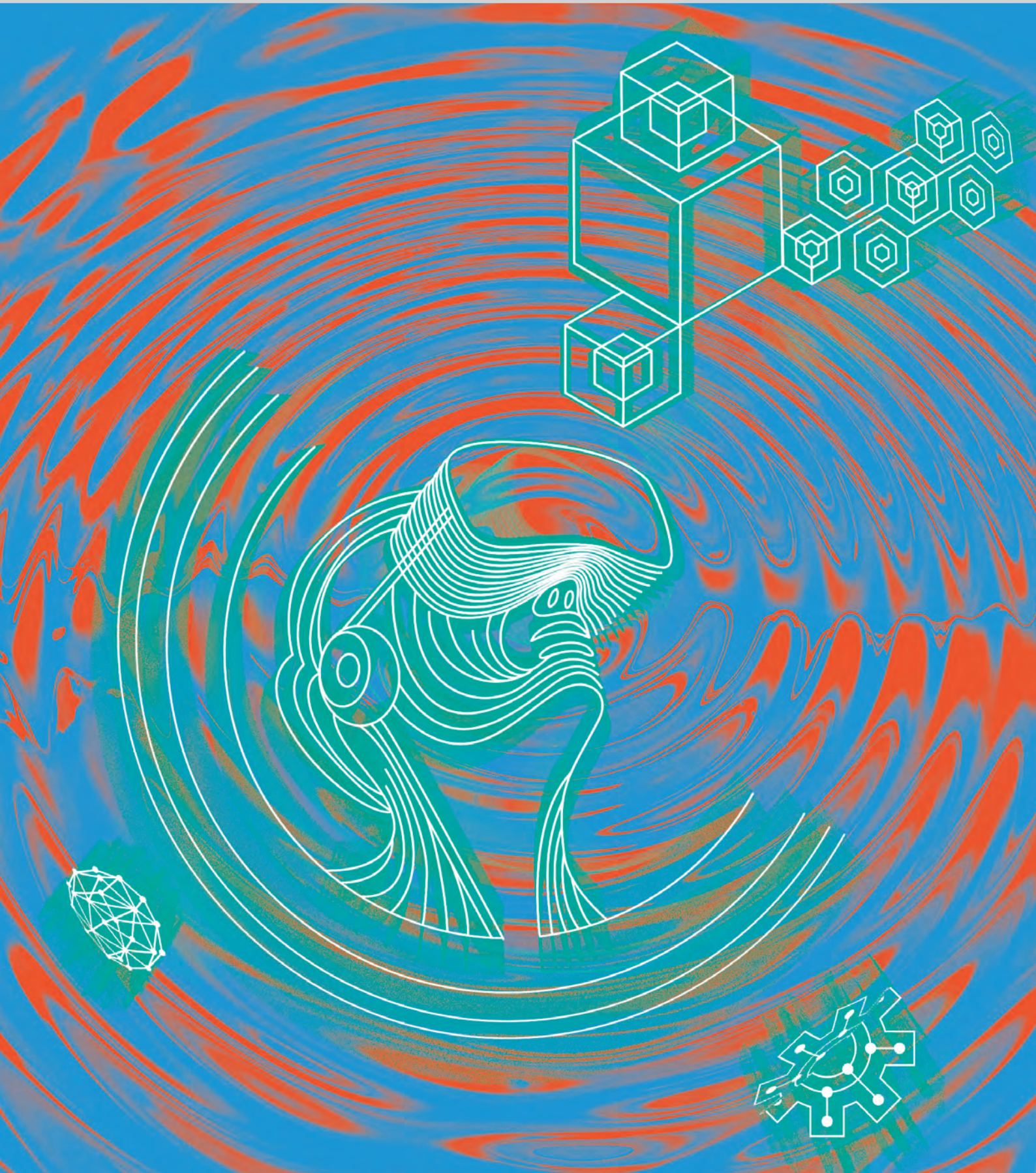
TATA CONSULTANCY SERVICES
Experience certainty.

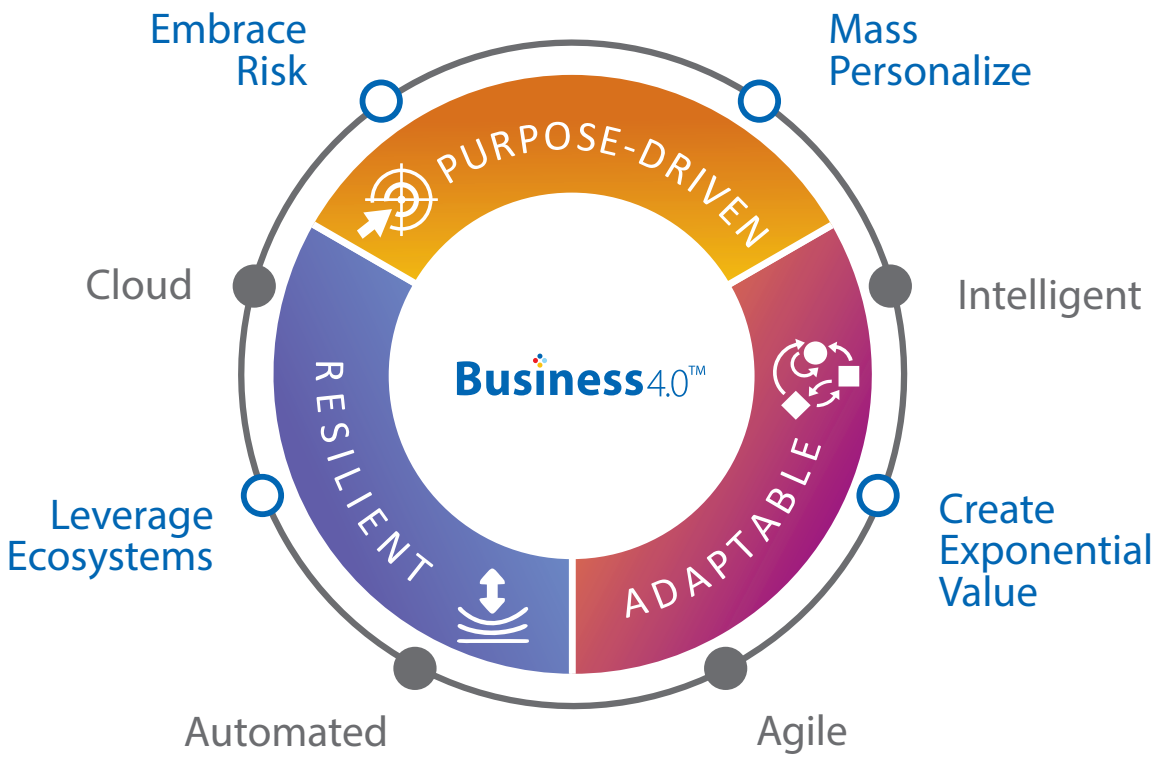
TATA

Reimagining Research
Volume II



PURPOSE - DRIVEN, RESILIENT & ADAPTABLE

with Business 4.0™



Reimagining Research

Volume II

TATA CONSULTANCY SERVICES

Book Title: Reimagining Research Volume II

Published by: Tata Consultancy Services; March 2020

Printed at: Rathna Offset Printers, Chennai

Typeset by: MPS Limited

Editorial Team: Seetha Srikanth, Content Lead, TCS Research & Innovation.
Ashmita T Rajaprathapan and Shobhana Suresh, Content Team, TCS Research & Innovation.

S Sahu, Freelance Writer.

Mentors: S Santhanakrishnan, Head – Process, TCS Research & Innovation.
Chakravarthi Sathyanarayana, Principal Consultant, TCS Research & Innovation.

Sundar Vinayakam, Principal Innovation Evangelist, TCS Research & Innovation.

Design Consultant: Sameer Karmarkar, Satisfice Designs Pvt Ltd

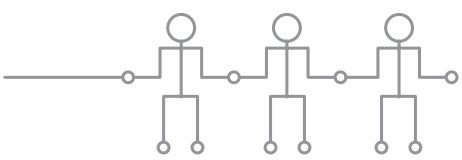
Cover and Separators Design: Falguni Gokhale, Design Directions Pvt Ltd

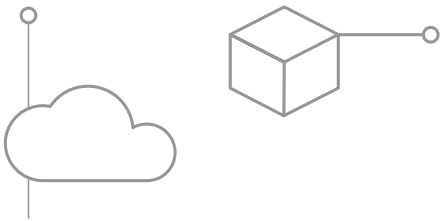
External Review for Tone of Voice: Netscribes (India) Private Limited

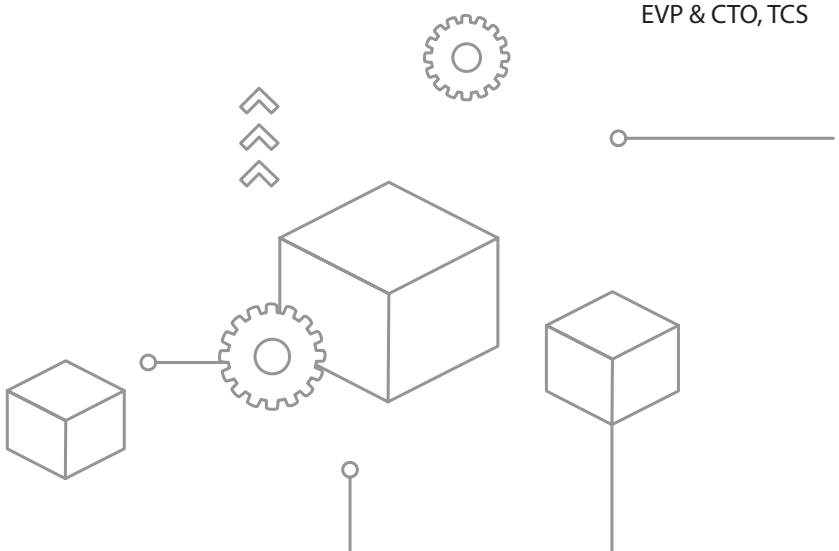
Reimagining Research

Volume II

TABLE OF CONTENTS











Innovation Forum

The TCS Innovation Forum is a premier, invite only event that annually engages 1,000+ senior business leaders from Europe, the United Kingdom, North America, Latin America, and Japan. Hosted by the Chief Technology Officer of Tata Consultancy Services, the event focuses on Innovation and Research through cutting edge showcases and a series of sessions comprising notable speakers from business, academia, government, and scientific institutions.

