DeRisk

CRISIL's insights on and analyses of regulations, macroeconomic factors, guidance and trends affecting the insurance industry

September 2021

AI+ML, the new cornerstone of modelling

Global Research & Risk Solutions

CRISIL
An S&P Global Company

Exponential growth beckons

Insurers possess, and have access to, vast historical – and ever-growing – datasets, which make them prime candidates for the adoption of artificial intelligence (AI)[1] and machine learning (ML)[2] for risk modelling.

That potentiality, however, did not manifest until the Covid-19 pandemic pushed insurers to scale up data mining. This trend is only expected to continue, and accelerate.

While global demand for AI/ML is expected to increase ~40% annually through this decade[3], there is a proxy signal for incipient insurer-specific interest: searches for the term 'Insurance AI' on Google Trends have shot up 400% over the past five years.

Practitioners are looking at other informal ways to learn; sites like Kaggle have 2200+ AI/ML competitions and thousands of data sets available for different use cases around data science (DS) and guidance-notebooks for Python, R and Julia languages, as of August 2021. Learning sites like LinkedIn have increased their AI/ML course offerings to 110+ compared with 30-50% less courses 18 months ago when the pandemic started, clearly showing the ongoing interest in this area. Other learning sites like Udemy, Google, Coursera, Microsoft EdX and Udacity too, have increased similar course offerings.

What are insurers doing with AI/ML?

Currently, insurers are learning about where AI/ML could be best used across product types, as well as the associated risks.

Some of the most valuable use-cases that adopt AI/ML models across categories (personal and commercial lines) are those focused on consumers, such as underwriting, pricing-actuarial, risk exposure, smart/digital analysis of documents and automated claims. Others include fraud, know-your-customer, cyber risk and loss control.

Insurers are also trying to gauge how AI/ML model risks differ from the traditional models because of three developments: [4]

- Vast increase in the amount of available data
- Increasing ability of software to find patterns in data, and continuously 'training' models using updated data, and
- Improvements in computational power and storage, allowing for broader rollout of fast, complex models

Who are the early adopters?

Non-life insurers have seen the greatest adoption, given their short-term contracts (less than one year). That's because such a duration of contract enables them to reflect experience into premium cost, i.e., modify the premium every year. This is crucial to the underwriting process, which aims to reflect consumer behaviour, and risk frequency/severity (e.g., pay-as-you-drive and climate-change insurance products).

In contrast, the life insurance business has a slower adoption rate as use-cases are still being identified.

What are the gains and risks?

AI/ML models have significant performance advantages over traditional regression and time-series methods. The key among them is the ability to identify complex, non-linear relationships from among a large numbers of input variables, allowing for more accurate predictions. However, ML models are a double-edged sword – they are often called 'black boxes,' due to the large numbers of inputs and complex inter-relationships between variables which they are able to identify. Their inner workings are opaque, at best.

Then there are differences in model governance between the new-age and traditional models, such as validation and monitoring. Identification of correct use-cases is another challenge, as is the requirement of new talent on DS and knowledge management.

---

[1] AI: the logical conclusion of machine learning – that computers should be able to improve and make independent decisions without any sort of human input or manual 'overlays',
[2] ML: computer algorithms and models which have (a degree of) built-in, automatic improvement based on new data and statistical methods,
[3] https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market
[4] Regulations.gov

This issue of DeRisk focusses on how insurers are using AI/ML models today. It details emerging risks even as insurers gain competitive advantage by adopting them in their value chains. It examines the impact of and gaps in regulation. It concludes by setting out the key learnings unfolding as newer use-cases are put to test and regulatory guidance takes shape.

Where do insurers stand?

Typically, 60-70% of the models at an insurance company are developed by the actuarial function[5]. While this distribution varies among life, general and health insurers, the overall model inventory dynamics has remained constant.

But a perceptible shift has occurred over the past 2-3 years due to multiple factors, including a talent relocation from the banking industry to insurance. New capabilities' focus, such as DS and Machine Learning Engineers (MLE), and the consequent rise of consumer-centricity, digitisation and automation, are ratcheting up AI/ML model adoption in insurance.

The potential of such models to improve profitability is high – through more efficient underwriting-pricing-claim processes, and faster response to policyholder queries such as appeals, questions or personalised services[6].

As insurers embed AI/ML deeper into their value chains, model risks related to governance, business efficiencies and DS capabilities, among others (see Exhibit 1 below), will be spawned.

