CAPCO

THE CAPCO INSTITUTE
JOURNAL
OF FINANCIAL TRANSFORMATION

ALTERNATIVE RISKS

AI augmentation for large-scale global systemic and cyber risk management projects: Model risk management for minimizing the downside risks of AI and machine learning

YOGESH MALHOTRA

ALTERNATIVE CAPITAL MARKETS

#49 APRIL 2019

THE CAPCO INSTITUTE

JOURNAL OF FINANCIAL TRANSFORMATION

RECIPIENT OF THE APEX AWARD FOR PUBLICATION EXCELLENCE

CONTENTS

ALTERNATIVE MODELS

ALTERNATIVE RISKS

ALTERNATIVE MARKETS

DEAR READER,

Welcome to edition 49 of the Capco Institute Journal of Financial Transformation.

Disruptive business models are re-writing the rules of our industry, placing continuous pressure on financial institutions to innovate. Fresh thinking is needed to break away from business as usual, to embrace the more rewarding, although more complex alternatives.

This edition of the Journal looks at new digital models across our industry. Industry leaders are reaching beyond digital enablement to focus on new emerging technologies to better serve their clients. Capital markets, for example, are witnessing the introduction of alternative reference rates and sources of funding for companies, including digital exchanges that deal with crypto-assets.

This edition also examines how these alternatives are creating new risks for firms, investors, and regulators, who are looking to improve investor protection, without changing functioning market structures.

I am confident that you will find the latest edition of the Capco Journal to be stimulating and an invaluable source of information and strategic insight. Our contributors are distinguished, world-class thinkers. Every Journal article has been prepared by acknowledged experts in their fields, and focuses on the practical application of these new models in the financial services industry.

As ever, we hope you enjoy the quality of the expertise and opinion on offer, and that it will help you leverage your innovation agenda to differentiate and accelerate growth.

Lance Levy, **Capco CEO**

AI AUGMENTATION FOR LARGE-SCALE GLOBAL SYSTEMIC AND CYBER RISK MANAGEMENT PROJECTS: MODEL RISK MANAGEMENT FOR MINIMIZING THE DOWNSIDE RISKS OF AI AND MACHINE LEARNING

YOGESH MALHOTRA | Chief Scientist and Executive Director, Global Risk Management Network, LLC

ABSTRACT

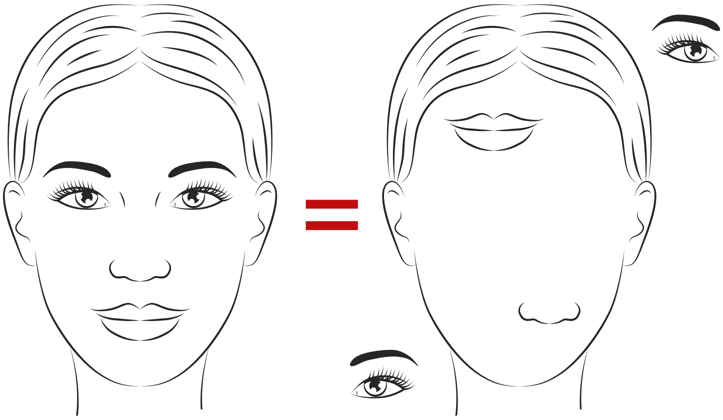
This article discusses how model risk management in operationalizing machine learning (ML) or algorithm deployment can be applied in national systemic and cyber risk management projects such as Project Maven. After an introduction about why model risk management is crucial to robust AI, ML, deep learning (DL), and neural networks (NN) deployment, the article presents a knowledge management framework for model risk management to advance beyond "AI automation" to "AI augmentation."

1. INTRODUCTION: PROJECT MAVEN

Project Maven, also known as "algorithmic warfare cross-functional team" (AWCFT), represents one of the first operational applications of Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), and Neural Networks (NN) technologies in defense intelligence. Its operational focus is on the analysis of full-motion video data from tactical aerial drone platforms, such as the ScanEagle, and medium-altitude platforms, such as the MQ-1C Gray Eagle and the MQ-9 Reaper. As noted by Maven CO, Air Force Lt. Gen. Jack Shanahan, "Maven is designed to be that pilot project, that pathfinder, that spark that kindles the flame front of artificial intelligence across the rest of the Department."

