data engineering is often duplicated across functions.

For insurers investing in data architecture, a common characteristic of poor strategy, is duplicative data engineering across different teams. In these cases, data scientists rarely feel confident about a centralised curation process. In some cases, there is no centralised curation process, or none that the data scientists are aware of, leading them to create their own, often duplicative, processes. IDC found that data professionals spend 67% of their time searching for and curating data.[15] In another insurance specific survey, nine out of 10 employees at the 100 largest firms in the US said that managing increased volumes of data was their number one challenge.[16]

Insurers also lack comprehensive data ontologies that define relationships among data.

In our view, a centralised data ingestion and curation capability can generate sizable ROI by overcoming such inefficiencies. This is an area where insurers in particular, have a long way to go. A recent survey found that as many as 75% of insurers lack a taxonomy to harmonise different types of data.[17] Most also do not have comprehensive data ontologies that define multi-dimensional relationships among the classified data in a data taxonomy (or in multiple taxonomies), an issue that is set to become more complex as data sources grow in number and diversity.

Factors to consider for best-fit use cases

Many failed implementations arise from a mismatch of algorithms to use case.

Successful MI-enabled implementations require matching of desired outcomes with the best-suited enterprise-ready algorithms and techniques. Not every algorithm works for every use case, and many failed implementations arise from a mismatch of algorithms to use case. Well-calibrated traditional statistical methods can offer similar results in terms of accuracy to advanced models, suggesting that data quality matters more than algorithmic innovation. Importantly, business use case and data availability should drive technique selection. Even with best efforts to match technique to use cases, a trial and error process ensues as different techniques are implemented and tested to determine which approach works best. Over time we

[15] *End-User Survey Results: Deployment and Data Intelligence in 2019*, IDC, November 2019, sourced from "Talend Accelerates Path to Revealing the Intelligence in Data", 27 February 2020.
[16] LexisNexis® Risk Solutions, op. cit.
[17] *Building New Data Engines for Insurers*, BCG, 5 November 2018.

expect consensus to develop with respect to which techniques work best with specific use cases. Factors to consider when matching technique/use case include:

- **Interpretability:** Questions to ask include: How much does the business need to understand? How much would it normally understand? What are the regulatory requirements and professional standards? In insurance, about a third of firms that have adopted MI are concerned that if regulators do not understand newer techniques, they could block or limit efforts to use new applications.[18] Also, depending on use case, firms may be required to be transparent to both regulators and customers. Typically, a high degree of model/process transparency is required, and the line between collecting data to improve service and compromising privacy is very thin. Getting the balance right can impact funding. Gartner forecasts that by 2022, projects will be twice as likely to receive funding if they have built-in transparency.[19]

- **Use case selection should consider the costs of different errors:** While models process large volumes of data rapidly, they can also lead to relaxed human oversight. Each error has a cost, and management must decide on acceptable levels of error tolerance to identify the point at which economic value can turn negative. In some use cases, all types of errors may have an equal impact, but in others one can prove more costly (eg, if a self-driving car ignores a pedestrian or an automated credit evaluation system extends credit to a company that later defaults). At other times, the cost of a false prediction may be greater than the savings associated with a true prediction. For example, an insurer looking to rapidly assess property claims using MI-based aerial image analysis may later have to increase reserves significantly because of non-visible damages (eg, under a roof).

  A false positive may prove less expensive in cross-selling campaigns (where the cost is a wasted email) than in underwriting or pricing (where the cost is accepting sub-standard risks). Table 2 demonstrates this trade-off in two scenarios: 1) propensity to buy in a cross-selling campaign; and 2) classifying a critical illness (CI) risk for underwriting decision. In cross-selling, the cost of approaching an unwilling prospect based on a less accurate propensity to buy prediction (false positive) is far lower (eg, USD 10) than the false-positive in the underwriting of a CI policy, when a risk classified as good is actually bad (eg, USD 1100).

Table 2
Assessing the impact of error costs

Propensity to buy	Predicted (unlikely to buy)	Predicted (likely to buy)
Actual (unlikely to buy)	(True negative) **3 000**	(False positive) **600**
Actual (likely to buy)	(False negative) **400**	(True positive) **6 000**

Cross-selling scenario	Number of predictions	Gain (Loss) per prediction in USD	Total gain (loss) in USD
True positive	6 000	100	600 000
False negative	400	Don't approach	–
False positive	600	(10)	(6 000)
True negative	3 000	Don't buy	–
Total	**10 000**		**594 000**

Critical illness classification	Predicted (Bad risk)	Predicted (Good risk)
Actual (Bad risk)	(True negative) **3 000**	(False positive) **600**
Actual (Good risk)	(False negative) **400**	(True positive) **6 000**

Underwriting scenario	Number of predictions	Gain (Loss) per prediction in USD	Total gain (loss) in USD
True positive	6 000	100	600 000
False negative	400	Don't underwrite	–
False positive	600	(1,100)	(660 000)
True negative	3 000	Don't underwrite	–
Total	**10 000**		**(60 000)**

Low cost of error — High cost of error

Note: This is a simplified illustration and may not capture all possible scenarios.
Source: Swiss Re Institute

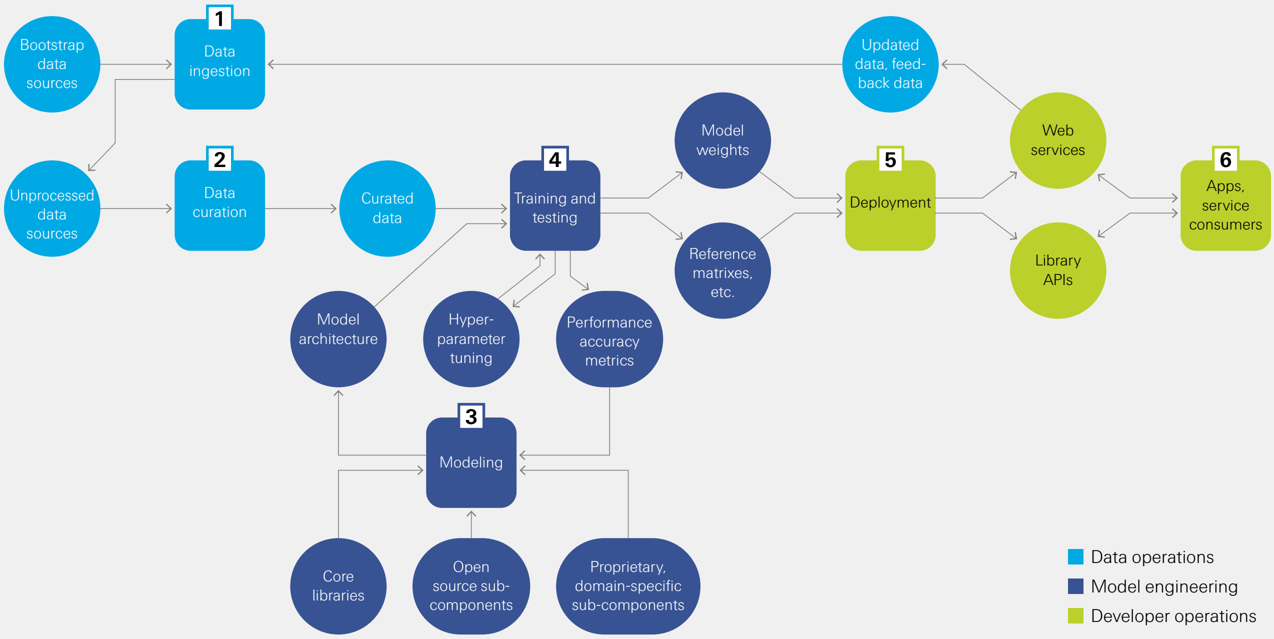
[18] LexisNexis® Risk Solutions, op. cit.
[19] *Can learnings from early projects give CIOs a head start with AI technologies?*, Gartner, 9 February 2018.