Exhibit 1: Challenges and current thinking among insurers on AI/ML models

Topic	Issue	How are insurers tackling them?
AI/ML governance	Bias, limits, and interpretability Validation and monitoring	• Identifying and mitigating negative implications around bias, drift, variance, and interpretability for models • Upgrading model-risk management (MRM) policies, practices, and frameworks to incorporate differences between traditional and AI/ML models
Business efficiencies	Adoption of AI/ML in specific use-cases Tools and partnerships	• Actively using AI/ML models, e.g. pricing, mortality, agency, marketing, customer behavioural, claims • Increasing efforts in adopting technology through developing, procuring and/or partnering with vendors and/or insurtechs
DS and ML engineering capabilities	Talent acquisition Knowledge management	• Developing and acquiring talent such as data scientists and MLEs to support AI/ML models, allocating and monitoring functions, contributions and efficiencies • Managing and upscaling talent to get into the practical aspects of business requirements through sharing and developing a knowledge culture

[5] Based on CRISIL's experience with multiple insurance clients
[6] https://research.aimultiple.com/ai-insurance/

The three sites of AI/ML risks

Traditional models, in use since the inception of the insurance industry, have well-known risks and materiality exposure at systemic, firm and consumer levels. But as insurers adopt AI/ML models, these risk levels may shift, with unknown or different impacts.

For example, clusters of financial institutions using similar AI solutions at the same time may cause a systemic risk. Specific insurers making wrong assessments and decisions due to model interpretability issues could trigger firm risk. Failure to properly capture the behavioural dynamics of consumers such as biases[7] could entail consumer risk (see Exhibit 2).

Exhibit 2: Key risks in AI/ML adoption in the MRM function of insurers

Systemic risks	Firm risks	Consumer risks
New entrants • Insurtech creating a systemic risk bubble/ waterfall due to low regulation around it **Unintentional errors** • Insurers making large mistakes unintentionally due to the use of AI/ML solutions in strategic decision-making areas • Dynamic interaction between AI/ML models leading to unforeseen circumstances or market crashes	**Regulation** • Late/unknown regulation • Dealing with 'too strict' or 'too lax' regulation • Unclear regulatory standards around acceptable practices, auditability, etc. **Incorrect decisions** • Interpretability/explainability • Bias and overfitting • Model drift • Disparate impact	**Ethics** • Bias inducing possible discriminatory practices that may reflect in classifying policies in wrong selection groups **Pricing** • Incorrect risk assessment due to wrong model use or model interpretability

Systemic risks

While is difficult to fully anticipate such risks, regulators are gradually moving from their 'hands-off' or 'wait-and-see'[8] approach to a more active role. They are organising AI/ML forums that involve financial institutions (banks and insurers), aiming to lay out clear regulatory initiatives[9] to the benefit of the industry and consumers alike. While financial crises in the past have not been triggered directly by the insurance industry, such events have nonetheless had an increasingly visible impact on it, primarily through investment portfolios[10].

But what concerns our insurance clients is how new market entrants, from fintech to insurtech, are not being held to the same regulatory standards. By definition, they are not insurers that need to follow solvency or risk-based capital requirements. Traditional insurers possess enormous datasets, but are constrained by their size, culture, legacy systems and regulatory/statutory requirements. Comparatively, fintech and insurtechs are considered flexible and agile, given that they are less regulated. This increases the possibility of a new version of the unintentional 'move fast and break things' (seen in the 90s in the tech industry), but in today's AI/ML era at the insurance industry.

---

[7] *For example, demographic blindness, i.e. use features uncorrelated with protected classes*
[8] *Given regulators worldwide have not adopted regulations to address many of the ethical and transparency issues emanating by the AI/ML models.*
[9] *This systemic, firm and consumer taxonomy is a reflection of the sentiment expressed by banks and insurers in the United Kingdom around AI/ML model adoption risks according to the recent discussion table organised by the Bank of England in July 2021. https://www.bankofengland.co.uk/minutes/2021/june/aippf-minutes-15-june-2021*
[10] *https://www.researchgate.net/publication/227461180_Insurance_companies_and_the_financial_crisis*

To be sure, regulators are attempting to close some of these gaps. Along with insurers, they may begin to differentiate between intentional and unintentional risks posed by AI/ML models. We expect that intentional risk – the business impact a model makes – will be particularly scrutinised[11].