Supported by a budget of U.S.$70 million, Project Maven, executed in collaboration with AI researchers from industry, aimed to achieve the distinction of deploying AI deep neural networks (DNNs) in active combat theater within six months from launch. Given that defense intelligence services are "drowning in data," AI and DL technologies, such as DNNs, provide essential respite by automating tedious work activities, such as counting cars, individuals, and, activities, and typing their counts into PowerPoint files and MS-Excel spreadsheets. The success of the project was bolstered by building partnerships with AI experts in industry and academia and with Department of Defense (DoD) communities of drone sensor analysts.

Figure 1: Limitations in spatial representations of features



Collaboration with top AI talent from outside the defense contracting base facilitated accelerated adoption of commercial AI, ML, and DL technologies. The above project focused on development of agile iterative product prototypes and underlying infrastructures along with ongoing user community testing. In addition, key AI system development activities, such as labeling data, developing AI-computational infrastructure, developing and integrating neural net algorithms, and receiving user feedback, were all executed iteratively and in parallel. AI techniques for imagery analysis are extremely capable, yet developing algorithms for specific applications is not simple. For instance, AI systems require labor-intensive classification and labeling of huge datasets by humans for training of DL algorithms.

> “Machine Learning deals with computer programs that try to learn from experience for prediction, modeling, understanding data, or controlling something.”

Maven needed individual labeling of more than 150,000 images for its first training datasets, with plans to have 1 million images labeled by January, 2018. Throughout the DoD, every AI successor to Maven will need a similar strategy for acquiring and labeling a large training dataset.

Maven's success is clear proof that AI-ML-DL is ready to revolutionize many national security missions. Having met sky-high expectations of the DoD, it is likely to spawn 100 copycat “Mavens” in DoD C4I (Command, Control, Communications, Computers, and Intelligence).

2. ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, DEEP LEARNING AND NEURAL NETWORKS

Project Maven focused on autonomous classification of objects of interest from still or moving images using computer vision enabled by AI, ML, and DL. MIT management scientist Tom Malone defines AI in intuitive terms, such as “machines acting in ways that seem intelligent.” MIT computer scientist Patrick Winston notes that: “AI is about the architectures that deploy methods enabled by constraints exposed by representations that support models of thinking, perception, and action.”[1] In contrast to general AI, which can solve many different types of problems, as humans do, most AI systems are narrow AI machine-based systems with the capabilities of addressing a specific problem, such as playing Go or chess.

According to MIT computer scientist Tommi Jaakkola, ML deals with computer programs that try to learn from experience for prediction, modeling, understanding data, or controlling something.[2] In the case of Project Maven, such ML is from a training set of labeled examples of images of objects to make future predictions for classifying instances of such objects. As computers process data as bits, images need to be translated into geometrical representations called “feature vectors” composed of such bits. Feature vectors are essentially arrays containing numeric identifiers representing the specific attributes or features of the respective object. The problem is hence translated from a set of images into a set of vectors: a vector being a two-dimensional matrix with only one row but multiple columns of numeric data.

The training set contains a set of labeled vectors and the test set contains a set of images to be classified consisting of unlabeled vectors to match with respective labels. Using vectors and labels, ML algorithm translates the problem into a geometric form wherein each vector represents a point in n-dimensional space. The solution involves developing an ML algorithm to divide n-dimensional space into specific parts, each of which corresponds to a specific label. For image classification, such geometrical

1 MIT AI-Machine Learning Executive Guide: including Deep Learning, Natural Language Processing, Autonomous Cars, Robotic Process Automation: https://bit.ly/2PXfIQH, MIT AI-Machine Learning executive education course videos.