Transformative, successful MI deployment requires a cohesive strategy that explicitly includes details of necessary capabilities.

Organisational maturity and willingness

Transformative MI deployments require much more than just model/algorithm development. Equally critical for success is a cross-functional, detailed strategy with senior executive sponsorship. An important part of the strategy is a focus on capability, in terms of technology, process re-engineering and staffing. An MI implementation workflow requires capabilities across data engineers, model engineers and software developers/IT operations (see Figure 8). It is also desirable to have staff with multiple skill sets who can translate requirements clearly across functions. In insurance, firms sometimes hire actuaries with programming skills to reduce miscommunication when actuaries hand over their models to development engineers without knowledge of statistics.

Figure 8
MI implementation workflow



Source: Swiss Re Institute

Machine intelligence in insurance

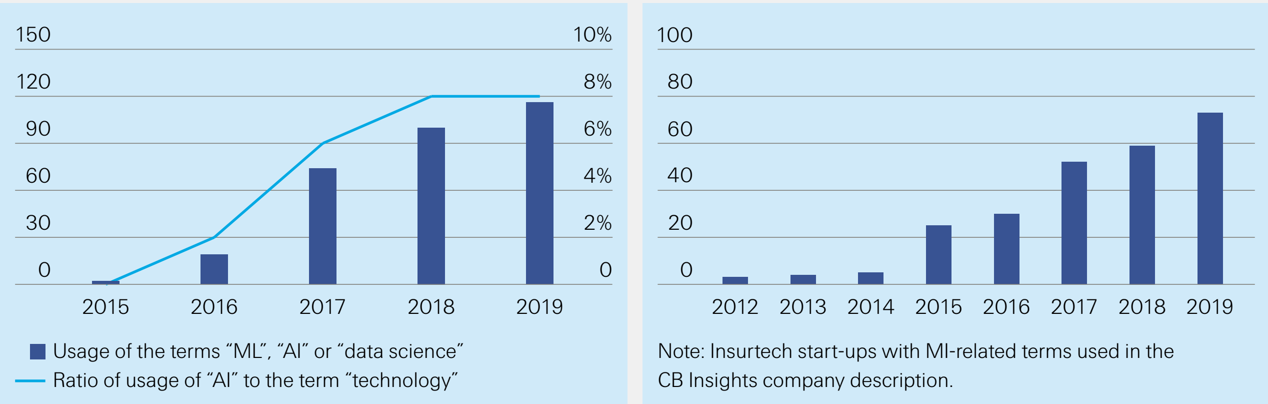
Insurers continue to experiment with newer MI approaches to build upon (and possibly replace) conventional MI techniques that are becoming standard practice in areas like customer analytics and claims processing. However, unlike in sectors such as social media, end-to-end transformation of insurance processes through MI-enabled systems remains elusive. Data availability, model interpretability, and privacy issues remain barriers to large-scale adoption. The cost of errors in insurance can also be high.

How things stand

Insurance executives remain optimistic about MI.

The insurance industry has lagged in implementation of MI-enabled systems. Still, a 2019 survey found that industry executives have high expectations about adopting ML in 2021. They were optimistic in the past too, with past surveys projecting high expectations for where they would be in 2019, although actual adoption last year was well below predictions.[20] It also appears that insurers have become more vocal about MI. Figure 9 shows that mention of MI-related terms (eg, AI, ML and data science) in investor annual reports has risen significantly, from two citations in 2015 to 116 in 2019. The growth in Insurtechs using AI/ML technology and MI-related patents filed by insurers in recent years mirrors this trend.

Figure 9
MI-related mentions in insurer investor reports (LHS), and trend in Insurtechs using AL/ML technology



Source: Annual reports of 30 leading insurers, CB Insights, Swiss Re Institute

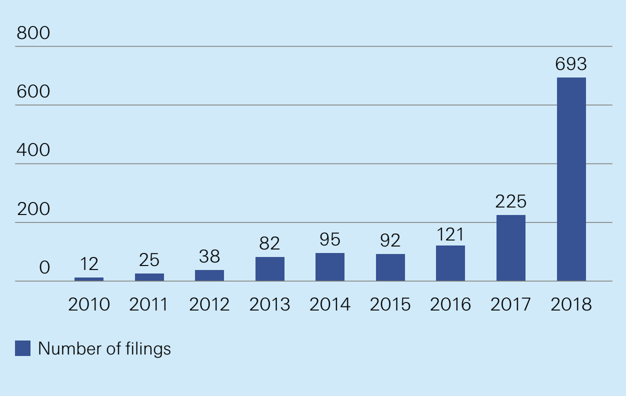
Strong growth in insurance-related AI/ML patent filings

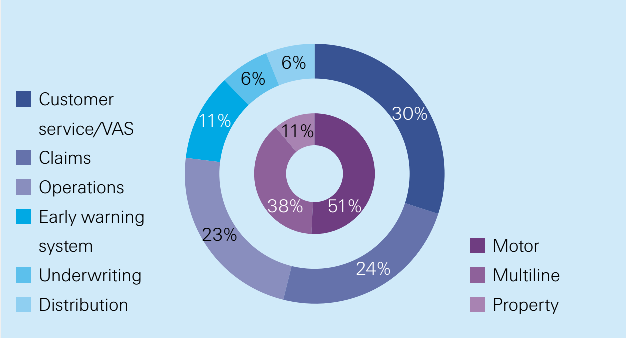
MI-related patents filed by insurers have increased exponentially in recent years.

We analysed patent databases and found that the number of MI-related patents filed by insurers has increased since 2010. Focusing on the most prolific patent filers among insurers in the US, in 2018 and 2019 more than half were for motor business, some in the area of autonomous vehicles. As MI-enabled processes enter business-critical vehicle systems, it expands demands for greater innovation in MI-enabled monitoring. For instance, many patent applications are for remote sensing, image processing and drone use for damage assessment.

[20] *Machine Learning: Today and Tomorrow*, Willis Towers Watson, 25 February 2020.

Figure 10
Growth (LHS) and composition of patents (RHS, 2018) at insurers





Source: Google Patent Database, Swiss Re Institute

Most recent innovations have sought to improve customer service, claims and operations.

In terms of use cases, most MI patents in insurance have been for functionality designed to improve customer service, claims efficiency and reduce losses. Some were aimed at generating early warning signals, such as to alert drivers to pedestrian or cyclist presence, or to alert vehicle operators of malfunction. Both reduce claims frequency as well as severity. Comparing China and the US, patent filing concentration is high in both markets, according to the Google patent database. However, the number of MI-related patents filed between 2010 and 2019 was more evenly distributed in the US, with 10 insurers accounting for 80% of activity. In China, less than five insurers accounted for 85% of MI-related patents filed.

Already-live AI and ML application deployment in insurance

The use of some conventional MI tools is standard in insurance.

Conventional MI such as generalized linear models have become standard tools in insurance for risk assessment and prediction models. More recently, enthusiasm for a range of AI and ML techniques such as deep and reinforcement learning has led some insurers to run pilots. In a few cases, early adopters of AI and ML are seeing benefits in select areas, such as faster claims settlement, more targeted cross- and up-selling, improved fraud detection and better risk scoring.