However, unintentional risk, such as interactions between high-frequency traders using computerised algorithms (investment risk), will remain elusive and difficult to manage, creating unforeseen systemic risks such as disappearing liquidity, similar to flash crashes[12] of the past. Other unknown risks may arise from the use of AI/ML models in other key areas of the insurance industry that could affect consumers.

Firm-level risks

At the firm level, insurers are especially concerned about regulatory risks from the use of AI/ML models. The hands-off regulatory approach seen to-date has left both unanswered questions and gaps in interpretation. For instance, vague definitions surround AI; insurers would prefer to see regulators provide context by focusing on distinctive technical aspects of AI/ML complexity, such as hyperparameters[13], material knowledge, skills, resourcing, sponsorship, and ownership[14].

What would be a sound regulatory approach? On a scale of 'too lax' to 'too strict,' finding middle ground may not be easy. Existing regulations around MRM, while insufficient for insurers, provide useful starting points. On the other hand, heavy-handed regulation could play into startups capturing modelling innovation (possibly creating a "shadow" sector of less-monitored competitors), or even permanently relocate innovation to friendlier jurisdictions. The industry would like to see regulators pay heed to these nuances.

The other major risks are around data. While data quality is the bedrock of models (and something most institutions already consider), other aspects such as traceability, auditability, accountability, and data attribution are no less significant. But these use-cases come with modelling challenges, such as:

- Interpretability: The output of many models, particularly those that best identify complex relationships between variables, may defy straightforward explanation. This is a frequent topic in technical symposia where the use of non-traditional variables is often discussed
- Overfitting: Models may identify spurious relationships that could change in the next model monitoring review or economic cycle – or even tomorrow. Also, model predictions may degrade in quality, particularly as forecasts are made further out.
- Disparate impact: With numerous variables, it is possible for models to unintentionally fit proxies for protected data categories, such as race, sex, or age[15]. Raising to a similar discriminatory problem in biased at the banks in their credit scoring[16].

Consumer risks

Broadly, these risks fall under risk identification and pricing and ethics. Various guidelines around AI/ML and its ethics can impact consumer-related risks (see Exhibit 3 for details).

Earlier academic[17] work has addressed these issues as well, including the 'The five Cs': consent, clarity, consistency and trust, control and transparency, and consequences.

In the world of pricing-underwriting and risk identification-selection, AI/ML promises new methods to identify potential customers from previously unusable data, e.g. consumer behavioural variables or patterns, including credit rating or other scoring criteria. Countries with less-established credit rating systems are seeing startups use AI/ML models to develop solutions using social media, bank records, and geospatial data to help with the value chain of the insurers.

Risk identification and consumer segmentation will need to vary based on the application, in direct proportion to the models' visibility to the customers and/or potential market impact.

---

[11] *https://www.bankofengland.co.uk/minutes/2021/june/aippf-minutes-15-june-2021*
[12] *https://www.investopedia.com/terms/f/flash-crash.asp*
[13] *AI/ML models are sensitive to the hyperparameters – parameters which determine how the data is fitted by the model. Selecting these is a mixture of art, systematic testing, and expert judgment.*
[14] *https://www.bankofengland.co.uk/minutes/2021/march/aippf-minutes-26-february-2021*
[15] *https://insight.equifax.com/qa-machine-learning-explainable-ai-in-credit-risk/*
[16] *https://medium.com/@kguruswamy_37814/mitigating-bias-in-ai-ml-models-with-disparate-impact-analysis-9920212ee01c*
[17] *Patil, D., Mason, H., Loukides, M.: Ethics and Data Science. O'Reilly Media, Inc. (2018)*

This will raise ethical issues. For example, EIOPA[18] stresses on proportionality, fairness and non-discrimination, transparency and explainability[19], human oversight, data governance of record keeping, robustness and performance. The OECD[20] principles rest on fairness and ethics, accountability, compliance, transparency, and security/safety/robustness.

The trade-off between risk identification and pricing-underwriting brings nuances and complexity. Customers face the risk of having their applications for insurance rejected (or premiums price increased) for what may seem arbitrary factors. Models should bring transparency to mitigate these perceptions from applicants and future policyholders. Models begin with larger, clean, accurate data, continue with processes and decision points, and end with clear actionable recommendations. Similarly, any regulation is likely to follow such a 'funnel' approach – start with data quality, continue with a review of model explainability and end with a review of disparate impact.