Reimagining Research

Gautam Shroff

Research plays a critical part in TCS' business. Being close to market leaders across industries offers TCS Research a ringside view of new technology adoption and the problems that need to be solved in the process. Interacting with customers, academics and technologists for the past three decades has been an immense source of learning for me personally. I would like to share some thoughts here on what I see ahead of us. And how we at TCS Research are reorienting ourselves to keep pace with these trends.

Three Game Changers

The pace of technology evolution is set to increase like never before. I see three important drivers:

- There is a growing availability of data about virtually everything, from people to machines to the environment. Everything that can be measured, tracked, and stored, will be, as communications and storage become cheap and ubiquitous.
- Super-linear improvements in the capabilities of Artificial Intelligence (AI) techniques to mimic, and in many cases, surpass human abilities, in tasks normally associated with requiring human involvement:
    i. Tasks requiring perception and understanding, e.g. recognizing objects, faces and text from images and video, transcribing voice to text, summarizing documents, and translating one natural language to another
    ii. Prediction tasks, e.g., detecting as well as forecasting a machine failing, or a patient having an adverse reaction or even a hospital re-admission
    iii. Prescriptive tasks, e.g., making complex decisions, such as which move to make next in a game, or which train to schedule on which track; tasks for which optimal solutions are, in principle, mathematically computable, but in practice, computationally intractable, but which humans manage quite well using experience
    iv. Participative tasks, e.g., physical machines (robots) operating autonomously in the real world, such as self-driving cars, or flexible industrial robots that can be "taught" new jobs or even learn by watching humans, rather than being programmed
- The movement of the traditional brick-and-mortar economy as well as governments toward integrating AI techniques into their IT systems. This will be the most impactful development as these are organizations responsible for 90% of the global GDP. They will better exploit their internal and external data so as to drive fundamental changes to their business models.

AI Meets the Brick and Mortar Economy

We are already seeing signs of this. Banks are seeking to go beyond being mere accountants of money and facilitators of payments. They exploit the data they are privy to information such as "who pays whom for what" to assist merchants in their advertising and pricing strategies.

Insurance is moving from provisioning to prevention; examples include mobile apps that track the driving behavior of customers and reward those that drive more safely as well as health monitors to

encourage fitter and healthier lifestyles. Not just autonomous vehicle makers, but all car manufacturers are using data from the hundreds of sensors that populate each modern vehicle to better predict failures and prevent costly recalls. Similar trends are observed for telecom operators, utility companies, retail chains, hospitals, and even governments. At the heart of each of these applications lies the intelligent analysis of data and convergence as traditional IT systems become AI-driven.

AI in the Fight against COVID 19

AI was leveraged in many ways to fight the COVID-19 pandemic. AI speeded up pharma research toward designing new molecules that may well fight the virus. It helped epidemiologists predict how the pandemic would spread and thereby helped Pune city administration manage lockdowns with as much efficacy as possible. In a time when most everyday actions were performed online (work, recreation, baking, shopping etc) AI played a critical role in ensuring safety in networks.

The Horizon-III Stars

What might transpire in 20 years or even 10, is difficult to predict in a frequently disrupted field like information technology. Here are some possibilities.

Quantum computing will likely be commercially viable. For this, the world has to re-invent its communications infrastructure to ensure security, as traditional encryption techniques have become easily breakable.

Machine-learning will become significantly more efficient, enabling the processing of vast amounts of data in real-time. This can be either via quantum or even DNA-based computing, or the development of neuromorphic chips based on the next generation of deep neural networks.

The biggest disruption will be at another grand convergence—better robots and cheap energy.

Robots might achieve human level dexterity. Energy-surplus may be achieved via solar, nuclear, or other means. This can drive the cost of production for virtually any goods to near zero; for with free energy, AI, and robots, we could both design and produce new materials including food, and manufacture whatever is needed for human consumption, at minimal cost. The economic impacts of a truly "free" economy are already being debated, albeit only as a matter of speculation at present.

Whatever might be the mechanism by which wealth is created and distributed in such an economy, one thing seems clear to me at least: leisure and digital entertainment will increase.

Advertising, in whatever avatar, will be needed. Data about consumers, consumption, and needs will increasingly drive the creation of new materials and products as well as manage their delivery and consumption for the purpose that we humans have always striven for—ever increasing prosperity and well-being in whatever sense society might choose to measure.

TCS Research will have to embrace these developments and help customers make smooth transitions at every evolution.

Research Paradigms at TCS

The Tatas believe in delivering the benefits of science to the common man. This vision led to Research as a separate activity in a dedicated facility in TCS with the establishment of Tata Research Development and Design Centre, TRDDC, in 1981 at Pune. Chairman F.C. Kohli and Director E.C Subba Rao attracted globally renowned researchers to the center. At this time, coding was artisanal. At TRDDC, Professor K.V Nori in particular worked on paradigms that would move computer science toward software engineering; from theory to a repeatable, scalable, and sustainable practice. We focused on model-driven development and object-oriented programming. We built dictionaries, compilers, translators, and program analysis tools. This resulted in some pioneering products for software automation (MasterCraft)™ and verification (TCS Embedded Code Analyzer); for industry (BaNCS), and for Society (Swach). The most important thing we learnt at TRDDC, however, was to look at a real world problem and abstract the issues to solve a large set of problems they indicated.

As our Research moved from TRDDC to global centers, we adopted a portfolio model in 2006, where we looked at research in the short- and long-term. This helped us deliver a number of solutions, platforms

and even lines of business (Ignio)™, while progressing with long-term research. Many of these stories have been published in our earlier book *Research by Design*, and in our website www.tcs.com/research and in our research reports. We also published a significant number of papers and filed for patents.

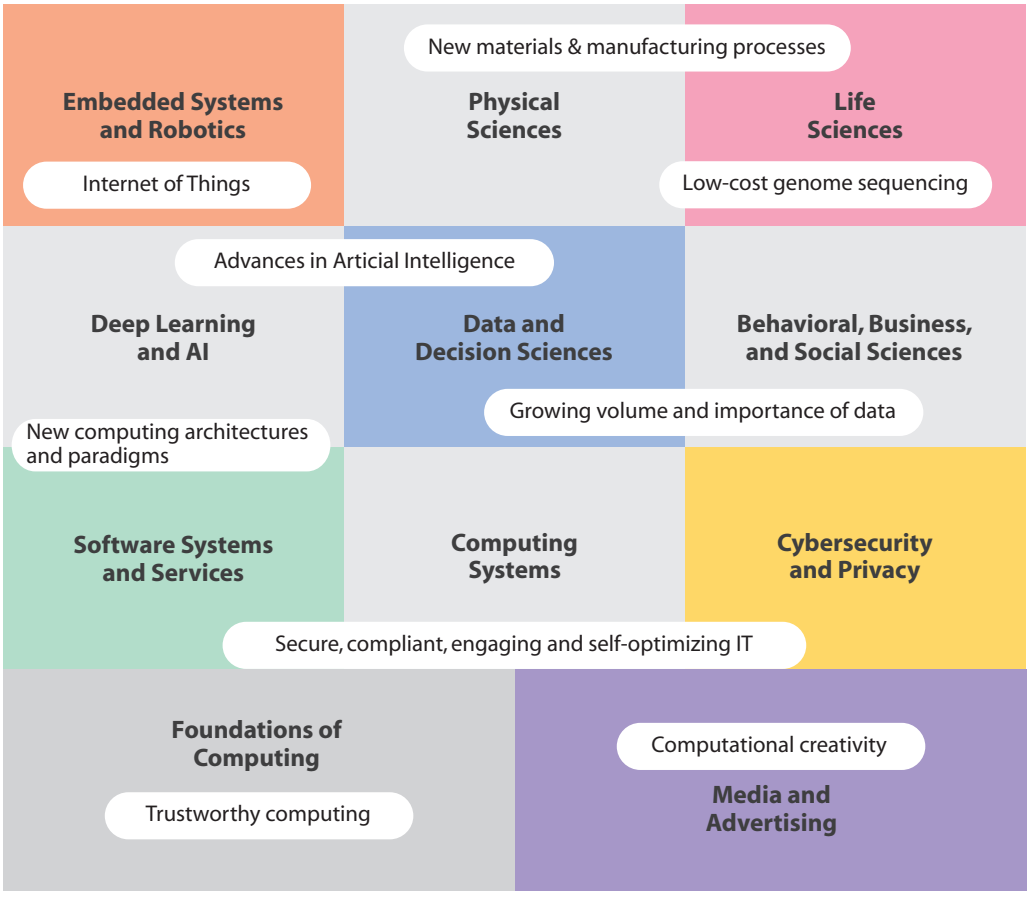
Now we are at another inflection point. If technology evolves faster, applied research must also speed up. As we do not want to stop looking at big research problems ambitiously and at depth, we have started industry-focused innovation units within the business, so as to keep one arm of research delivering innovation continuously for our customers. These will draw on the greater scientific rigor and long-term exploration taken up by the central research team. From the central team's explorations, we have initiated a formal mainstreaming of research as future-facing business lines.

A People Culture in Research

From its inception at TRDDC when professors from CMU did seminal work, the aim has always been to attract talented researchers and offer them a distinct career path. The research community in TCS has grown with many shining stars. We have rich connects with academia around the world providing a channel for exchange of ideas with some of the best minds in the world. Mentoring of junior researchers is something we take seriously. The research advisory board meets with researchers at regular intervals, reviewing work, and also providing outside-in perspectives. We encourage researchers to keep their thirst for knowledge alive by offering sabbatical opportunities. We insist that they learn and be esteemed by peers: publications and patents are critical metrics for TCS researchers. A vital factor to the success of our research is that researchers are well respected by the business. TCS Research leaders have direct access to the CxOs in the company.

Current Research in TCS, with an AI Lift

While today we are organized into eleven research areas, elements of AI run through each of these: the original Software Systems and Services group that created model-driven software engineering is now focusing on AI-driven program synthesis and meta-model based AI systems. Physical Sciences,



with its expertise in computational materials engineering, has developed research programs for IoT Analytics using AI, and is discovering new materials via data-driven techniques.