- **Modernising claims analytics:** Much of claims processing is still manual. A number of insurers now have pilots in triaging, routing, validating and corresponding with third parties, the ambition being that some degree of automation will materially reduce the cost of claims processing.[21] Simpler tasks like assessing high-volume losses and processing well-specified items are more likely to be successfully executed by MI-enabled systems. Areas where insurers report higher savings from MI include those where information is better structured, such as documentation in standardised formats.

- **Fraud detection and claims mitigation:** ML techniques are well suited to use cases involving large classification of data and anomaly detection, such as fraud detection. Increasingly, insurers are evaluating and deploying ML-based fraud solutions that augment internal data with new sources of information, including third-party IoT and public data. Insurers are also using ML to create entirely new loss mitigation offerings, which can in turn lead to lower claims. Such is the thinking behind Direct Line's telematics programme, for example, which uses ML to identify individuals who need coaching to become better drivers.[22]

21 "The challenge of full automation", *insuranceinsider.com*, 2 April 2020.
22 *Direct Line Group saves young drivers over £50 million in motor premiums*, Direct Line, 1 Feb 2019.

- **Distribution channel optimisation:** Another area where ML is seeing application is in agent recruitment and retention. Insurers have started using ML-enabled systems to identify individuals most likely to become successful producers. These systems can also improve producer-client matching. For example, Discovery does real-time, automatic matching of call centre agents to members with whom they are likely to have the highest affinity. The model has been operational since 2018 and customers on calls where affinity was matched reported greater satisfaction.[23]

- **MI in customer experience:** MI has been deployed at enterprise scale in many social media and online retail contexts. Some insurers have sought to do the same, with the ambition to increase the effectiveness of targeted marketing. Despite early successes, insurers discovered that rushing out MI-inspired initiatives may not necessarily generate the desired outcome. For instance, anecdotal evidence suggests that targeted digital advertising based on previous interaction with a product can actually turn a customer off. This result suggests that MI models could benefit by using insights gained from behavioural economics to disentangle interaction effects.

- **Underwriting:** Given the level of confidence needed to deploy new technologies in underwriting, fully AI and ML-enabled underwriting systems still do not exhibit levels of accuracy necessary to be used at scale. This also means that MI cannot be relied on to completely replace risk assessments, except in simpler lines. This said, some examples related to supervised learning, can complement and or eventually replace parts of existing processes in insurance. These include smarter mechanisms for triage and routing, which may be more effective than current business rules, eg, triage between depths of investigation (full vs. simplified underwriting), safely waive additional evidence (lab tests, physician statements) or allocate referrals to the right level of seniority in the organisation (junior underwriter vs. medical officer).[24]

- **Pricing:** This is subject to regulatory approval, and the traditional approach involves fitting a GLM to historical claims and premiums. More accurate pricing models based on newer machine learning techniques cannot be put into production immediately, as results may be difficult to explain both internally and externally to regulators. There may also be other constraints to using the data like cost and lack of access to data.

Inadequacies in existing implementations

Challenges to scale use of newer MI tools remain.

The challenge to scale AI and ML models continues to hinder deployment of newer MI technologies at the enterprise level across core workflows in the insurance value chain. The following are processes where MI could potentially be implemented at scale and the associated still-existing obstacles that hold back broader adoption:

1. Collecting and curating relevant structured and unstructured data. Here obstacles include data privacy regulations and incentives (eg, firms or agents unwilling to share relevant data), fragmented access processes, inadequate data usage contracts, and still-difficult to systematise data curation processes.

2. Assessing, understanding, and processing relevant input information. NLP techniques are still inadequate given the difficulties in interpreting tacit and subtle informational aspects, and data quantity and quality remain poor.

[23] *Insurance trend #1: Get to know me*, EFMA, 14 November 2019.
[24] *What's new? The next wave of insurance automation complemented with new technologies*, Swiss Re, 25 November 2019.

3. Underwriting approval and pricing: Intelligent automation integrating humans and machines is still a massive design challenge. Seasoned executives and underwriters do not trust algorithms given examples of "obvious" misses.

4. Monitoring risk portfolios and managing claims: Challenges to efficiency improvement remain in terms of creating lower-cost systems that have better false positive/false negative trade-offs than human-centric methods. Data processing architectures still treat data in "pools" rather than the "rivers" necessary to accelerate time between data collection and usage. Data not transitioned into actionable insights near immediately hinder MI-enabled system usefulness.

5. Improving capital allocation across liability segments: Prediction models still fall short in terms of reliably supporting better capital allocation. Data are incomplete and biased in many liability segments, and systems are still designed around current processes that are not MI-ready.

Poor MI integration hinders system deployment potential

Poor integration of MI-enabled systems across processes can hurt project outcomes.

System design and management often fall short when insurers attempt to implement MI into existing cross-functional processes. Too few resources are dedicated to integrating models and algorithms into workflows, leading to poor cross-functional coordination. In an interview with Swiss Re Institute, one insurer seeking to eliminate unnecessary underwriting questions said it leveraged banking transaction data to offer accelerated underwriting to prospects. The MI-enabled underwriting model performed well in classifying individuals into standard and sub-standard risks. The marketing department, however, did not invest in a propensity-to-buy exercise nor modify its sales process, which nullified the benefit of the system.

Data collected from IoT devices currently have limited integration into underwriting and pricing.

Another challenge is that new data (especially collected from wearables) for underwriting and pricing purposes may not necessarily lead to more accuracy in underwriting (See Figure 11). For example, tracking the number of steps one walks may not materially improve one's health. In many cases, the outcome is the opposite: an individual who walks more may also think he/she has license to eat more because he/she is fitter. To this end, there has been a tendency to over-estimate the extent to which collecting and crunching these data actually changes risk profiles. The industry will struggle to adopt IoT data without a clearer understanding of how these insights on behaviour correlate with actual risk experience.

Figure 11
Degree of difficulty in incorporating new data sources into processes

	Easily obtained-traditional data Data via application, FNOL, medical info, motor vehicle records	**Now easier to obtain digitally-traditional data** Tax filings, property registrations, births & deaths, crime records, government data	**New data sources** Electronic health records (EHRs), third party data bases, photos, images, voice, data from wearables & IoT devices, behavioural insights, reviews
Predictive power	Established	Partially established	Exploratory phase
Integration challenges	Addressed	Being addressed	Less known
Acquisition cost	Low	Stabilising, was high	Still expensive
Regulatory barriers	Minimal	Moderate	High

Source: Swiss Re Institute

Recommendations for current-day MI initiatives

We expect that successful implementation of MI-enabled systems in end-to-end processes will reap many productivity benefits for insurers, with the ultimate outcome of boosting profitability. However, for many firms there is still a long way to go to being fully "MI-ready". This does not negate the positive benefit that existing, often small-scale, MI projects can deliver. The following are some recommendations to improve the likelihood of success in current initiatives.

Focus first on components of a process that are amenable to MI.

Invest incrementally. Insurers should start with a focus on process steps amenable to MI, rather than attempt large-scale transformations. Successful MI-enabled system implementations should start with narrowly defined objectives and follow clear milestones rather than aim for full automation. A number of processes - even in higher-volume lines - can be too complex to fully automate. A good example is auto insurance: one accident can include several smaller claims, each of a different type (eg, bodily injury, vehicle damage, car rental) involving different parties and suppliers, and therefore requiring expert human intervention.

Choose use cases that augment employee effort.