2 Ibid.

Figure 2: GAN: CNNs see all images on the right as ostriches

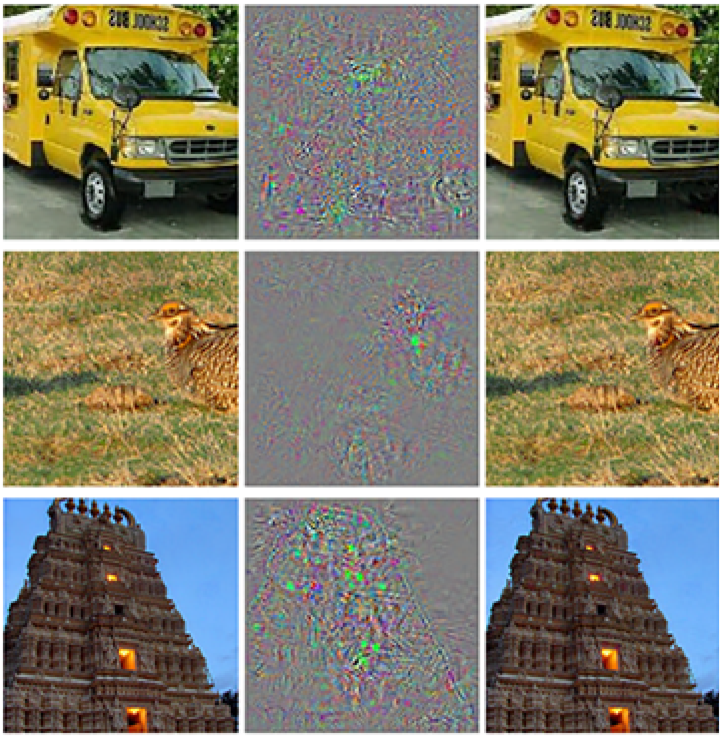
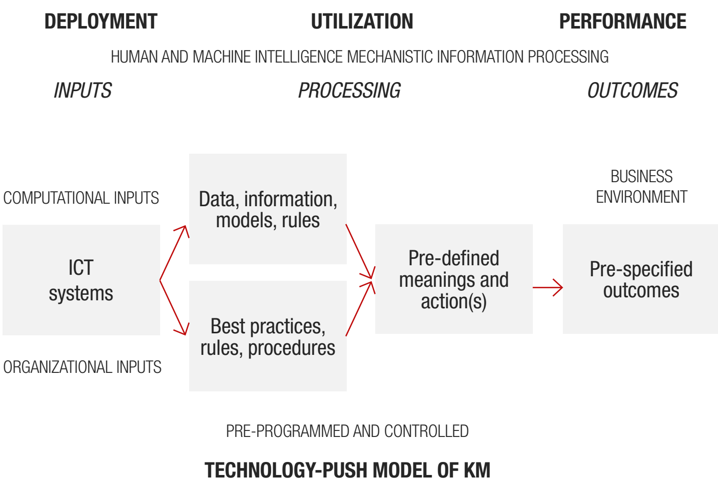


Figure 3: Technology-push inputs driven models: suitable for static and deterministic environmental and operational contexts



transformations use image filters to distinguish between low-level and high-level features such as edges (i.e., boundaries between objects and combinations of edges, curves, parts, and the object).

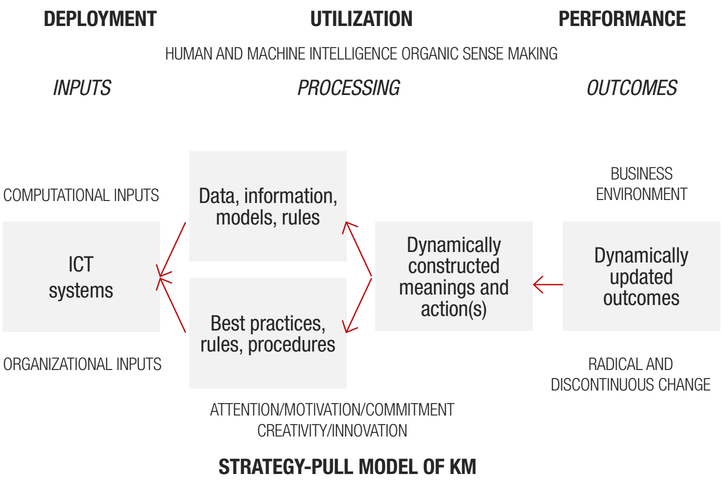
The image signal traverses different transformation layers for processing low- to high-level features with the ML solution being specification of transformation layers and how low-level features are combined. More granular specification and precision is feasible using multiple layers of transformation, with the number of such layers representing the **depth** of the model and the ML problem becoming a **deep** learning problem. Such DL architectures, which are based on fine tuning of millions of parameters across multiple layers of mathematical and geometrical transformations, pose **interpretability** and **trustability** challenges.

Algorithms called neural networks (NNs) are deployed to automate processing of text, voice, and images once they have been trained using millions of example images of such objects. NNs containing multiple transformation layers are called deep neural networks (DNNs). Three general types of DNNs are in common use for text, voice, and image processing. Convolutional neural networks (CNNs) are commonly used for classification of visual images and are an example of feedforward neural networks that have acyclic nodes with all inputs and outputs independent of each other. Recurrent neural networks (RNNs), in contrast, are used for natural language processing (NLP) of sequential information and contain cyclic nodes with outputs being dependent on previous computations. Long short term memory networks (LSTMs) are an extension of the most commonly used type of RNNs that better capture long-term dependencies for sequential information flows given much longer-term memory than vanilla RNNs.