Life Sciences, with its focus on genomics and meta-genomics, leverages machine learning for both personalized medicine as well as new explorations in synthetic biology. Embedded Systems and Robotics is developing a business in intelligent and flexible robotics for Industry 4.0 applications as well as for emerging connected healthcare ecosystems. The Cybersecurity research area is addressing the many privacy concerns engendered by AI via its tools for end-user control and content that are finding rapid traction given the growing global concerns in this arena. To deal

with the growing importance of data, both the Data and Decision Sciences as well as the Deep Learning and AI research areas deal with cutting-edge research such as applying reinforcement learning to real-world optimization problems, or applying and enhancing the latest deep-learning techniques for sensor analytics, text, and image processing. With data about individuals today being so prevalent and rich, the Behavioral Business and Social Sciences research is, among other things, attempting to study collections of people much as physicists study collections of molecules, as well as mechanisms for technology to add to entertainment and the inevitable advertising that goes with it. Last but not least, software applications, AI or otherwise, need to be engineered for high performance and use the latest in computing technology, which is the focus of the Computing Systems research area, such as developing high-performance, deep-learning frameworks for e-commerce or keeping tabs on emerging trends such as quantum computing.

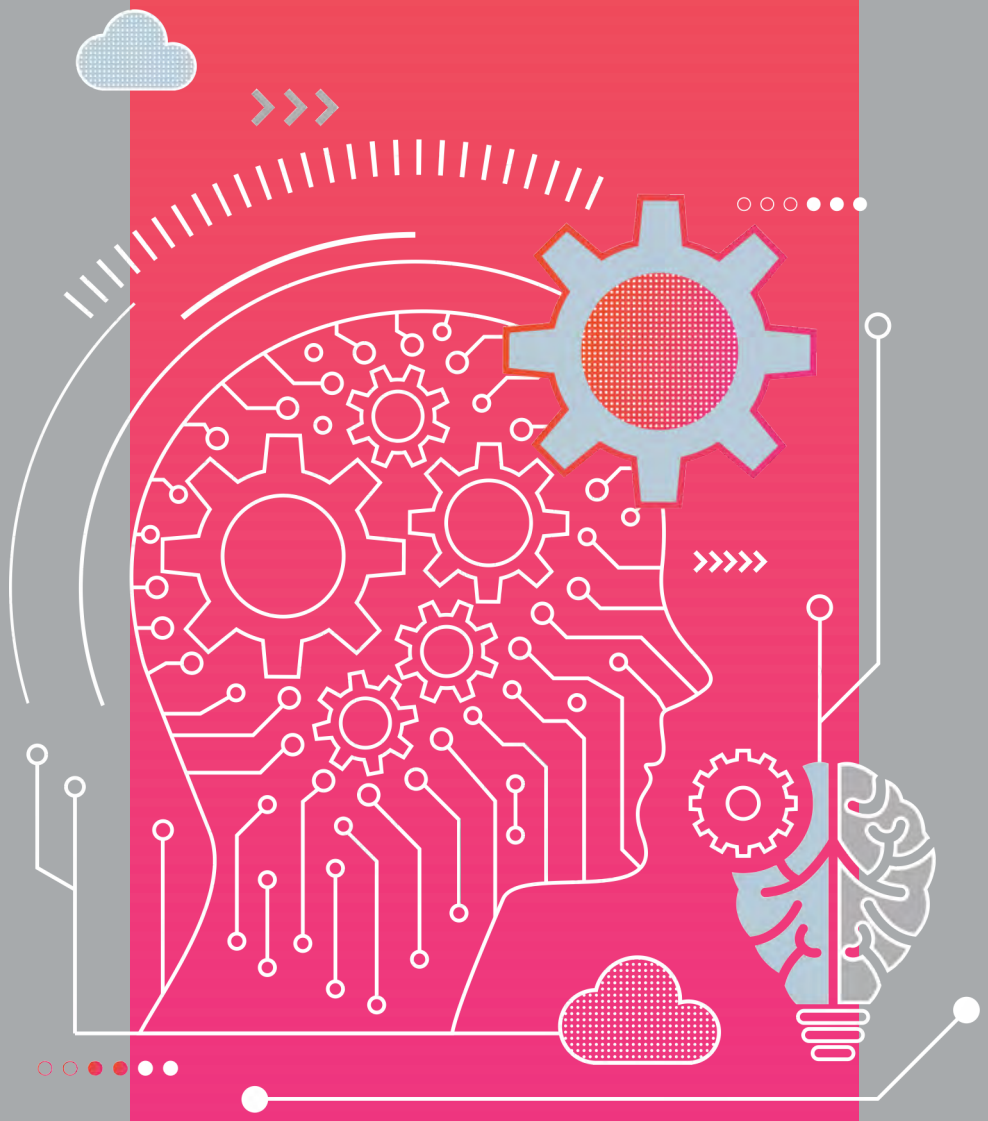
In this book, which is the sequel to the volume published in 2018, we highlight examples from TCS Research across two distinct themes, the industrialization of computing on the one hand, and the digitization of industry and society on the other. The former theme is organized into sections covering Data and Computing related topics, while the latter into sections on Industry, Society, and Enterprise. Each of these chapters in a section highlight work in a selected research group or program aligned to the respective theme. While not exhaustive, this collection should give the reader a very good flavor of what TCS Research is all about. In the Business 4.0 era, when customers are looking at digital and beyond, these essays provide perspectives on what lies ahead in a number of areas. We hope these essays can start conversations about technology and stoke thought leadership.

Gautam Shroff

Gautam Shroff is a Vice President and Chief Scientist and heads TCS Research. In 1994, he was conferred the 'Young Scientist', Award from the Indian Department of Atomic Energy. Gautam has published over 60 research papers in the areas of computational mathematics, parallel computation, distributed systems, software architecture, software engineering, big data, information fusion, virtual reality as well as artificial intelligence including machine learning, deep learning, Bayesian inference and natural language processing.

He has written two books: "Enterprise Cloud Computing" published by Cambridge University Press, UK, in October 2010; and "The Intelligent Web", published by Oxford University Press, UK, in 2013 (paperback ed. 2015). He completed his B.Tech degree in Electrical Engineering from IIT Kanpur in 1985, and Ph.D in Computer Science from Rensselaer Polytechnic Institute Troy, NY in 1990.

Digitally Embedded Intelligence



TCS Pace Port™

A physical manifestation of the TCS Pace™ promise, TCS Pace Ports are spaces that bring to life agile innovation. Each Pace Port is close to an academic research institute of repute and a COIN accelerator. It offers design thinking and collaborative work spaces. Pace Ports bring customers, academic faculty, students, start-ups, and TCS Research to brainstorm and create rapid prototypes of innovative solutions.

Intelligence in the Data World

Anand Sivasubramaniam

Data analytics is not a new subject. Statistical and machine learning techniques to analyze data have existed for decades. So why has there been a huge hype and pervasiveness of big data in the past decade alone? The answer to this lies in the two broad advancements that have contributed to the prevalence and success of big data.

The first is related to the availability of the data itself. Digitized data has typically been housed in IT systems, which have conventionally been siloed. Increasingly, these IT systems are being interconnected to cooperatively tackle large-scale problems, making access to disparate data seamless. Digitization is spreading beyond the virtual world, with the rapid evolution of the Internet of Things (IoT) in this era. Physical world sensing is providing voluminous streaming data, encompassing diverse domains – from personal healthcare and environmental data to physical infrastructures at the city scale. Finally, social interactions in the networked world are also providing digitized data that was not easily available until now.

The second advancement is related to the technological progress in hardware and software to run the data analysis for meaningful purposes. As Moore's law predicted, plentiful transistors, over several decades, have taken on the immense challenges posed by big data applications. Equally, software advancements in distributed computing, cloud analytics, and open source software have democratized the accessibility to data analysis techniques for the population at large.

With these two trends firmly taking root, the onus is now on us to collect meaningful data from whatever sources we can access, using the available computing power to derive actionable insights or for decision making. The next set of articles give examples of such exercises. The first article, "The Seamless Integration of 5G", discusses seamless network provisioning with real time network predictions and how crucial it is as a business requirement. The article "Reading Images" describes the need for information extraction from document images and how this has been done using relation extraction and natural language. The next article, "Sensing for Health", highlights how we enable better and more reliable decision making to interpret better health systems whether it is referring to human beings or machines. The article, "The Who, What, and How of Voice" discusses natural conversational languages in noisy conditions and how our patented analysis tools and algorithms help in addressing several voice related challenges faced by enterprises.

Anand Sivasubramaniam

Anand Sivasubramaniam is a Consulting Advisor for Tata Consultancy Services and heads Data and Decision Sciences Research. He is a Distinguished Professor of Computer Science and Engineering at Pennsylvania State University, USA. Anand's research interests are in computer systems design and evaluation and in leveraging information technology to solve real-world problems using data from the field. He has published over 250 research papers in highly prestigious conferences and journals. He has served on the organizing and program committees of several conferences and on the editorial boards of journals. He is an ACM Fellow and an IEEE Fellow. Anand received his B.Tech in Computer Science from IIT Madras in 1989, and a subsequent Ph.D. degree, also in Computer Science, from Georgia Tech in 1995.

Seamless Integration of 5G

A cognitive framework for dynamic network provisioning

Hemant Kumar Rath,
Bighnaraj Panigrahi, et al.

IN BRIEF

Industry 4.0 is emerging as an important industrial revolution, bringing a variety of use cases which can enable customization of services and products, exponential business growth, and ease of operations. Their realization can only be made possible with the support of fifth-generation cellular (5G) networks. While 5G brings disruptive changes, its coexistence with other networks is a reality. Therefore, seamless network provisioning with real-time network predictions, in a manner that is vendor agnostic, is crucial. In this chapter, the Cognitive Controller Framework is introduced. It brings prediction, adaptation, and management of the network and applications based on the business requirements in an autonomic manner.

The success of the entire process of "sensing to actuation" depends on various factors, such as the type of applications or sensing, the kind of underlying network, QoE, or KPI requirements, and real-time provisioning of the network

The fourth industrial revolution or Industry 4.0 [1–4] is enabling customization of products, flexibility in product design and life cycles, fast decision-making, and efficiency of the production process. This requires distributed decision-making and automation of the entire industry process. This, in turn, requires reliable exchange of information, be it between the sensors/smart devices/controllers and/or actuators. With this revolution, Internet of Things (IoT)-enabled devices are set to rule the industries, irrespective of whether it is a single process or an entire industry. Consequently, this brings opportunities in terms of design and planning, deployment and operations, seamless integration, privacy and safety, and decision-making processes, among other areas.