MI can be deployed in functions with fewer regulatory restrictions. While wholesale replacement of some insurance processes may require regulatory approval, augmenting existing processes with selective MI is possible with few regulatory restrictions. Important here is how an MI-enabled system is deployed. Many MI deployments to augment human-centric processes fail by adding process costs without improving overall efficiency or profitability. Involving staff in re-engineering process discussions and introducing small deployment steps can be the difference between successful and unsuccessful implementation.

Use newer approaches like deep learning to complement more conventional techniques.

Combine new and conventional model approaches. Some newer MI (eg, deep learning) methods can be used to supplement more conventional ones (eg, GLM). AI or ML methods may improve data curation, facilitate better process design, and address weaknesses in aspects of the conventional MI (eg, incorporate output from unstructured data.) Insurers should use simple, interpretable models as a baseline for AI or ML, especially in areas that are regulated. For example, a large US insurer acknowledged that because the industry relies so heavily on GLMs, its experiments with deep learning are still focused on developing a multi-variate rating plan. In this case, the end result was to use deep learning to develop new insights; final implementation incorporated these insights to improve the GLM process.[25]

Centres of excellence should foster connections between centralised and local teams.

Foster collaboration between centralised and distributed data science teams. Best practice programs bring uniformity across divisions. At many insurers, if an analytics team in one division builds a successful algorithm for a particular issue, there is little structure to facilitate its adaptation in other divisions. Larger insurers like QBE are building playbooks that all divisions can consider, including algorithms to accelerate claims settlements, identify fraud, improve loss reserving, and suggest when claims cases may become lawsuits.[26]

[25] *Trick or Treat? Application of Neural Networks in Insurance*, KPMG, 10 January 2019.
[26] "QBE, Unlocking the secrets to technological transformation", *Claims Magazine*, April 2019,

Progressing enterprise-scale deployment

With emerging understanding of how MI-enabled systems can improve data ingestion and curation, and augment existing analyses, there has been growing recognition of the applicability of new approaches. These include, for instance, hybrid physics/ML-based models and causal-inference algorithms to improve the predictive power of MI systems. However, failure in enterprise scale MI-system deployment is more often due to larger organisation constraining characteristics. To this end, insurers should focus more on trust, technology, talent and tenacity.

Innovation and new approaches

New and innovative MI can improve existing approaches.

More advanced MI techniques often require more and better data, and more compute power. The lack of one or both can hold back deployment of MI-enabled systems at enterprise scale. Often the difficulties are specific to models or algorithms, which in turn can (not always) be the reason for failed deployment. Where improvement to data quality and/or compute power is challenging, an alternative way to address model problems, which has been the focus of more recent research, is to develop a new approach less sensitive to these issues. Examples are reinforcement learning or ensemble modelling, such as hybrid physics-based and ML models. Table 3 highlights exciting areas of innovation in MI that have the potential to help overcome key problems in existing approaches.

Table 3
Schematic showing positive developments in MI

Key development	Challenge it is addressing
Combining physics-based models with ML	Improve accuracy, interpretability of MI models, while improving predictive and exception-handling capabilities of physics-based models. Strong applications in critical maintenance activities, early warning systems, etc.
Progress around using ML for causal inference	More informed decision-making with higher level of confidence. Better understanding of the impact of interventions. Huge application in sensitive domains, like, healthcare, defence and even insurance.
Advances in visualization tools for decision support	Improve interpretability and diagnosability of complex MI systems. Applications in NLP, image processing, etc.
Better model interpretation techniques	Improve interpretability of current black box MI techniques, while improving accuracy of more interpretable but currently low accuracy techniques like CART.
Intelligent automation: Re-designing workflows	Automated data curation, insight discovery and sharing. Model prototyping in production languages. Potential to save significant time on development as well as ongoing maintenance.
Privacy-preserving analytics	Governments, corporations, academia all join hands to help improve weighting of model parameters and thus the model performance without compromising on data privacy.

Source: Swiss Re Institute

Combining physics-based models with data-driven approaches

Data-driven AI and ML systems often fail to incorporate physical and scientific knowledge.

Purely data-driven AI and ML-enabled systems are not robust. As insurers move from "detect and restore" to "predict and prevent", they may find that data-driven AI and ML-enabled systems for complex applications are not straightforward because they fail to incorporate physical and scientific knowledge into learning and prediction. Often available data are insufficient, noisy and/or biased, which makes it even more important to compensate with theory-based models. Using (typically inadequate) data with current AI and ML algorithms leads to inconsistent results, with the outcome that trained models do not generalise well to out-of-sample testing.

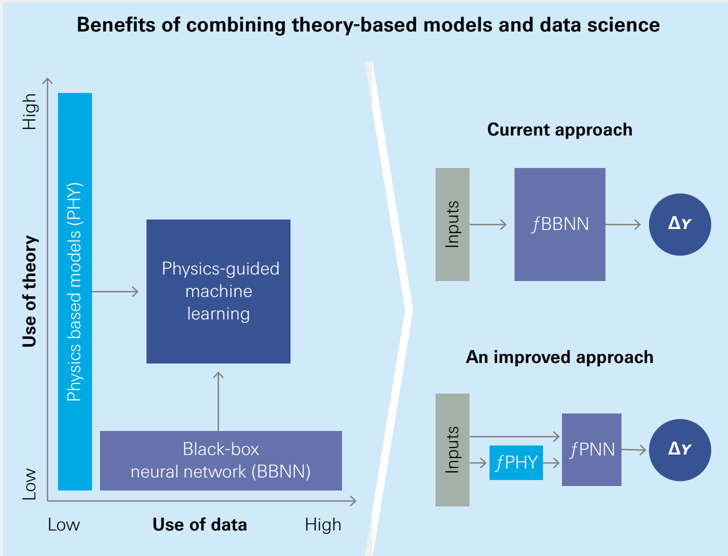
The community is exploring the continuum between physics-based and ML models.

On the other hand, pure mathematical physics-based models may fail to capture the full range of complex interactions characterised by physical systems of interest to insurers (eg, climate, behaviour, urban resilience, health, etc.) To bridge this gap, some insurers and technology developers are exploring hybrid physics- and AI/ML algorithm-based models This hybridisation is called theory-based data science, or physics-based ML, or ML that incorporates the laws of physics. Newer AI such as reinforcement learning, GANs, neuromorphic computing, and agent-based simulation techniques will further expand the possibilities at this hybrid intersection of physics-based models and MI.

Figure 12
Dichotomy between theory-based models and data science models

Characteristics of physics-based versus data-driven models

	Interpret-ability	Robust-ness	Scala-bility	Predictive power
Physics-based model	High	High	Low	Low to Medium
Data-driven MI model	Low	Data dependent	Medium to High	Data dependent



Note: PNN refers to physics-guided neural networks
Source: A. Karpatne et al., *Theory-guided data science: A new paradigm for scientific discovery*, Cornell University, 2016

Successful combination of the two requires clear lines of communication.

Figure 12 is a two-dimensional view of the dichotomy between physics-based and data science models. Science theory-based models (y-axis) can have knowledge gaps with respect to certain processes that are either too complex to understand or too difficult to observe directly. At the other end of the spectrum, data-driven models (x-axis) use large volumes of data but are agnostic to underlying scientific theories. A complementary approach can take advantage of the unique ability of ML to extract patterns from data, while also benefiting from scientific knowledge.

This combination is being experimented in areas such as predicting breakdowns.