3. WHY MODEL RISK MANAGEMENT IS MOST CRUCIAL TO ROBUST AI-ML-DL USE

As noted earlier, CNNs are commonly used for classification of still or moving images, such as in the case of Project Maven for autonomous classification of objects of interest. Geoff Hinton, a pioneer of CNNs, noted recently that: "I think the way we're doing computer vision is just wrong. It works better than anything else at present but that doesn't mean it's right." Simultaneously, his lecture notes[3] highlight "Why convolutional networks are doomed," observing that: "sub-sampling loses the precise

Figure 4: Strategy-pull outcomes driven models: suitable for complex and uncertain environmental and operational contexts



spatial relationships between higher-level parts such as a nose and a mouth. The precise spatial relationships are needed for identity recognition." (Figure 1)[4]

Mathematically, CNN ignores spatial relationships between the lower-level features such as eyes, nose, and, mouth; hence it computes the above two images in Figure 1 as being equivalent. Computer scientists and neuroscientists also note the challenges of interpretability and trustability that the fallibility of AI, and in particular DL, pose. Patrick Winston of MIT describes advances in AI in the past years as "computational statistics" rather than AI, observing that machines don't have common sense: "The computer that wins at Go is analyzing data for patterns. It has no idea it's playing Go as opposed to golf, or what would happen if more than half of a Go board was pushed beyond the edge of a table…"[5] Tomaso Poggio of the McGovern Institute for Brain Research at MIT, notes that "These systems are pretty dumb. We have not yet solved AI by far. This is not intelligence."[6]

The latest and, deemed greatest, innovation in AI-ML-DL is called Generative Adversarial Network (GAN). GAN is comprised of two nets, the "generator" generates new instances of data and the "discriminator" evaluates them for authenticity. The discriminator, which is a standard CNN, tries to determine whether a specific instance of data belongs to the actual training dataset or not. The generator is like an inverse CNN, which given random numbers generates an image. The goal of the generator is to pass fake images as authentic to the discriminator which then evaluates the images for authenticity based on its ground truth of real images. As seen in Figure 2, ML models are vulnerable to adversarial examples: small changes to images can cause computer vision models to make mistakes such as identifying a school bus as an ostrich. Human eyes cannot discern that images on the right are distorted versions of those on the left; CNN sees the three as ostriches.[7]

4. A KNOWLEDGE MANAGEMENT FRAMEWORK FOR MODEL RISK MANAGEMENT

For static and deterministic environmental and operational contexts, predictive modeling underlying AI-ML-DL is most optimal (Figure 3). Problems are defined in terms of static features (or attributes, characterizing respective objects) and feature vectors (i.e., mathematical arrays containing numeric representations of such features) that can be resolved optimally by pre-programmed and controlled mechanistic human and machine intelligence. As noted earlier, feature vectors are essentially arrays containing numeric identifiers representing the specific attributes or features of the respective object, a vector being a two dimensional matrix with only one row but multiple columns of numeric data.