The seamless integration of *things* with networks and applications requires a thorough understanding of physical attributes, the use of intelligent and adaptive interfaces, real-time decision-making processes, and real-time actuations. This requires a full-fledged and robust network which should ensure high reliability, availability, latency guarantee, and high performance at a low operational cost. Besides these requirements, such a network also necessitates computational resources at various levels of

Fact File

TCS Research: Network-Embedded Infrastructure–Communication

Outcomes: Cognitive Control Framework which can be integrated with enterprise networks or telco networks to provision Industry 4.0 use cases

Principal Investigators: Hemant Kumar Rath

Techniques Used: Microservices, Cognitive Learning

Industries Benefited: Enterprise Networks and Telcos

Patents: Unique patents Filed: 1 Framework and 13 Supportive Patents; Granted: 15 Grants in Different Geographies; In Process: 3

Papers: 10 Conference Presentations

the network with a focus on the edge of the network. This can only be achieved with advanced compute, communication, and data management techniques, or through a control framework.

There are several use cases both from a business and from a research and innovation (R&I) perspective. Broad areas include smart warehouses and warehouse automation, digital manufacturing, enterprise solutions, smart retail and supply chains, and intelligent transport systems, among others. These bring significant opportunities for service industries in terms of system integrations, design and deployments, transformation of the existing operational setup, security and privacy enhancement, and seamless operations, among other areas. For all these, the basic requirement depends on a robust, always available, highly reliable, and flexible network. Though there are existing network and communication technologies, such as fixed wireless, cellular networks with 2G/3G/4G operations, optical fibers, and others, Industry 4.0 applications need the support of 5G communications in the process of their realization [5].

5th Generation (5G) Communication

5G is envisioned to be a network technology which promises to deliver extremely low latency (1–10 milliseconds), high reliability (99.9999%), and availability, in addition to gigabit/sec rate of transmission, virtualization, and softwarization. 5G enables a multilayered communication service framework spanning connectivity, platforms, integration, and operations. It is a combination of heterogeneous technologies such as New Radio (NR), Long Term Evolution Advanced (LTE-A), Wi-Fi, LTE Licensed Assisted Access (LAA), Optical Fiber, Fixed Wireless, Wired Broadband, and others.

At the connectivity level, it involves offering several connection options, programmable network slicing (to provide customized network services), mobile edge computing,

The Cognitive Controller is based on an innovative Sense–Analyze–Decide–Response (SADR) framework approach

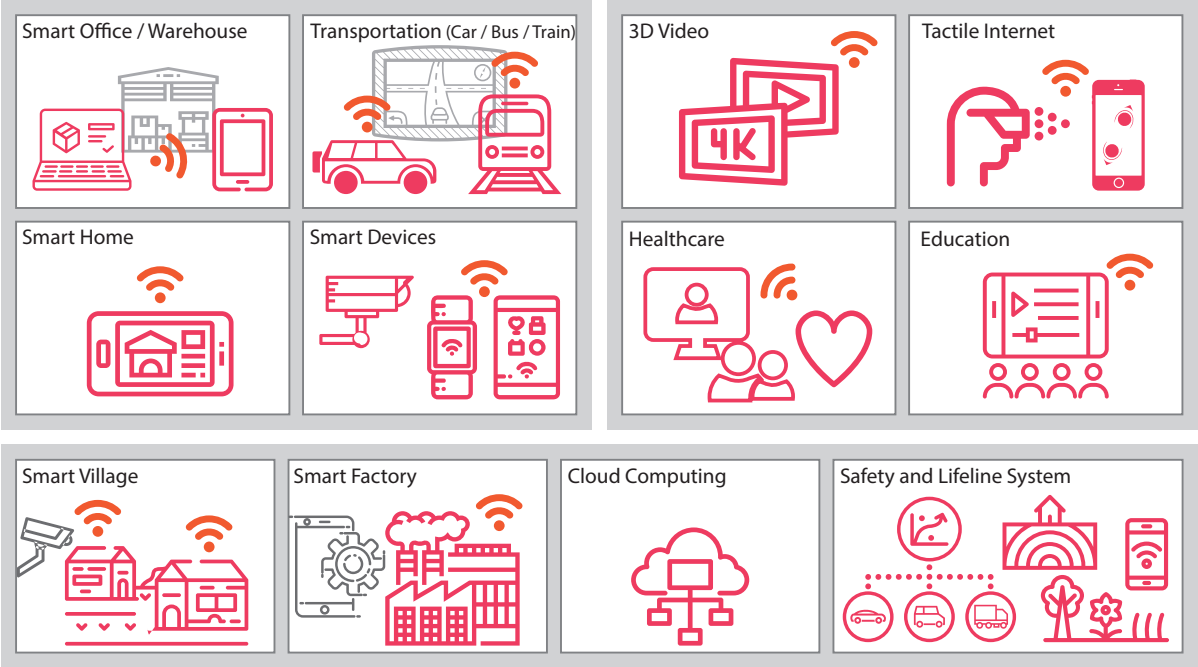


Figure 1: *Smart-X Use Cases with 5G*

and more. 5G is expected to play a significant role in the broad areas of Smart-X: smart villages, smart agriculture, smart cities, smart warehouses, smart industries, smart transport, smart buildings, and much more (see Figure 1) [6].

In all the Smart-X use cases discussed above, data or information related to the system state, real-time camera feeds, mobility information, environmental condition, and surveillance, among others, pose a major challenge. For example, in use cases such as smart warehouses/ retail or digital manufacturing, robots and drones are deployed to sense or collect the information through various applications which has to be communicated to a central location (or a cloud) or to distributed locations including the edge or fog nodes. The data needs to be analyzed before a decision is made or generating the next set of instructions to the robots or drones. The success of the entire process of "sensing to actuation" depends on various factors, such as the type of applications or sensing, the kind of underlying network, Quality of Experience (QoE) or Key Performance Indicator (KPI) requirements, and real-time provisioning of the network.

For a sustainable and QoE/ KPI-assured service model, the usual mode of operation has to be changed to a *resource-aware* or *network-aware* mode. Hence, there is a need for a platform or a framework which can bring network awareness to the applications and application awareness to the network managers. Though 5G provides technology for network provisioning according to different applications' KPI requirements, it does not clearly define the "what" and "how" parts of it.

The Decide module comes up with online/ offline rules on network provisioning or application configurations which are executed through the Respond module

No comprehensive methodology for network-aware application provisioning is also available. Moreover, applications and networks are treated in silos but not as a single entity, which is a major drawback of the current system.

TCS is focused on developing a framework which can be used to dynamically provision the network based on the application need (i.e., "sensing to actuation"). The underlying network is assumed to be the heterogeneous 5G network. While the initial focus is concentrated on regular high-speed Wi-Fi access networks (including Wi-Fi6), integration with 5G (NR) is planned for the future.

The Future

The proposed cognitive controller through the SADR-based approach brings a seamless experience for users, not only for the existing use cases but also for the upcoming ones. While the current framework supports Wi-Fi-based enterprise use cases, integrating other enterprise network technologies such as Wi-Fi 6, private 5G, light fidelity (LiFi), or visual communication, among others, the microservices-based architecture will be extended to support Telco use cases, where miniaturized versions of the Cognitive Controller can be deployed over a distributed fog/edge setup to support large-scale networking. To this end, the Cognitive Controller framework is being integrated with several application programming interfaces (APIs) and network interfaces for enterprise use cases, including Telco use cases.

Cognitive Controller using the SADR Framework

TCS has designed a cognitive network and application controller framework applicable for Industry 4.0 use cases. It provides multidomain service orchestration spanning admins, networks, and user domains; automated analytics; and end-to-end fault, configuration, accounting, performance, and security support (FCAPS). It is based on an innovative **SADR** framework approach.

The modules of this framework are deployed through a microservices-based architecture. The **Sense** module senses the application and network parameters which are analyzed through the **Analyze** module. Information related to the analysis of the usage of devices and dynamics of the network is used to adapt application characteristics in terms of measurable QoE parameters intelligently. Machine learning (ML) and other techniques are used to understand the current system state and also predict the future system state. Appropriate rules are formulated which are fed back for adaptive data filtering, dynamic network slicing, and network provisioning, among other purposes. This also directs the applications to decide what needs to be communicated under the predicted circumstances.

The **Decide** module comes up with online/offline rules on network provisioning or application configurations which are executed through the **Respond** module. In case the decision rules are not executed through the Respond module (due to admin or security or any other reason), they can be passed to the admin through a prescriptive method. The rules used in the Decide module are framed through an integrated emulation platform and/or through learning-based artificial intelligence (AI) techniques. At present, open-source discrete event emulator NS-3 is also integrated with this framework.

Figure 2 illustrates the broad architecture of the proposed SADR-based cognitive controller framework. This framework can be used for various use cases. For example, in robotics and drone-based applications, sensing can be performed by robots or drones and the sensed data can be communicated over the legacy network or over 5G (Wi-Fi/LTE-A or similar) to a central platform, where the Analyze and Decide modules operate. Based on the need of the applications and the network, appropriate actuation commands are passed to the robots and drones while appropriate network provisioning in terms of handoff, end-to-end bandwidth, or channel management/slicing, among others, are passed to the underlying networks.

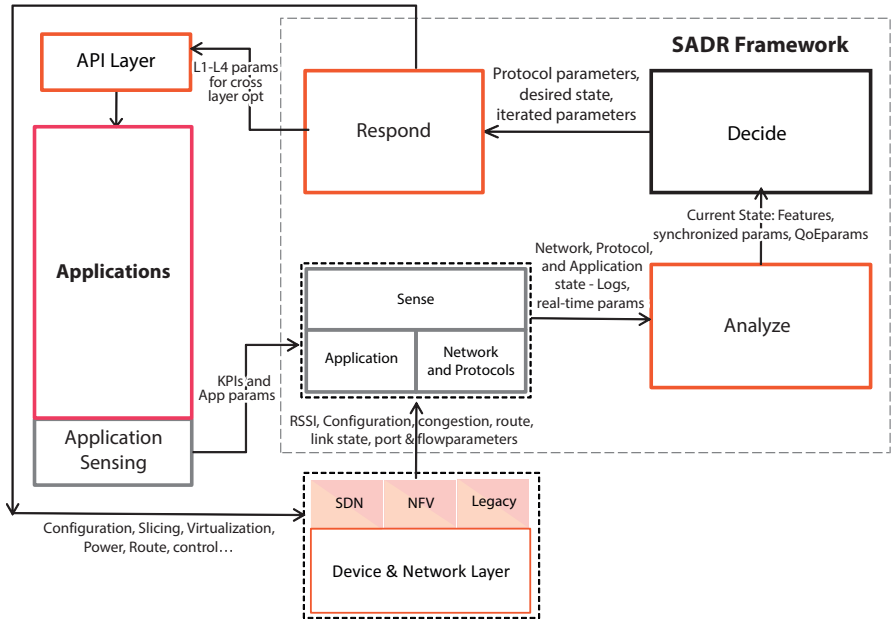


Figure 2: *SADR-Based Cognitive Controller*

Cognitive Controller Deployment — Microservices-based Framework

The microservices-based deployment architecture helps in enabling a highly flexible framework where different services are developed to support the Sensing, Analysis, Decision, and Response modules as listed above. Microservices are being deployed over a middle-layered platform. It supports software-defined networking (SDN), legacy network options, and network function virtualization (NFV). The services developed here can be plugged-and-played to design use case-specific solutions as and when required.