This combined approach is being experimented in areas such as predicting breakdowns and remaining useful lifetime for industrial systems. These are areas where physics-based models can be incomplete and data-driven models can be hampered by poor representativeness of training data. Researchers use physics-based performance models to infer unobservable model parameters related to equipment health, which can be combined with sensor readings to generate a data-driven prediction model.[27]

Progress in combining causal-inference tools with MI

MI-enabled systems learn connections, but typically cannot reason cause and effect.

A fundamental assumption of classical statistics and ML is that the distribution of the training data is the same as the distribution of the data in practice. This is often not the case in real life as, for example, new regulation or any other intervention can change the distribution of the data. A general property of causal models is that they are robust to such changes and more interpretable.

[27] M.A. Chao, C. Kulkarni, O. Fink, et.al. *Fusing Physics-based and Deep Learning Models for Prognostics*, Cornell University, 2 March 2020.

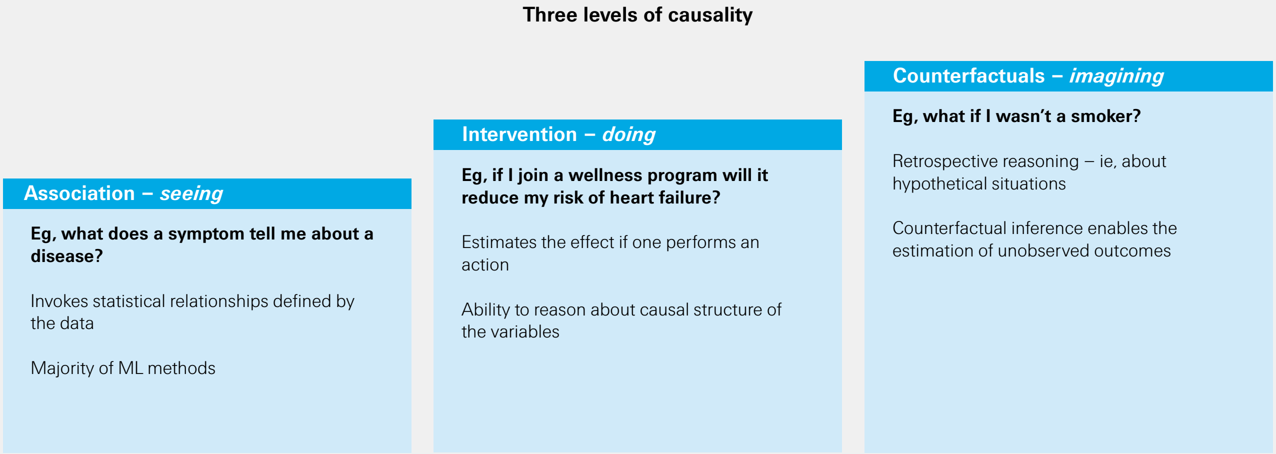
Three levels of causality: seeing, doing and imagining.

Figure 13 shows three levels in the ladder of causality. Level 1 is associational, and asks "how will seeing X change my belief in Y?" For instance, what does a particular symptom tell me about the presence of a disease. Level 2 explores questions which cannot be answered from past data alone. The questions address behaviour changes in response to interventions ("what happens to Y if I do X?"). Level 3 involves imagining, answering counterfactual queries like “what if I had acted differently?"

Causal inference facilitates more adaptive intelligence…

Shifting to such a causal-inference paradigm creates more adaptive intelligence, which aligns with a more precise definition of AI. Computer scientist Judea Pearl's book from 2000 (*Causality: Models, reasoning, and inference*, Cambridge University Press, 2013) and his more recent *The Book of Why* (Basic Books, 2018) explore a collection of techniques that can be used in conjunction with various MI techniques to extract causal connections in contrast to just identifying associations. It is important to note that associations arise from almost all MI ranging from conventional techniques to the newer AI and ML algorithms.

Figure 13
Levels of causality needed in insurance



Source: Swiss Re Institute

… and can thereby improve the predictive power of MI systems.

Causal inference arises from hypothesising causal relationships with a range of drivers in the context of directed acyclic graphs (DAGs) based on the best available scientific understanding.[28] Then, different techniques can be used to “prune” the graph to distinguish causal drivers from confounders. Combined with other MI techniques, causal inference can be a powerful tool to improve the predictive power of particular models and feed into more robust risk-management systems. This combination follows nicely from hybridising physics-based models with newer MI approaches as scientific theories provide guidance as to which variable inter-relationships should be targeted in training, fitting, or estimating relevant MI models.

Using visualisation to generate actionable insight

Data visualisation tools help non- data scientists understand the output from MI-enabled systems.

Even if an enterprise successfully implements an MI-enabled system, the output often remains restricted to discussion among the firm's data scientists. This limits the system's influence. Inflexible decision-making processes and immature software make it hard for data scientists to transform system output into actionable insights that decision-makers can use. In a recent survey, more than 70% of US insurers said they were concerned that non-data science staff did not understand AI and ML

28 The directed acyclic graph causal framework allows for the representation of causal and counterfactual relations amongst variables.

outcomes.[29] Insurers will need to develop more customised visualisation and decision-support tools that work for their specific needs. One seemingly counter intuitive recommendation is to include non-technical designers as part of the MI-enabled system deployment team. These non-technicians should work with executive decision makers from project beginning, not end. Many powerful systems are not productively used because the output is confusing to decision makers.

Off-the-shelf business intelligence tools now allow for custom visuals that can be used to communicate model findings.

Many new features were added to visualisation tools in 2019, based on well-known JavaScript visualisation libraries such as D3, jQuery and R. Gartner predicts that by 2022, 40% of ML model development and scoring will take place in tools (eg, business intelligence (BI tools) that do not have ML as their primary goal.[30] For instance, Microsoft has made it possible to integrate Python scripts within PowerBI, its popular BI tool.[31] AutoML is already available in visualisation and BI tools and currently supports classification and regression models.[32] There will likely be additional model types in the future, and the ability to export ML models to interactive computing environments like Jupyter Notebooks, facilitating model refinement "on the fly."

Progress in model explainability and interpretation techniques

Today there is more emphasis on explainable AI and ML, especially for techniques with higher accuracy.

As newer MI tools demonstrate productive potential in the enterprise context, more emphasis is placed on "Explainable AI and ML". That is, algorithms with higher accuracy levels (relevant for specific business use cases) need more explanation before they will be acceptable across a broader range of business contexts. Successful enterprise-wide and decision-critical system deployments require explainability and interpretability. The past years have seen progress in explaining complex models, such as SHAP (Shapley Additive exPlanations) values and Local Interpretable Model-Agnostic Explanations (LIME). Optimal classification trees are also being proposed to improve accuracy while dealing with the problem of interpretability (see *Case study: Optimal Classification Trees*). Whether these approaches are sufficient for regulators and internal governance units at insurers is still unclear.

Case study: Optimal Classification Trees

Decision Trees are interpretable but may have lower accuracy.