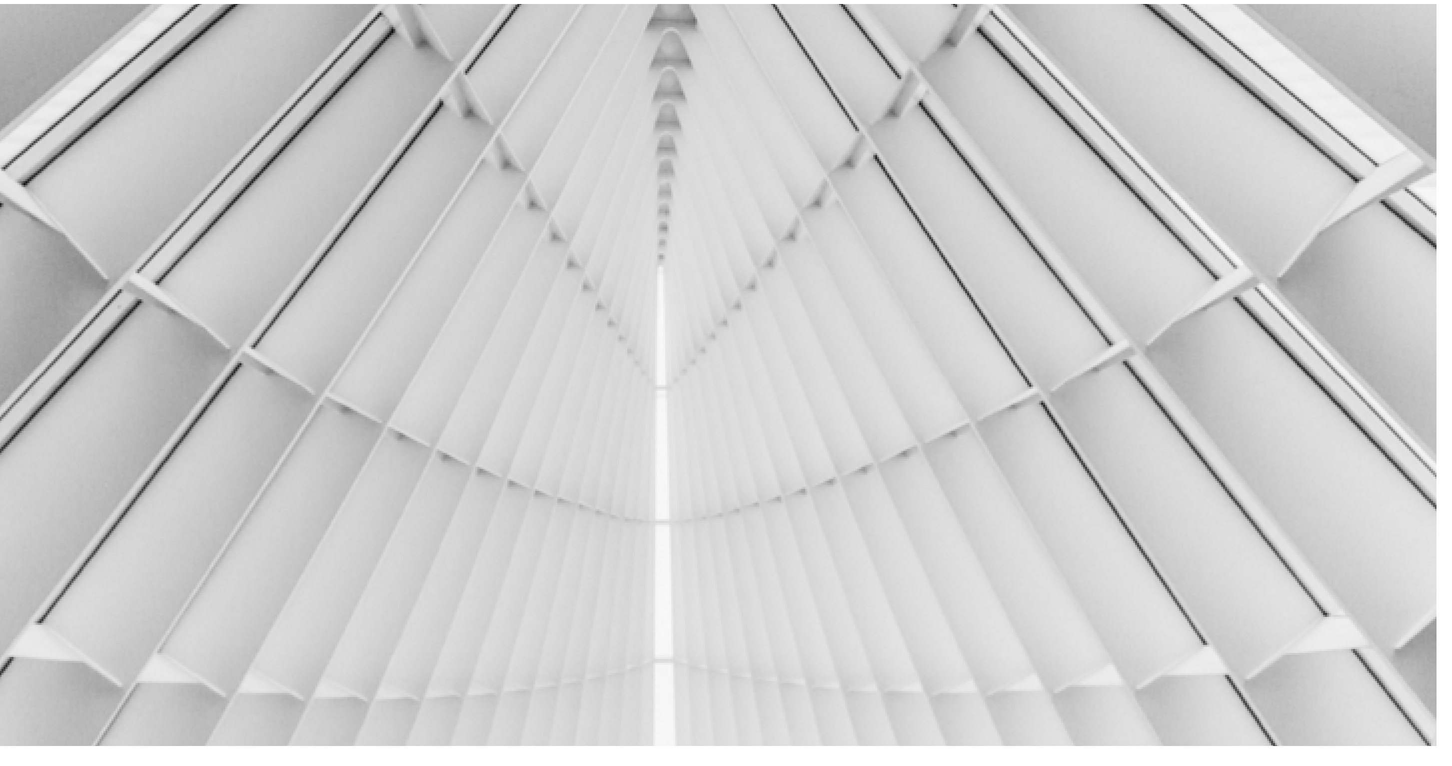
However, in contexts characterized by complexity and uncertainty, as in Figure 4, predictive analytics based on historical data do not meet the dynamic target given pre-specified outcomes. Hence, anticipation of surprise is needed along with requisite variety to tackle dynamic uncertainty and complexity.[8]

Model risk management (MRM) is needed for environmental and operational contexts that do not match static and deterministic criteria with pre-defined and pre-programmed problems and solutions. MRM is a function of the variance in both inputs and outcomes, as observed in Figures 1 and 2, respectively. Use of any statistical or

3 Hinton, G., "Taking Inverse graphics seriously," lecture notes, Department of Computer Science, University of Toronto, https://bit.ly/2Ud0KTy
4 Pechyonkin, M., 2017, "Understanding Hinton's Capsule Networks. Part I: Intuition," Medium, November 2, https://bit.ly/2AcPGg0
5 Refer to footnote 1
6 Ibid.
7 Elsayed, G. F. S. Shankar, B. Cheung, N. Papernot, A. Kurakin, I. Goodfellow, and J. Sohl-Dickstein, 2018, "Adversarial examples that fool both computer vision and time-limited humans," Cornell University, May 22, https://bit.ly/2U9LITF
8 Malhotra, Y., 2005, "Integrating knowledge management technologies in organizational business processes: getting real time enterprises to deliver real business performance," Journal of Knowledge Management 9:1, 7-28