Further Reading

1. "Industry 4.0: The Industrial Internet of Things", Alasdair Gilchrist, ISBN-13 (electronic): 978-1-4842-2047-4, 2013.
2. Internet of Robotic Things, Carnegie Mellon University, https://engineering.cmu.edu/news-events/news/2017/06/16-internet-robotic-things.html.
3. "Network based communication for Industry 4.0", Discussion Paper, Federal Ministry for Economic Affairs and Energy (BMWi), 2016.
4. Dr. Andreas Müller, Robert Bosch GmbH 2015, Contribution to the discussion session on "Radio communication for Industrie 4.0" on May, 28th, 2015 at the ITG-expert committee 7.2 radio systems.
5. View on 5G Architecture (5G PPP White Paper), Ver. 2, Jul 2018 (www.5g-ppp.eu).
6. B. Panigrahi, H.K. Rath, R. Ramamohan, A. Simha, "Energy and spectral efficient direct Machine-to-Machine (M2M) communication for cellular Internet of Things (IoT) networks", IEEE International Conference on Internet of Things and Applications (IOTA), 2016.

Hemant Kumar Rath

Hemant is a Senior Scientist at TCS Research & Innovation, Bangalore. He has more than 20 years of experience in academics, research, and industry. His research interests include QoS in Networks, 5G Network Design, Software-Defined Networks, and ICN, among others. Hemant holds an M.Tech. and a Ph.D. from IIT Bombay and is a Senior Member of IEEE and IARCS. He has published over 50 research papers and presented many talks at national and international conferences and has filed over 35 patents in the broad areas of communication and networking. He is also active in national and international standardization in the areas of networking and communications. Hemant is also an Adjunct Faculty at IIT Bhubaneswar.

Bighnaraj Panigrahi

Bighnaraj is a Scientist at TCS Research & Innovation. He is working on various projects in Networks such as ICN, 5G communications, Cloud Robotics, and Tactile Internet, among others. He has research experience of more than 10 years in different fields of wireless and next-gen networks. Bighnaraj holds a B.Sc. degree in Physics as well as an M.Sc. and an M.Tech in Electronic Science and Computer Science, respectively. He obtained his Ph.D. from IIT Delhi where he worked on cross-layer optimization problems for Wireless Sensor and Ad-hoc Networks. He has over 35 publications in international conferences and journals and has several granted patents to his credit.

Samar Shailendra

Garima Mishra

Balamuralidhar P

Arpan Pal

Reading Images

Information extraction from document images using deep learning and program synthesis

Lovekesh Vig

IN BRIEF

With the increased availability of phone cameras and scanners, it has become common for important documents to be digitized as scanned images. While these can be digitally stored, to make the information usable, the documents must be made machine readable. TCS' Deep Learning group has looked at improving visual perception and knowledge extraction aspects of the problem. We present a glimpse of its solution approaches. These include Deep Learning based vision APIs, Cycle GANs for noise removal, text entity identification, knowledge graph construction and neuro-symbolic program synthesis to extract textual information from scanned images.

Non machine-readable documents are a nightmare for modern businesses. Contracts, bank guarantees, trade finance notes, scanned ID cards, claims, and many such documents come as handwritten text, images, or PDFs that cannot be edited, mined for insights, or stored in an easily searchable fashion. Using these involves cumbersome manual digitizing processes and increases risk to businessess. Information extraction from scanned documents is still an open problem. TCS' Deep Vision team works on a variety of vision problems toward data extraction from images processed using deep learning techniques. Here, we discuss some problems and methods of approaching them.

Complexities in Documents

While OCR accuracies have significantly improved due to advancements in deep learning, these alone are insufficient for effective extraction of visual information from scanned documents. Most documents have a rich set of visual entities in the form of tables, text boxes, blocks, charts, and arrows. Until recently, vision algorithms were not powerful enough to accurately identify and extract these visual entities. This resulted in the propagation of errors downstream in the extraction process.

Real-world documents often have a combination of both handwritten and printed texts, which makes

Information extraction from scanned documents is still an open problem

Fact File

TCS Research: Deep Learning and AI

Outcomes: DeepReader Platform for Information Extraction from Image Documents

Principal Investigator: Lovekesh Vig

Academic Partners: Prof Ashwin Srinivasan, BITS Goa

Techniques Used: Inductive Logic Programming, Deep Learning, Knowledge Graph Mining

Industries Benefited: All verticals that use scanned documents for their business processes such as Finance, Manufacturing, Utilities, Insurance, and Healthcare

Patents: 8

Papers: 10+

A pressing challenge is determining the correct template in which to classify the test document; a deep Siamese network can be used to achieve this

text harder to localize and identify. Moreover, the scanning process can often introduce noise (glare/ distortion/stains) in the documents which confuses the text recognition algorithms. Therefore, any real-world deployment has to address these vision challenges in order to be effective.

Improving Visual Perception

Denoising

Camera shake, improper focus, imaging noise, coffee stains, wrinkles, low resolution, poor lighting, etc., introduce noise in scanned images. Denoising can be achieved using generative adversarial networks (GANs) that have been applied to different image-to-image translation problems, such as super resolution, style transfer, and product photo generation.

Document identification

A pressing challenge is determining the correct template in which to classify the test document. A deep Siamese network can be used to achieve this.

Processing handwritten text

A handwritten text recognition (HTR) system, which uses a convolutional feature extractor followed by a recurrent encoder–decoder model for mapping the visual features to a set of characters in the image, may be used to digitize text characters from handwriting.

Extracting Document Entities

To preprocess a document for information extraction, a number of visual elements need to be identified. Deep vision APIs may be leveraged in order to identify these elements.

Page lines

To read page lines, a horizontal clustering of the words based on the Euclidean distance between connected component clusters may be used. After localizing page lines, each page line patch can be sent through a standard OCR engine (such as Tesseract, Google Vision, or Abby FineReader) to get the text corresponding to that line. Sending smaller patches results in higher accuracy from the OCR engine.

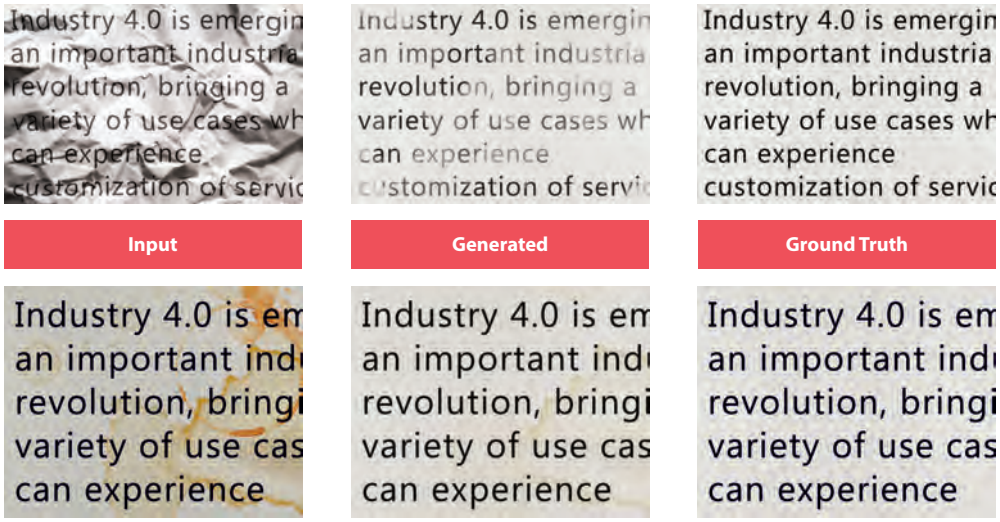


Figure 1: *Sample outputs of the denoising process*

Text block

A text block is a set of lines which begin at approximately the same x coordinate and the vertical distance between them is not more than twice the height of the line. The lines that lie inside a particular text block or box may also be identified separately as block or box lines.

Boxes

To read boxes, the image can be first eroded followed by thresholding and inversion. After that, a comparison of the area of each connected component with the area of its bounding box can be made. If the area of the connected component is within a percent of the area of the bounding box, then it can be deemed that the connected component is a box.

Creating a Knowledge Base

In addition to addressing vision problems, there are challenges involved in understanding the complex visual structure between the entities in the document. The visual relationships between the different entities detected in an image are often critical to understanding the information present in the image/ document. For example, a text label might only make sense if viewed in the context of the entity it is connected to by an arrow. The spatial relationships between entities can be recorded and populated as a high-level relational schema. The relational schema must be designed as generic and exhaustive as possible.

Building schema

Once all the entities are identified, relations between the entities need to be populated and stored in the database. The corresponding database schema should be designed to facilitate subsequent information extraction. All the entities are associated with their spatial coordinates, and this information conveys the whereabouts of the neighboring text entities.