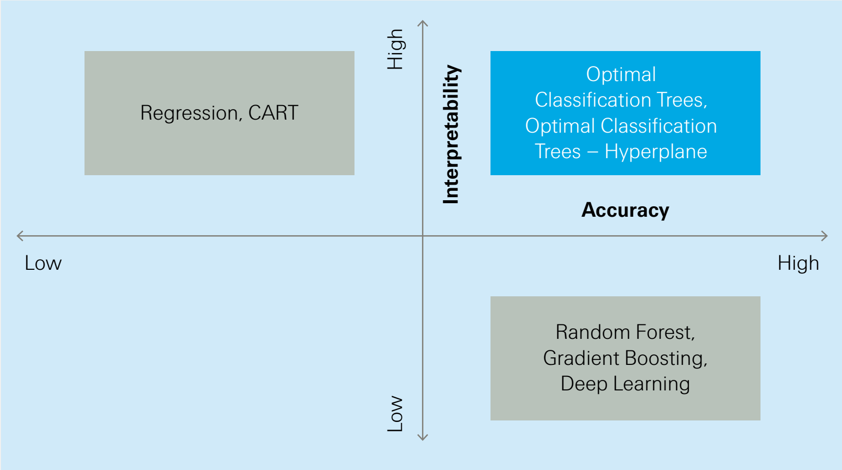
Decision trees are highly interpretable and explainable to a non-technical audience. However, such models may lack stability. A slight change in data can cause a large change to tree structure, making them less appropriate for regulated areas like insurance pricing. Another shortcoming is that every split in the tree is decided on a standalone basis without considering the possible impact of future splits in the tree. This can lead to trees that do not adequately capture underlying characteristics of data sets, potentially leading to weak performance when classifying future data.

Optimal classification trees improve accuracy, while maintaining interpretability.

A helpful solution associated with a top-down approach is to create the tree in a single step (ie, jointly decide all tree nodes). Each split is therefore determined with complete information of all other tree splits. In 2017, Bertsimas and Dunn proposed a technique called optimal classification trees to improve decision-tree accuracy.[33] This technique uses mixed-integer programming (MIP) to learn optimal classification trees. MIP comes with a suite of off-the-shelf solvers and algorithms that can be leveraged to effectively prune-out the search space.

[29] LexisNexis® Risk Solutions, op. cit.
[30] *Gartner Magic Quadrant for Analytics and Business Intelligence Platforms*, Gartner, February 2020.
[31] "A Tour of Artificial Intelligence Features in Power BI", *blue-granite.com*, 5 December 2019.
[32] AutoML is a ML capability that enables developers with limited ML expertise to train models specific to their business needs.
[33] D Bertimas, J. Dunn, "Optimal Classification Trees", *Machine Learning*, vol 106, July 2017.

Figure 14
How techniques map with regard to interpretability and accuracy.



Source: D. Bertsimas,J. Dunn, *Machine Learning under a Modern Optimization Lens*, Dynamic Ideas, 2019

It is now possible to model prototyping in the same programming language used in deployment.

Intelligent automation: re-designing workflows

The data science vendor space has matured to cater to both expert and citizen data scientists to build, train, deploy and manage analytical models.[34] MI techniques are increasingly used to simplify analytical processes such as data preparation, insight discovery and insight sharing. Newer AI and ML techniques still face the challenge that language program prototypes cannot scale at an enterprise level, but we expect that new developments could help overcome such obstacles. Model prototyping will be possible in the same AI and ML-oriented languages that are used for industrial grade deployment. For example, Amazon Sagemaker recently announced an open source library and API to prototype deep learning models in Java.[35] Internal engineers now expect to save 30% in development time.[36]

Better tools can augment existing data curation workflows.

Inadequate data curation workflows continue to materially hamper successful deployment of enterprise MI-enabled systems. Fortunately, better tools are becoming available to improve existing data curation. These platforms augment collecting, labelling and feeding data into supervised learning models and standardised workflows. More sophisticated libraries and software packages allow models that are better able to generalise, meaning that a wider set of problems can be solved (eg, Tensor Flow for ML models). Even with better tool availability, the feedback is mixed: some platforms facilitate seamless integration across diverse tools, while others still struggle with a plethora of tools that do not necessarily work together.

Privacy-preserving analytics

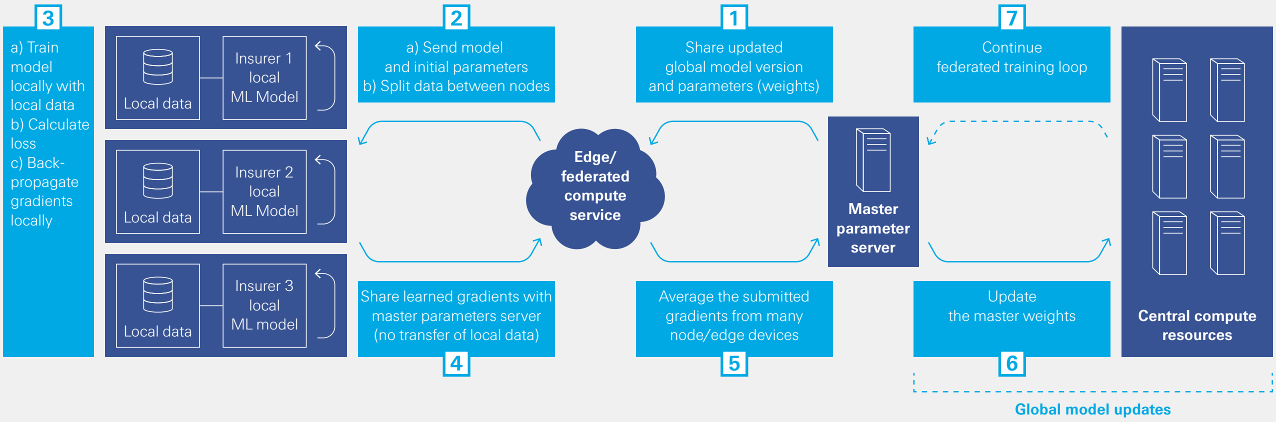
New protocols offer better data privacy protection than standard anonymisation techniques.

Given that MI-enabled system performance is often boosted with more data, industry players would benefit if they were to share data. That said, standard anonymisation protocols are not secure enough. New protocols are creating new opportunities. Secure multi-party contribution protocols can unlock derived analytics from non-public data across multiple insurance companies. These new protocols facilitate a higher level of data privacy protection beyond what is typical for standard anonymisation techniques. In this way, a consortium of insurers can contribute data to generate derived analytics for MI applications that would benefit all contributors (see Figure 15). The differential privacy techniques eliminate the possibility that even

[34] *Solution Criteria for Data Science and Machine Learning Platforms*, Gartner, 6 September 2019.
[35] *Introducing Deep Java Library: Develop and deploy Machine Learning models in Java*, Amazon Web Services, 3 December 2019.
[36] S. Sivasubramanian, *Leadership session: Machine learning*, Amazon AI Amazon Web Services, December 2019.

the researchers and data engineers working on the pooled data can look back into individual contributions while still facilitating more sophisticated MI.

Figure 15
Federated learning to achieve privacy preserving analytics



Source: Swiss Re Institute

Master models leverage local data at each insurer to help them learn from each other without sharing data.

Figure 14 shows how once the model has been instantiated, the parameters and weights are pushed out (in steps 1 and 2) through a web service to individual insurers, which each run the global model on their local data (eg, claims records), to find out how accurate the master model is (step 3). After this is completed, each insurer (step 4) offers feedback and shares the findings or learned gradients (ie, what is different between the two models). This feedback is combined across insurers, and updated weights are submitted to the federated services, and reflected in the global model (steps 5 and 6). The cycle can continue until a certain level of accuracy is obtained in the master model.

Trust, technology, talent and tenacity

Often organisational constraints are the reason for failure in the deployment of MI-enabled systems.

Model-related problems are not the only reason for failure in deployment of MI-enabled systems at enterprise scale. More often, organisational constraints such as poor use case planning, lack of properly trained staff and poor communication are the sources of failure. In the insurance sector, a change in mindset could help. Insurers need to better understand the value that MI-enabled-systems can deliver from an end-to-end, enterprise perspective. Small-scale pilot projects for emerging technologies make sense as part of an initial R&D project or targeted assessments. A tendency among insurers thereafter has been to launch enterprise-wide deployment of the pilot, without due consideration of other non-technology design-related issues.