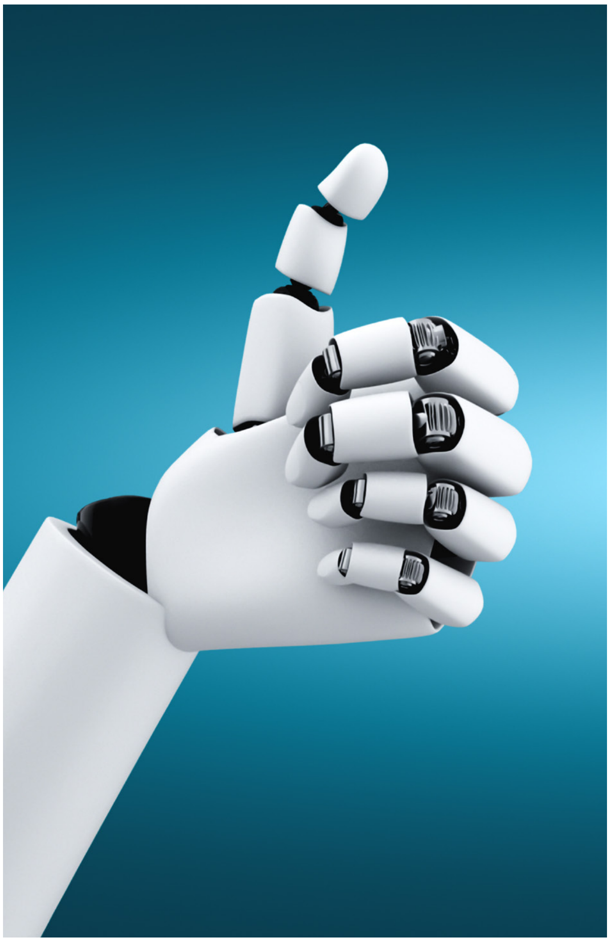
This would bring newer issues around bias, inducing possible discriminatory practices that may reflect in unfavourable selection groups for policies and pricing or incorrect risk assessment due to model use or interpretability challenges due to low transparency.

The formative language of AI/ML regulation

As we've seen, regulation ties in with all manner of risk. Guidance regarding model validation and documentation will continue to evolve, with regulators continuously reflecting on industry case studies and advice[21]. At this stage, it is not clear on which side of the stringency-leniency spectrum future AI/ML regulations will aim for. But it could be said that the level of restriction will determine the potential impact on the systemic, firm, and consumer risks. For an illustration, the banning[22] of financial institutions to track or score consumer behavioral aspects would lead to a decrement in model performance, if new AI/ML models are not approved for use.

Further, industry lobbying groups, such as the US-based Bank Policy Institute, express the need for consistent regulatory treatment for banks and nonbanks in financial services.[23] As things stand, large organisations, though masters of massive datasets and vast resources, are also weighed down by legacy systems and higher regulatory asks. That creates opportunities for fintech and insurtech startups and other niche players to enter the market and avail of regulatory arbitrage.

Exhibit 3 below lists key regulations or guidelines in various jurisdictions where insurance companies have large presence, and that affect AI/ML modelling. While these are still at formative stage, it is clear they are aimed at managing different risks.



[18] The European Insurance and Occupational Pensions Authority
[19] A model needs to be intuitively understandable in order for it to be trustworthy. Reasonable to differentiate required level of interpretability based on model's impact or risk The most recent booklet (August 2021) from the OCC contains multiple references around explainability. https://www.occ.treas.gov/publications-and-resources/publications/comptrollers-handbook/files/model-risk-management/index-model-risk-management.html
[20] Organisation for Economic Co-operation and Development
[21] On July 1, 2021, US banking regulators concluded a request for information (RFI) on the use of AI by banking institutions. Analogous consultations are likely to emerge, but focused on the insurance industry as well.
[22] https://www.pymnts.com/news/artificial-intelligence/2021/eu-proposes-restrictive-new-ai-regulations/
[23] https://bpi.com/existing-safeguards-encourage-responsible-ai-innovation-but-neglect-risks-from-nonbanks/

Exhibit 3: Key regulations that affect AI/ML modelling [24]