This information may then be used to infer different logical

The spatial relationships between entities can be recorded and populated as a high-level relational schema

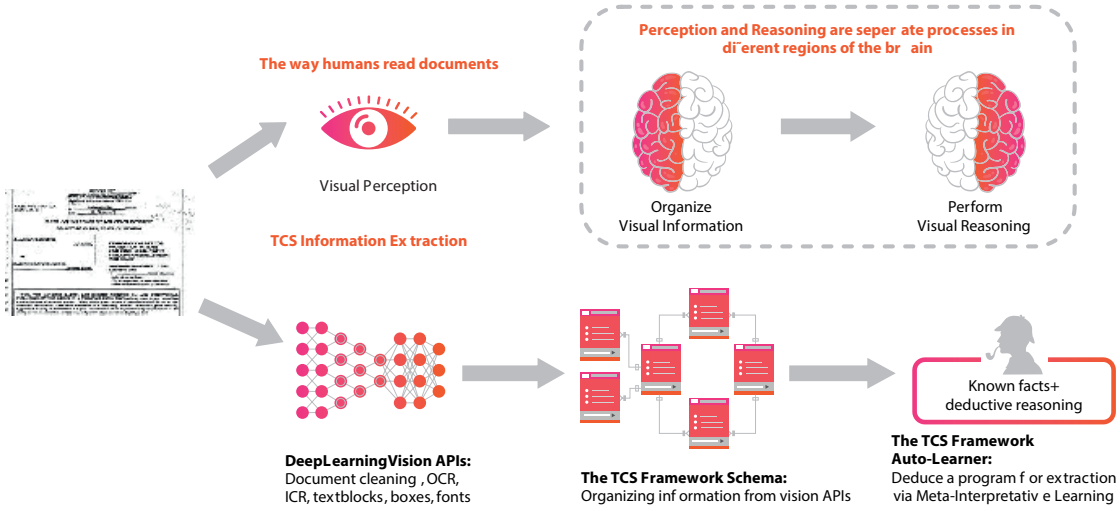


Figure 2: *TCS Information Extractor*

and spatial relationships. The line entities identified from the vision components will include information about the line text, individual words in the line along with the line and word coordinates. Using the coordinate position of words, extraction of words to the left/right/above/below of other words may be made and maintained in the schema.

Similarly, raw vision data to maintain the words and the lines in which they occur in the schema can be a separate table.

Incorporating universal and domain-specific background knowledge

Real-world documents use some word variations that humans understand easily, but machines find difficult to parse. These can relate to address fields where we can differentiate a zip code from a phone number: abstractions such as 'St.' for street, semantically similar but differently abbreviated words such as "amount"/"amt," commonly occurring text pairs separated by a punctuation such as "SWIFT: SCBLUS33," etc. Universal background knowledge and domain special background knowledge need to be incorporated into the system. The APIs for correcting OCR spelling mistakes may also be added to make extraction cleaner.

Fetching the Data

Once the relational schema is populated with data from the document, it can then be queried like a regular database schema using SQL. Thus, the original problem of information extraction from images has been abstracted to the problem of simply querying a set of tables. This allows for a wider audience that may not be familiar with vision but is familiar with SQL to specify extraction rules/steps. While querying in SQL certainly makes the information extraction solutions more accessible to a wider range of programmers, it may be possible to reach an even wider audience if the system had a conversational interface, obviating the need for familiarity with SQL. The user, when interacting through

Raw vision data to maintain the words and the lines in which they occur in the schema can be a separate table

natural language can ask for a variety of different fields. Things may be further simplified if an intent identifier is built to classify the NL-Query and ascertain the intention which may be an extraction query, a request for creating or saving a workflow, book-keeping, etc.

Future Work

TCS' Deep Learning team has built a proprietary framework for the extraction of relevant data from image documents using the said approaches. It handles vision challenges involved in processing real documents such as removal of background noise and handwritten text recognition with a GAN-based architecture. Spatial relationships between visual entities are captured via a rich predefined relational schema. A Siamese network is used to identify the template of a document. An NL interface has been integrated to allow novice users to converse with the system for rapid and easy data extraction.

The current TCS framework can be used to process a wide variety of real-world documents that adhere to a finite number of templates; however, several enhancements are required for better generalization. Organizing universal background knowledge by developing a framework for prioritizing rules for a particular context and allowing for easy interoperability across domain ontologies for easy integration of domain-specific knowledge are part of the future plans.

An aspect that is being explored is learning the rules for extraction automatically via examples. This would allow users to specify the extraction fields automatically without writing down elaborate extraction rules. The underlying philosophy is that while the innate human ability to see and discriminate objects in the real world is not yet fully understood, the ability to read and understand documents can be performed via a combination of background knowledge and visual/textual cues. With the progress of deep learning in being able to capture the visual elements more accurately, perhaps document understanding can now be accomplished to a much more satisfactory degree.

TCS Framework for Information Extraction from Document Images

Trade finance, a critical business operation for banks, encompasses processing multiple activities including payment and collection, domestic as well as international guarantee, letters of credit, etc., for imports and exports of goods and services across borders between stakeholders. The transaction also happens across different banks. In short, the trade finance life cycle is a complex process. For the purpose of payment, all vendors have to be onboarded to the system. Furthermore, these documents are also required for future references.

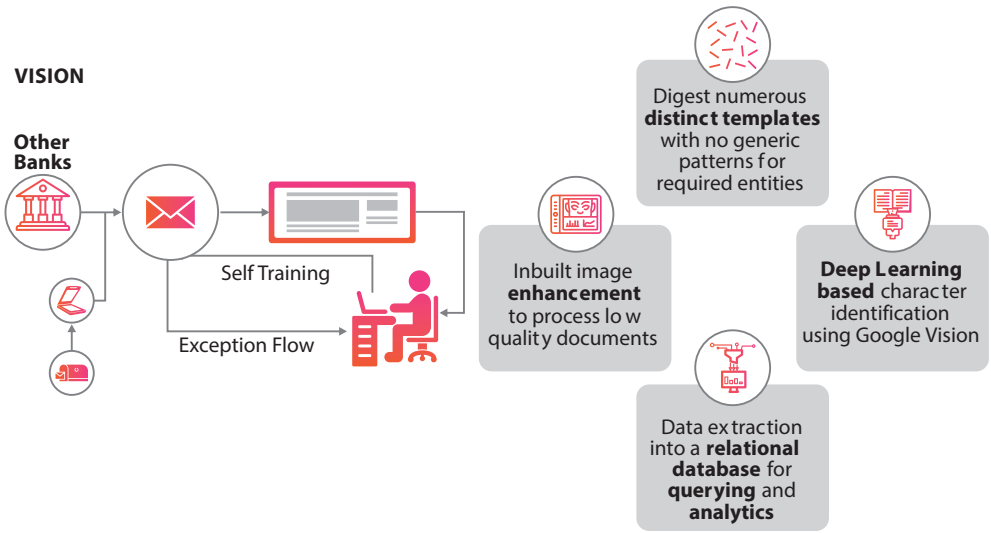
One of our customers, a European bank, receives several documents for processing trade finance payment instructions on a daily basis. These include scanned images of documents, which include handwritten text. As per the bank's internal process, information such as records of goods shipped, place of shipping, consignment value, and details of receipt of goods had to be extracted and stored digitally for processing and for future reference. The bank was extracting information manually. While it ensured some level of accuracy, it was slow and manpower intensive. The bank was using AWS vision API for extraction. However, there were some limitations. While it was able to recognize text with clear templates where fields appeared in the same format, the tool failed when the fields exhibited variations. For example, if the field was shown as "sender/seller" or if the location of the field varied depending on the document or when the value of a field was either a text or a numeric or an alphanumeric, then extraction became a challenge for the bank. Additionally, it could not support handwritten text, noisy backgrounds, watermarks, boxes, text overlaps, disconnects between text and images, and other visual complexities.

Enter TCS Document Image Information Extraction Framework

TCS suggested the use of its IP-based framework along with the existing tool as it would complement the functionalities. With deep learning-based vision APIs and novel DB schema, the TCS framework would be able to extract information from text images efficiently. Accordingly, a pilot was run for extracting payment elements to create a SWIFT payment message using the TCS framework. It was trained to extract relevant fields using scanned documents that the bank provided. Furthermore, the TCS framework was also configured with the business rules of the bank.

During the pilot, the TCS framework was trained on more than 30 documents from multiple banks. Accuracy improved with each cycle of training. For each document, 13 fields had to be extracted. These fields varied in title and location in a document.

Automating generation of SWIFT messages for payment from trade finance related scanned images



As technical components built into the tool included an image enhancer, an image-to-text converter, relationship builder (TCS IP), template identifier, extractor, trainer, the TCS framework was able to capture relationship of fields.

Key benefits

- **Increased accuracy levels:** In the pilot phase, 85% fields for all documents were extracted correctly by the TCS framework, achieving higher accuracy for the test set.
- **Automated processes:** Different banks have varied form structures. The TCS framework helped in extracting and collating information into neatly labeled and structured columns. This allowed automating payment instruction generation quickly.
- **Enabled future analytics and insights:** As complete documents were processed and stored in the form of neatly labeled and structured columns in the DB, the DB became a source of value for any future analytics and for drawing critical insights.
- **Standardized processes**: The pilot enabled the bank to capture business rules and standardize them so that they could be implemented uniformly across the bank.

The TCS framework was trained with the bank's documents and templates, the knowledge base was broadened using the bank's business and validation rules, and the inbuilt technical components of the tool-enabled relational schema. This was done with the help of TCS' Digital Unit, BFSI NorthWest Europe. A combination of all these features resulted in enhanced accuracy with every cycle of document extraction leading to higher efficiency levels in the bank's trade finance business process operations.

Inputs from Chandershekher Joshi

The TCS framework was conceived in collaboration with Prof. Ashwin Srinivasan from BITS Goa and Dr. Gautam Shroff, with valuable inputs from Rohit Rahul, Monika Sharma, Vishal Sundar, and Vishwanath D.

Lovekesh Vig

Dr. Lovekesh Vig is a Senior Scientist at TCS Research and Innovation and leads the Deep Learning research area. His principal research interests are in neuro-symbolic integration, meta-learning, machine learning for optimization, and information extraction. Lovekesh's research group at TCS Research works on developing AI solutions to enterprise problems at the intersection of Vision, NLP, Robotics and Sensor Analytics. Lovekesh completed his Ph.D. in Computer Science from Vanderbilt University.

Sensing for Health

Of machines, materials, and people

Tapas Chakravarty,
Avik Ghose, et al.

IN BRIEF

In the era of smart machines, even real time insights are late. Today, predictive and prescriptive are the operative words. While human expertise can perceive symptoms of disorder in systems – men, materials, or machines – to an extent, it is subjective and inconsistent, and more importantly, usually occurs after a certain amount of damage is already done. Simply put, human perception is just not "predictive" enough.

Sensor fitted machinery generates vast amounts of data which often goes to corroborate observations from human perception. But curiously enough, there are still spaces that neither human perception nor sensors can reach easily. These spaces have been of special interest to our team. In this article, we have highlighted how we enable better and more reliable decision making to interpret health of systems.