Insurers should also focus on non-model issues holding back wider adoption.

To successfully transform their enterprises with MI-enabled technology, we recommend that insurers stop relying on proof-of-concept, small-scale pilots of model/algorithm approaches. They also need to focus on the salient, non-model characteristics of end-to-end enterprise deployment: trust, technology, talent and tenacity (see Table 4).

Table 4
Non-model considerations for enterprise-scale of MI-enabled systems

Key findings	Implications for the current model	Outlook
Trust: Develop an algorithmic risk and digital ethics framework	Better equip MI-enabled systems against risks, eg, adversarial attacks	Balance different definitions of fairness and incorporate self-monitoring into MI-enabled systems from the design phase
Technology: Balance internal versus external expertise	Understand how procuring MI differs from traditional software to reduce risks and maximise ROI	Develop approaches to harmonize fragmented technologies.
Talent: Develop talent and skills	Identify how MI can complement current actuarial-science- based approaches	Encourage all staff to learn new MI-related tools and leverage citizen data scientists
Tenacity: Foster a dynamic tech-informed culture; engage with regulators	Use sandbox approaches to test MI at scale	Educate regulators. Keep humans in the loop

Source: Swiss Re Institute

Trust: develop an algorithmic risk and digital ethics framework

Insurers must gain deeper understanding of the consequences MI may have on the services they provide.

Life-altering decisions can be automated via algorithms, and embedded biases within algorithms may often be inadequately monitored and documented. This can result in liability for companies using decision-support algorithms that incorporate bias (in most cases, unintentionally) should victims choose to litigate. Seven out of 10 US carriers are already concerned about bias in ML models.[37] Even if an MI-enabled-system outcome is solely or mostly responsible for undesired consequences, "the algorithm did it" is not an acceptable excuse. In a survey carried out in 2019, nearly 50% of firms using MI solutions across sectors said they have a formalised framework to consider ethical use, bias risks, and trust implications; 25% had created a senior management position specifically to ensure compliance.[38]

Existing frameworks were not designed to govern behaviour by large-scale algorithmic systems.

There is also scope for automated technology-based solutions that detect bias and generate risk scores for algorithms, which allow insurers to assess the malpractice risk of specific algorithms. Insurance solutions can be considered to protect companies using such algorithms against liabilities resulting from embedded bias. Insurers should have a stronger voice in the societal debate about questions of fairness in algorithmic decisions and join forces with researchers to address these issues (eg, 'FAT machine learning community[39]). In the last decade, academics have published several definitions of fairness, not all of which can be achieved at the same time. Since creating a generalised state of fairness is not feasible, insurers may need to choose which conditions to keep and which to discard.

Technology: balance internal versus external expertise

With increasing commoditisation of MI categories, insurers need a detailed MI procurement and knowledge-transfer strategy.

IT support for MI will be especially challenging as technology teams try to manage the balance between: (1) running the business in the face of increasing requests for various IT services, along with; (2) innovation and research. More than half (59%) of CIOs and IT decision makers surveyed recently were unable to deliver on all their projects in 2019, creating a backlog for 2020.[40] As the range of MI-related offerings continues to grow, IT units will need to modify procurement approaches designed for buying traditional software to reflect MI procurement. For example, insurers may need to restrict agreement terms to shorter period (eg, no more than three years to protect from lock-in.[41]) There will need to be more emphasis on tool/system flexibility

[37] LexisNexis® Risk Solutions, op. cit.
[38] Global survey of 2 473 firms that use AI solutions. *IDC Survey Finds Artificial Intelligence to be a Priority for Organizations But Few Have Implemented an Enterprise-Wide Strategy*, IDC, 8 July 2019.
[39] *Fairness, Accountability, and Transparency in Machine Learning*, FAT/ML, see https://www.fatml.org
[40] *New report shows 3 out of 4 organizations expect negative revenue impact if they don't digitally transform in next 12 months*, MuleSoft, 13 February 2020.
[41] *Lack of Focus on AI Licensing Will Result in Higher Costs, Risks and Long-Term Headaches*, Gartner, 11 September 2019.

and interoperability. Insurers will also need to create and execute knowledge-transfer plans to ensure continuity between external providers and their own staff, both in IT and the business.[42]

Talent: develop talent and skills

Attracting and retaining people with MI skills remains a major challenge for insurers.

Lack of sufficient staff to analyse data is among the three biggest challenges preventing insurers from becoming more data-driven, according to a recent survey.[43] Awareness of MI is growing among actuaries. MI is endorsed at prominent actuarial conventions, with papers on how actuarial science can incorporate deep learning in areas such as mortality modelling, claims reserving, telematics analysis and non-life pricing. Still, retaining MI talent remains a challenge. Insurers invest in skills development programmes for employees, but many struggle to create near-term opportunities and incentives to apply MI in a way that interests skilled MI developers and data scientists. About one-sixth of respondents in a survey cited difficulty in hiring and retaining people with AI skills as a significant barrier to broader AI adoption in their organisation.[44]

Tenacity: a dynamic, tech-informed culture and engage with regulators

Involving the business and executives throughout the MI development lifecycle is key.

Beyond investing in foundational MI-enabled capabilities, insurers must focus on high-level business workflows and opportunities productively transformed by these new technologies. In recent years, many insurers have funded proofs of concept and pilots in the MI space. These efforts provide preliminary guidance but do not transform business. Going forward, key project components for making MI-enabled systems productively transformative go well beyond the technology. They include enterprise technology architecture design, business workflow re-engineering, co-creation with executives data on visualisation, and extensive change management programs. Having business people involved throughout the identification, testing, evaluation, and implementation process is key to achieving success.

Some issues require more discussion, such as complying with regulatory requirements.

Regulatory risks regarding tech-linked innovation in insurance present challenging hurdles. The risks mostly centre on questions of data management and use. The General Data Protection Regulation (GDPR) in Europe emphasises important questions for managing data privacy, which is particularly relevant for MI-enabled systems, which often merge and mix different data sources for risk assessment. Some issues require more development and discussion such as complying with GDPR principles focused on "use for legitimate purposes only" and conditions for use in "high risk" cases (eg, medical and health, profiling).

Sandbox approaches could help overcome barriers to adoption of enterprise-scale MI in insurance.

Further, restrictions on cross-border data transfers can also impede development and application of cross-border solutions, and slow regulatory approval of new tech components like cloud solutions. Given the complicated and subtle nature of many MI-enabled solutions, inadequate understanding of MI possibilities and drawbacks could slow industry adoption. More sandbox efforts – particularly experiments at enterprise scale – are required to overcome regulatory barriers and foster a deeper understanding related to data privacy management and MI capabilities among regulators and insurance executives.

[42] Gartner, op. cit. 9 February 2018.
[43] Willis Towers Watson, 25 February 2020, op. cit.
[44] *AI adoption in the enterprise 2020*, O'Reilly, 18 March 2020.

Conclusion

Insurers must shift focus from technology development to enterprise transformation to realise the potential of MI-enabled systems.

Despite significant advances in MI-enabled image recognition and customer analytics for example, productive, enterprise-scale transformation based on MI-enabled systems in the insurance sector has proven elusive. Some trends have a long arc and will most likely continue current trajectories such as integrating computer vision into underwriting systems; other shorter-term trends like semi-automating fragmented data curation systems could change quite quickly.