On data privacy risks	On ethical and governance risks
• Europe - General Data Protection Regulation (GDPR, 2018) • US - Health Insurance Portability and Accountability Act (HIPPA, 1996); California Consumer Privacy Act (CCPA, 2020) or California Consumer Privacy Right Act (in effect until 2023)	• Europe - 1. EIOPA AI governance principles, 'Towards ethical and trustworthy artificial intelligence in the European insurance sector' (2021) - consumer and insurer value chain focus; covers product design and development, pricing and underwriting, sales and distribution, customer service, loss prevention and claims management 2. EC's 'Ethics guidelines for trustworthy AI' • US - 1. NAIC's principles on AI (2020) - relies on the OECD AI principles adopted by 42 countries, including the US 2. OCC Model Risk Management Booklet (2021) • UK - 1. Department of Health and Social Care (2019) 2. FCA's 'The future of regulation: AI for consumer good' (2019) and 'AI-transparency-financial-services-why-what-who-and-when'(2020)

Note: EC- European Commission; NAIC - National Association of Insurers Commissioners; FCA - Financial Conduct Authority

Along with the practical aspects of AI/ML adoption, regulatory trends will also likely include[25]:

- Model-risk assessment: Clearer guidelines on what are considered 'high-risk' models, and tiered levels of controls
- Model documentation: Given the opacity of AI/ML models, regulators are likely to require additional model validation and tiered roll-outs, particularly for high-risk models[26]
- Accountability and data quality: Insurers will be held to higher standards on their data quality and third party data providers will need respond to the call to adopt best practices[27]
- Customer data privacy: Regulations (such as GDPR) will transmit to other markets, affecting client-facing sides of the business, e.g. consumer protection laws at the federal level in the US
- Explainability and bias: This will encourage (or in client-facing applications, mandate) the adoption of methods which would improve transparency and answer the questions of "what are the most important variables?" and "why was this decision made?"
- Better governance drive: Models will be required to have clear thresholds, manual review, executive ownership and accountability and tiered risk thresholds for escalations, in case of high firm and/or systemic risk

[24] Source: https://www.eiopa.europa.eu/sites/default/files/publications/reports/eiopa-ai-governance-principles-june-2021.pdf; https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1; https://www.fca.org.uk/news/speeches/future-regulation-ai-consumer-good
[25] https://wp.nyu.edu/compliance_enforcement/2021/04/27/the-future-of-ai-regulation-the-rfi-on-ai-from-u-s-banking-regulators/
[26] https://www.bankofengland.co.uk/minutes/2021/june/aippf-minutes-15-june-2021
[27] Adopt playbooks and eliminate 'garbage in = garbage out' issues

CRISIL's perspectives and the road ahead

In this section, we sum up CRISIL's main learnings gained from working with various insurance clients, on trends in AI/ML model adoption:

Best practices are evolving: At present, no consensus exists around best practices for AI/ML model validation. But change is inevitable as it is in the interest of all parties. Firms would face lower regulatory risks, regulators would ensure that standards are followed and consumers would gain trust in the technology. Yet, it is unclear just how rigorous such audits may become. Data quality, modelling approach, as well as accuracy, explainability, and fair treatment of customers may all become benchmarks.[28]

Meanwhile, as the insurance market continues to grow, experimentation will continue. The following exhibit captures new AI/ML business use-cases in insurers' value chain and learnings from them.

Exhibit 4: Key learnings around AI/ML adoption

Topic	Insurer's focus and key learnings
Business use-cases	Consumer-focused solutions for agile response, e.g. pricing and chat bots
	Fraud and cyber security models, e.g. identifying unusual behaviour
MRM governance	Clear governance demarcation from traditional models, e.g. policy-level
	More frequent model ongoing monitoring and calibration, e.g. monthly
Data management	Automation of data quality and accesiblity, e.g. service-now platforms in place
	Enhanced approval process for data usage and data quality playbooks
Talent and knowledge management	Expansion of DS and MLE teams
	Developing a collaborative culture and ongoing learning across teams
Partnerships and accelerators	Identification and partnership with firms that can help adopt AI/ML faster
	Use of start-up and AI/ML expert firms in key areas

A lot depends on data quality

Models using machine intelligence are appealing for their novelty and promise of improved results. But they are more reliant on the quality of input data than traditional models. Complexity improves performance but requires larger amounts of data. As AI/ML models pick up complex interactions between variables, more data and variables also increase the risks of introducing bias. Poor data quality can skew outcomes.

Shorter validation cycles, more transparency necessary

As model validators, we are challenged by the complexity and low transparency of AI/ML models. Given that these models adjust to input data, traditional output validation is not effective here. Model validation under AI/ML also requires a shift from 1-3 year cycles to more frequent monitoring, say, cycles no longer than each quarter.