In the era of smart machines, even real time insights are late. Today, predictive and prescriptive are the operative words. Human expertise can perceive symptoms of disorder in systems – men, materials, or machines – to an extent. A plant operator can sense an anomaly in the vibration signature of, say, rolling element bearings in a gear transmission system or sniff out the distinct rotten-egg smell that signifies a leak or oxidation or corrosion in machinery. A food taster can sample a dish to determine its quality and taste. A doctor can understand a patient's condition to a certain extent through observation. But human perception is subjective and inconsistent, and more importantly, usually occurs after a certain amount of damage is already done. Simply put, human perception is just not "predictive" enough.

Sensors, which take the form of tiny electronic pieces, have been changing the game over the last few decades. The machinery in new industrial setups is often equipped with sensors and legacy systems are retrofitted with physical/soft sensors wherever possible. These sensors generate vast amounts of data which often goes to corroborate observations from human perception. Together, they offer keener insights.

But curiously enough, there are still spaces that neither human perception nor sensors can reach easily. These

Fact File

TCS Research: Embedded Sensors and Devices

Outcomes: TCS Unobtrusive Sensing Platform, TCS Human Sensing Platform

Principal Investigator: Tapas Chakravarty

Academic Partners: Indian Institute of Technology, Kharagpur, and Singapore Management University

Techniques Used: Radar/Wearable/Mobile-based sensing for human location, Activity and Physiology detection, RF/Ultrasound/ Optical sensing, Metamaterial-based sensor design

Industries Benefited: Manufacturing, Healthcare & Life Sciences, Smart Cities, Retail, TTH, Government, Banking, and Insurance

Patents: 33

Papers: 40

spaces have been of special interest to our team. Our vision is to enable human decision-makers to better interpret the health of systems, so that decision-making becomes easier, more reliable, and consistently effective.

The "Sensing Concept Map" (Figure 1) provides an overview of sensing modalities and inferences drawn from the same (via the additional information they provide) for augmented human perception of humans, materials, and machines. The objective is to detect their activities/operational states/defects in their natural habitat, sense subtle variations in behavioral patterns,

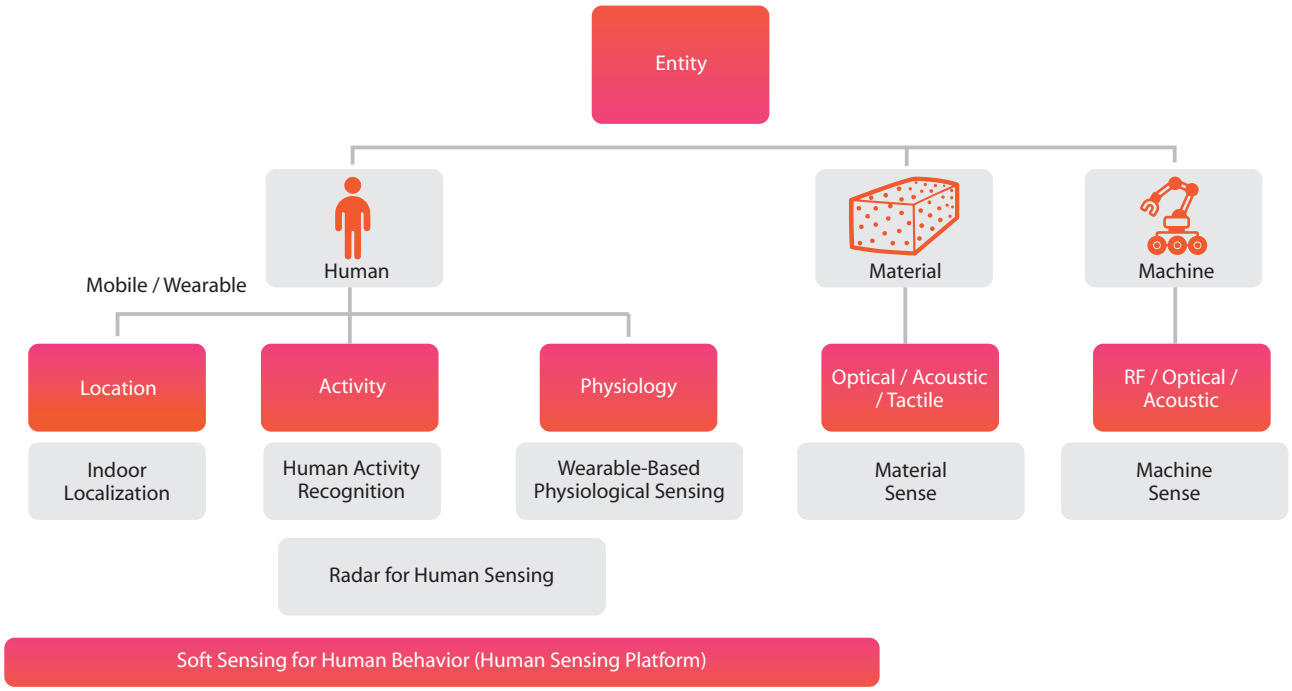


Figure 1: *Sensing Concept Map*

draw rich insights, and enable accurate predictive models.

Human Sensing

Human beings are fundamentally interested in learning more about themselves. Health is a prime concern. Multiple sectors focus on human health: pharma, healthcare, insurance, and the government, to name a few. With sensors, organizations operating in these spheres can now help get a better view of our health, movements, and activities in order to predict disorder and prescribe remedies.

To understand human parameters as well as the human context, our team employs a variety of sensing mechanisms to observe human location, activity, and physiology among other relevant factors. For location sensing outdoors, we primarily rely on GPS or A-GPS. For indoor scenarios, we have actively worked on indoor localization using personal devices. Here we rely on Radio Frequency (RF) transceivers present in devices such as Wi-Fi and Bluetooth Low Energy (BLE) peripherals. For indoor location sensing, we created a novel radiowave propagation path loss model which can locate a person with high precision. For activity recognition, we relied on inertial sensors like accelerometers and gyroscopes to understand human gestures, posture, and gait-based activities. We realized that the barometer sensor helps provide additional context to recognize activities like sitting, standing, or falling. We have also looked at physiological parameters like fatigue, heart-rate, breathing-rate, and the breathing power of individuals using wearable devices.

We wanted to obtain human parameters in sizeable detail while letting the subject lead her daily life in her natural setting, in order to remove the element of "observation

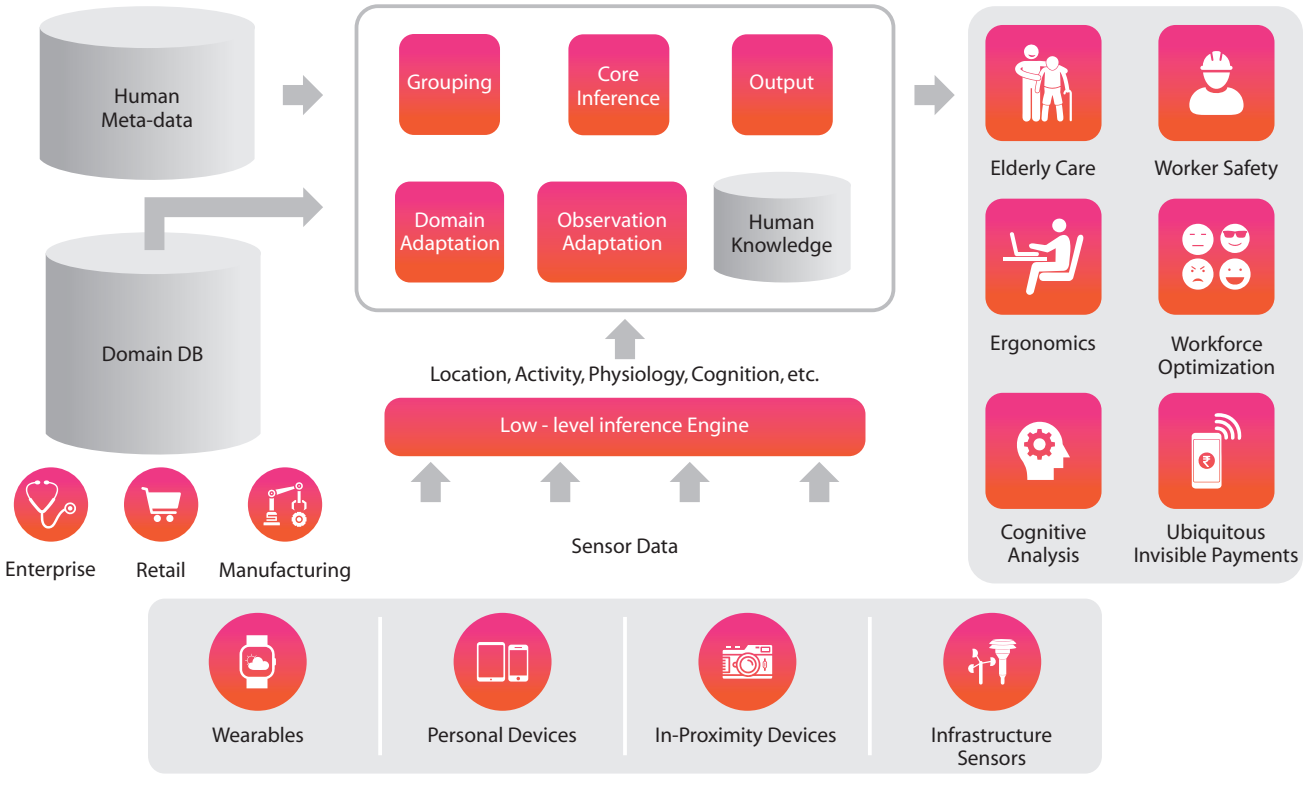


Figure 2: *Human Sensing Platform*

bias" from human behavior. Here, it becomes clear that electromagnetic imaging-based human parameter elucidation is the future holy grail of human sensing. A step forward in this direction led us towards the design of small, reliable, energy-efficient and customizable (deployment-ready) millimeter wave (MMW) radar units for applications in our daily lives. Radar-based human sensing offers three distinct advantages:

- The ability to sense at a distance
- Preservation of privacy
- Insulation from the effects of ambience (illumination, temperature, and fire among other factors)

Multiple small-sized, networked radar modules can be spread across an apartment so that the occupant's location, movement, and physiology are seamlessly recorded at every moment.

We have used these low-level inferences from human sensing to infer a higher level of human knowledge, of which a few case studies are illustrated below. In this respect, we have envisaged a human sensing platform as shown in figure 2.