Investments in data collection and curation capabilities will be a key differentiating factor.

Data have become paramount in any strategy to fully exploit the potential of MI in insurance. While longer time series of structured data and efforts to find novel data continue to be an important component of this narrative, unstructured data (eg, text, audio, and video) have become a new opportunity not yet fully exploited. Incumbents with proper tools and organisation will differentiate themselves as better curated and novel data become a component of their competitive edge.

Figuring out which specific tools are realistic and deserve investment is also critical.

Newer MI in the AI and ML spaces are among the over-hyped technology areas that have yet to be implemented in a materially profitable and transformative way within the insurance value chain. For example, chatbots powered with the best in natural language processing are still rolled out as the solution to confusing menu systems and as a tool to reduce the size of call centres. The predictions for customer support transformation were wildly unrealistic. This said, the collection of tools available to insurers will continue to evolve. Much work remains in determining the specific tools to use in these spaces.

Insurers should accept that project-completion timelines will be longer than many expect.

An important consideration is the difference between targeted proof-of-concept value and successful enterprise deployment. Insurers and their technology partners will benefit from more investment and experimentation with MI at the enterprise scale. Many insurers already experiment with MI in narrow contexts. The failures to date result from the inability to scale profitably these narrowly focused experimental projects. Everyone involved in enterprise MI deployment should accept that project-completion timelines will continue to be much longer than most executives expect.

Effective MI deployment will rely on a range of factors including cultural and regional attitudes towards privacy and regulation.

Regulatory compliance will continue to be a critical component of any strategy to leverage data and digital tools. One area in this context that will be particularly onerous for any firm expanding its use of data, particularly in the area of personalisation and customisation, is data privacy. New regulations will continue to come at a fast and furious pace, furthering the advantage of large insurers already equipped to manage compliance. Cultural norms, attitudes with respect to data privacy, and regulation differ substantially across regions. Multi-national insurers that robustly address this heterogeneous, and often fragmented regulatory landscape in their MI implementations will differentiate themselves from competitors.

Building "trust" will encompass how data are managed, and how customer needs are met.

MI-enabled systems have profitably transformed other industries. This promise continues to drive MI-related investments in the insurance industry. Executives, technology architects, project managers, and analysts must shift their focus from technology development to enterprise transformation to realise this business-value potential. Key success factors include building trust with clients and regulators, implementing enterprise-oriented technology, fostering cultures that retain suitable MI-trained talent, and engendering hurdle-clearing tenacity.

Recent *sigma* publications

2020
- No 1 Data-driven insurance: ready for the next frontier?
- No 2 Natural catastrophes in times of economic accumulation and climate change
- No 3 Power up: investing in infrastructure to drive sustainable growth in emerging markets
- No 4 World insurance: riding out the 2020 pandemic storm
- No 5 Machine intelligence in insurance: insights for end-to-end enterprise transformation

2019
- No 1 Emerging markets: the silver lining amid a challenging outlook
- No 2 Natural catastrophes and man-made disasters in 2018: "secondary" perils on the frontline
- No 3 World insurance: the great pivot east continues
- No 4 Advanced analytics: unlocking new frontiers in P&C insurance
- No 5 Indexing resilience: a primer for insurance markets and economies
- No 6 Global economic and insurance outlook 2020/21

2018
- No 1 Natural catastrophes and man-made disasters in 2017: a year of record-breaking losses
- No 2 Constructing the future: recent developments in engineering insurance
- No 3 World insurance in 2017: solid, but mature life markets weigh on growth
- No 4 Profitability in non-life insurance: mind the gap
- No 5 Global economic and insurance outlook 2020
- No 6 Mortality improvement: understanding the past and framing the future

2017
- No 1 Cyber: getting to grips with a complex risk
- No 2 Natural catastrophes and man-made disasters in 2016: a year of widespread damages
- No 3 World insurance in 2016: the China growth engine steams ahead
- No 4 Insurance: adding value to development in emerging markets
- No 5 Commercial insurance: expanding the scope of insurability
- No 6 Life in-force management: improving consumer value and long-term profitability

2016
- No 1 Natural catastrophes and man-made disasters in 2015: Asia suffers substantial losses
- No 2 Insuring the frontier markets
- No 3 World insurance 2015: steady growth amid regional disparities
- No 4 Mutual insurance in the 21st century: back to the future?
- No 5 Strategic reinsurance and insurance: the increasing trend of customised solutions

2015
- No 1 Keeping healthy in emerging markets: insurance can help
- No 2 Natural catastrophes and man-made disasters in 2014: convective and winter stormss generate most losses
- No 3 M&A in insurance: start of a new wave?
- No 4 World insurance in 2014: back to life
- No 5 Underinsurance of property risks: closing the gap
- No 6 Life insurance in the digital age: fundamental transformation ahead

2014
- No 1 Natural catastrophes and man-made disasters in 2013: large losses from floods and hail; Haiyan hits the Philippines
- No 2 Digital distribution in insurance: a quiet revolution
- No 3 World insurance in 2013: steering towards recovery
- No 4 Liability claims trends: emerging risks and rebounding economic drivers
- No 5 How will we care? Finding sustainable long-term care solutions for an ageing world

2013
- No 1 Partnering for food security in emerging markets
- No 2 Natural catastrophes and man-made disasters in 2012: A year of extreme weather events in the US
- No 3 World insurance 2012: Progressing on the long and winding road to recovery
- No 4 Navigating recent developments in marine and airline insurance
- No 5 Urbanisation in emerging markets: boon and bane for insurers
- No 6 Life insurance: focusing on the consumer

Published by
Swiss Re Institute
Swiss Re Management Ltd
Mythenquai 50/60
P.O. Box
8022 Zurich
Switzerland

Telephone +41 43 285 2551
Email institute@swissre.com

Authors
Jonathan Anchen
Dr Jeffrey Bohn
Rajeev Sharan

***sigma* editor**
Paul Ronke

Managing editors
Dr Jeffrey Bohn
Chief Research and Innovation Officer

Dr Jerome Jean Haegeli
Swiss Re Group Chief Economist

The authors thank Evangelos Avramakis, Luca Baldassarre, Binay Biswal, Aakash Kiran Raverkar and Jürg Schelldorfer, and also Julia Brandenberg, Katherine Chen, Charlie Dang, Ashish Dave, Mustafa Dinani, Yannick Even, Nate Jensen, Nuno Mesquita, Sudipto Pal, Ashok Shetty, Charilaos Tsarouchas, Francesca Volpe, Guan Wang, Kelvyn Young, colleagues from Swiss Re Client Management and Solutions, and Gianluca Antonini from IBM, for reviewing and providing ideas to improve this publication.

Explore and visualize *sigma* data on natural catastrophes and the world insurance markets at *www.sigma-explorer.com*

The editorial deadline for this study was 30 June 2020.

sigma is available in English original language and Chinese.

sigma is available on Swiss Re's website: www.swissre.com/sigma

The internet version may contain slightly updated information.

Translations:
Spanish: Traductores Asociados Valencia S.L.

Graphic design and production:
Corporate Real Estate & Logistics / Media Production, Zurich

MIX
Paper from responsible sources
FSC® C022369
FSC www.fsc.org

Printing: Multicolor Print AG, Baar

Order no: 270_0520_EN

Swiss Re Management Ltd.
Swiss Re Institute
Mythenquai 50/60
P.O. Box
8022 Zurich
Switzerland

Telephone + 41 43 285 2551
swissre.com/institute