mathematical model entails model risk since the specific results are not measured but estimated using the specific statistical and mathematical models. An important insight from model risk management research and practices is that there is unlikely to be any perfect model (all models are wrong), and the best results can be obtained from combining the results from models based on different inputs (some models are useful) – "All models are wrong, but some are useful" – George E. P. Box. Hence, instead of relying on any one specific quantitative model, using a range of different plausible quantitative models, which can be robustly discriminated from one another, is a recommended strategy for minimizing the model risk. When results from multiple models are combined, analogous to the use of "ensemble models" such as in ensemble learning, the variance in the range of estimates across the respective models provides a succinct measure of model risk. The papers and presentations downloadable from the author's SSRN page (https://papers.ssrn.com/author_id=2338267) discuss multiple specific examples of model risk management in the context of complex systems, spanning quantitative finance and hedge fund trading systems and cyber risk insurance systems to AI-ML-DL-GAN applications in Space and Defense projects such as Project Maven. One example is the recent invited presentation to the CFA Society on Hedge Fund Chief Investment Officer Practices on using Auto-Machine Learning (Auto-ML) for Model Risk Management (https://bit.ly/2tlg3b7). The current article spans the focus from Cybersecurity, Finance, and, Insurance to broader applications of AI-ML-DL-GANs in the Defense & Space risk management contexts, such as the Project Maven.

> "In dynamic, complex, and uncertain environments, anticipation of surprise is more important than predictive analytics based on historical data as the past may not be the best predictor of the future."



Specific examples will include multiple variations of the CNNs and related models being used to address the limitations of any one given model. Furthermore, the capsule networks (CapNets), which are proposed as a solution for ameliorating many of the limitations of CNNs noted earlier, provide additional diversity in terms of different plausible models that can be robustly discriminated between. Broadening the range of estimates based upon diverse models provides a better assessment of risk in terms of variance.

5. CONCLUSION: BEYOND "AI AUTOMATION" TO "AI AUGMENTATION"

As illustrated in the case of GANs, small changes to images not discernible to humans can cause computer vision models to make mistakes, such as seeing a school bus as an ostrich. While it is easy for humans to see a bus as a bus, it is hard for AI-ML algorithms to do so. Many simple tasks that anyone can do, like recognizing objects or picking them up, are much harder for AI-ML-DL as a recent report by the consulting firm Deloitte notes.[9] On the other hand, many of the issues related to algorithmic bias may be traced back to bias in training data or the design of algorithms and models. The same report notes that "AI algorithms must be complemented by human judgment."

Remarking on the certainty of knowledge, Morris Kline had noted: "Insofar as certainty of knowledge is concerned, mathematics serves as an ideal, an ideal toward we shall strive, even though it may be one that we shall never attain. Certainty may be no more than a phantom constantly pursued and interminably elusive."[10] Emanuel Derman observed: "Models are at bottom tools for approximate thinking. The most important question about any model is how wrong it is likely to be, and how useful it is despite its assumptions. You must start with the model and overlay them with common sense and experience."[11]

There is no right model as the world changes in response to the ones we use. In addition, changing environmental and operational contexts make newer models necessary. Hence, knowing and applying the leading-edge developments in AI-ML-DL-GAN models is important for ensuring systemic and cyber risk management progress and growth aligned with world developments. It is, however, equally important to know the limits and boundaries of the models and related assumptions and logic by deploying "audacious imagination, insight, and creative ability"[12] as noted by the mathematician Morris Kline.

[9] Guszcza, J., H. Lewis, and P. Evans-Greenwood, 2017, "Cognitive collaboration: why humans and computers think better together," Deloitte Insights, January 23, https://bit.ly/2wetBzl
[10] Kline, M., 1980, Mathematics: the loss of certainty, OUP
[11] Derman, E., 1996, "Model risk," Goldman Sachs Quantitative Strategies Research Notes
[12] Refer to Footnote 10

ABOUT CAPCO

Capco is a global technology and management consultancy dedicated to the financial services industry. Our professionals combine innovative thinking with unrivalled industry knowledge to offer our clients consulting expertise, complex technology and package integration, transformation delivery, and managed services, to move their organizations forward.

Through our collaborative and efficient approach, we help our clients successfully innovate, increase revenue, manage risk and regulatory change, reduce costs, and enhance controls. We specialize primarily in banking, capital markets, wealth and asset management and insurance. We also have an energy consulting practice in the US. We serve our clients from offices in leading financial centers across the Americas, Europe, and Asia Pacific.

WORLDWIDE OFFICES

APAC
Bangalore
Bangkok
Hong Kong
Kuala Lumpur
Pune
Singapore

EUROPE
Bratislava
Brussels
Dusseldorf
Edinburgh
Frankfurt
Geneva
London
Paris
Vienna
Warsaw
Zurich

NORTH AMERICA
Charlotte
Chicago
Dallas
Houston
New York
Orlando
Toronto
Tysons Corner
Washington, DC

SOUTH AMERICA
São Paulo

WWW.CAPCO.COM

CAPCO