[28] Minutes of the Artificial Intelligence Public-Private Forum - 26 February 2021 | Bank of England

Maintain constant awareness of biases

But many things remain the same as with traditional models. Use existing modeling regulations as a starting point for compliance. Modellers should continue watching out for look-ahead and survivorship biases. Protected characteristics and variables with spurious economic explanations should not be used as inputs. Sharing of protected data across the organisation should be closely reviewed for regulatory violations, such as the use of inside information. For practitioners, improving input data quality often proves more important than fancier models with more variables.

Get the ‘why’ right

Modellers should document their reasons for choosing black box models over more explainable alternatives. This may include higher standards for the documentation of model inputs, and adversarial comparison testing to challenger models. Requirements for documentation may vary depending on the level of visibility afforded by the model used.

Keep standards high, watch for down-the-line risks

Finally, models developed by vendors and other third parties must be held to the same standards as internally-developed models. Wherever model outputs are used by downstream models or interact within a network, the dependencies and increased risks will need to be taken into account. Remember that regulators will pay close attention to dynamic interactions, particularly in segments such as algorithmic trading, where haywire models may lead to systemic risk.

How CRISIL can help you	Details
Planning	• Insightful approach: We identify new business use-cases, redirect investments and help you adopt best-in-class business models towards increasing transparency and steering transformation
Governance	• Streamlining internal policies: Understand and differentiate between traditional and AI/ML models by: (i) requiring updates at each reporting date; (ii) making key-driver assumptions; and (iii) establishing a system to assemble multiple information sources
MRM	• Holistic adoption: We help accelerate the adoption of data science into the MRM function and imbibe best practices around model use, conceptual soundness, theoretical framework, implementation, data review, and model validation • Ongoing monitoring: A new way of tracking model performance is emerging for AI/ML models, particularly for monitoring data drift (inputs outside the trained range) and concept drift (change in functional relationship between target and predictors)
Processes and guidelines	• Credit rating: Assess your readiness and game-plan change to submit information to credit rating agencies • Strong orientation to user experience: We provide elaborate guidelines to help users navigate AI/ML challenges • Data/analytics: We help establish data management and analytics processes for social media and customisation of products • Software: We assist in integration and migration of processes from legacy systems
Documentation	• Model cards: CRISIL develops short documents aimed to increase transparency in model reporting to encourage dialogue and ensure all stakeholders have access to both technical and non-technical information
Talent	• Team structure: We support cross-functional teams mapped to key objectives • Training needs: We educate stakeholders on business needs, particularly MRM

Contacts




Stephen Knights, PhD
Director, Risk and Analytics
London, UK
stephen.knights@crisil.com

Alberto Ramirez, FCA, MAAA
Actuarial and Insurance
Practice Leader
Chicago, US
alberto.ramirez@crisil.com

Valeriy Filatov, CFA
Analyst, New York, USA

Arshiya Sood
Analyst, Delhi, India

About CRISIL Limited
CRISIL is a leading, agile and innovative global analytics company driven by its mission of making markets function better.
It is India's foremost provider of ratings, data, research, analytics and solutions, with a strong track record of growth, culture of innovation and global footprint.
It has delivered independent opinions, actionable insights, and efficient solutions to over 100,000 customers.
It is majority owned by S&P Global Inc, a leading provider of transparent and independent ratings, benchmarks, analytics and data to the capital and commodity markets worldwide.

About CRISIL Global Research & Risk Solutions
Global Research & Risk Solutions is the world's largest and top-ranked provider of high-end research and analytics services. We are the world's largest provider of equity and fixed income research support to banks, and the foremost provider of end-to-end risk and analytics services to trading and risk management functions at world's leading financial institutions. We provide corporate research and analytics solutions to operations, strategy, and sales and marketing teams of corporations globally. Coalition provides analytics and business intelligence to 14 leading global investment banks. We operate from 8 research centers in Argentina, China, India and Poland, working with clients across time zones and languages. Being part of CRISIL enables us to attract and retain top quality talent. We have over 2,300 employees, 75% of whom hold advanced degrees in finance, accounting and management. We employ the largest number of CFAs and CFA aspirants in India. We have won top honours at the World HR Congress on Talent Management and HR Project for the year 2015. We have also won the NASSCOM Exemplary Talent Practices Award (NExT Practices) for skill development for two years in succession in 2011 and 2012. The award recognizes us as a firm that has the vision to proactively invest in its people and get them future-ready.
We are committed to delivering cutting-edge analysis, opinions, and solutions. This underscores our proposition of being the best people to work with.