The human sensing platform allows us to use sensor data to draw low-level inferences related to location, activity, physiology, and cognition. This can be further augmented with sensor-driven insights into the psychological state (example GSR, EEG, ECG). Hence, when psychology, location, and activity are correlated using an algorithm, it can yield a better model of the mathematics of human behavior. Domain adaptation is used for this purpose along with the knowledge base of

Active Assisted Living

Active Assisted Living (AAL) projects have humans living independently or aided by a smart environment. We worked with these environments to monitor human subjects using ambient sensors and intelligence. However, AAL also brings in the paradigm of actuation, which in turn leads to new challenges in the domain. When a robot and a human coexist in proximity, the robot needs to have a deep behavioral understanding of the human to effectively interact with her. This becomes more challenging in AAL scenarios as the subjects (such as geriatric subjects) are often cognitively challenged.

The scope of such research work is to continuously, but non intrusively, monitor the subject using wearables and infrastructure-based sensors, and then make intelligent inferences on the wellness, safety, and behavioral parameters of the subject to create a deep model of the person. Such deep models can be used to study the wellness, cognition, and emotion of the human subject. In our current research, we concentrate on monitoring the elderly. However, the insights gained from the study can be generalized to other domains involving human factors and human understanding.

Driver Behavior Modeling

In this study, we try to understand the skill and aggression levels of a driver based on mobile sensor data. Every driver attempts to minimize risk through a psychologically induced "task difficulty" homeostasis—a state of equilibrium that is maintained by weighing out one's own driving skill against the demands of driving under specified conditions. In the process of navigating the vehicle through a plethora of difficult road conditions, a driver's innate ability as well as traffic psychology comes to the fore. This ability can be modeled by monitoring the driving pattern over long periods of time, thus giving rise to the driver's unique behavioral profile. Driver behavior models have many practical uses like uniquely identifying the driver as well as greatly improving road safety by inducing safe driving styles.

human nature which is derived from research in behavioral sciences.

Machine Sensing

Spectral vibration analysis of a machine offers critical information about the machine's current state. Rotating machines experience anomalies due to situations such as mass imbalance, mechanical wear and tear, and misalignments, among others. Such anomalies result in vibrations induced on non-rotating parts, noisy environments, increased harmonic distortions in power supply, and other undesirable outcomes. To address this challenge, we built a unique non-contact vibration measurement sensing system termed the "RF Guided Stroboscope". Optical stroboscopes are traditionally used for the visual inspection of vibrating parts. Additionally, if stroboscopes could extract phase information from vibration signatures, it could lead to major possibilities in prognosis. However, such stroboscopes are not cost-effective in large deployment scenarios. We tackled this issue through a smart combination of the best of both worlds, namely the stroboscope and radar. In figure 3, the schematic is presented.

This system combines two separate sensing methodologies in a unified approach for inference. Initially, the Doppler radar captures the spectrum of all vibrations in its field of view (FOV). The set of deduced frequencies is then conveyed to the stroboscope. The stroboscope, in turn, illuminates the FOV using an LED panel, one frequency at a time, thereby enabling the identification of all the parts. This system is completely automated and significantly faster than a standard stroboscope, which needs to search for the unknown frequency. Coupled with a detailed spectral analysis of vibration (amplitude and frequency of harmonics), this low-cost and non-contact system offers an elegant solution for legacy machines as well

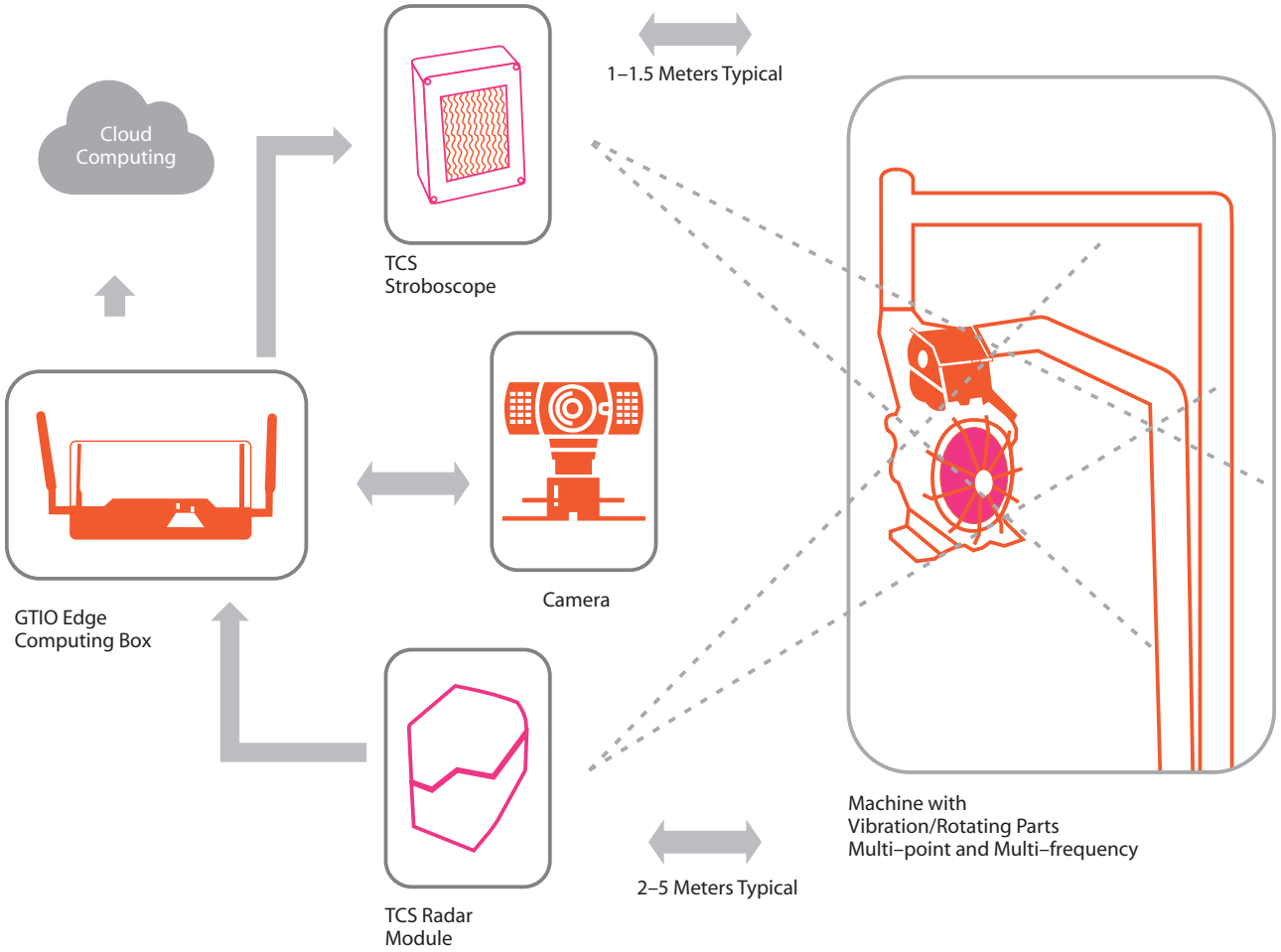


Figure 3: *Schematic of RF Guided Stroboscope for Machine Inspection*

as for extreme environments like high temperature operations.

Material Sensing

Micro-defects in materials can cause major problems. This is clearly evident during production processes in manufacturing and fabrication. For a "zero-defect" production process that ensures product integrity and reliability, defects need to be detected upstream of the actual process to prevent problems from snowballing later. Optical techniques such as infrared thermography, speckle imaging, digital holographic techniques, microscopy, and THz imaging aim to augment standard vision-based inspection approaches for inspections with high sensitivities and specificities. As part of our technology roadmap, we are putting together a portfolio of sensors comprising elements from optics, acoustics, and electronics. These elements are finely tuned for specific use cases in this context. Based on whether a large piece of machinery or a small additively manufactured component is being inspected, we use principles such as light scattering, spectroscopy, acoustics, and interferometry to develop rugged, cost-effective, versatile, multi-modal platforms.

The Roadmap

Humans need technologies that help them perceive and interact in real time. This becomes possible when their natural communication channels are augmented with synthesized information. Audio and haptic augmentation systems form

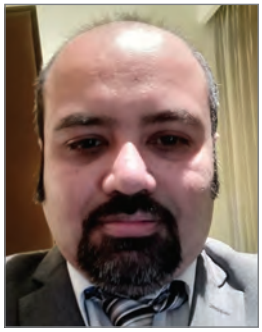
an interesting "non-graphical display" alternative to vision. We also propose to stretch the sensing paradigm from merely sensing location, activity, and physiology, to actually recognizing emotional states using emerging sensing paradigms like "Volatile Organic Compound (VOC)", which is the biochemical state of the human subject, in order to perceive actionable insights on the subject using the domain of HII. Again, with the radar modules in place, a natural extension of this concept would be to build models of human-articulated motion like gait analysis, which takes us in the direction of "person identification", a potential solution with many significant uses. Similarly, a configurable and high-precision ultra-wideband (UWB) radio frequency (RF) imaging system can act as a set of powerful "eyes" in dynamic factory environment(s), thereby enabling worker safety applications. System reliability in a dynamic environment is ensured only when the proposed system automatically adapts itself to the changes unfolding in the scene. Possibilities include the study of human range-of-motion (ROM) and activity using smart fibers enabled with opto-electronic sensing systems. Next-generation vision may include compound bionic eyes which use microfluidics-based liquid lenses.

The future is near with exciting developments in "metamaterial"-based sensor designs as well as "quantum sensing". Metamaterials (beyond materials) are artificially synthesized materials with unusual properties, not seen in our world. Since the beginning of this century, the subject of metamaterials has been intensely explored. However, we are yet to see genuine engineering applications. Metamaterial-based sensors will significantly enhance the performance of sensors while reducing their size. Quantum sensing allows us to reach the absolute limits of measurement resolution and sensitivity, and in doing so, extends our reach far beyond classical measurement strategies. Recent studies demonstrate the possibility of completely interaction-free measurements that are theoretically impossible according to classical sensing principles. Such interaction-free measurement techniques will eventually enable us to see the "unseen".

Tapas Chakravarty

Tapas Chakravarty is a Principal Scientist with TCS Research and Innovation. His major area of research is in the field of sensor design, RF/ Microwave components, wireless platforms, statistical modelling & solutions. Tapas has over 28 years of professional experience in diverse domains. He has led and executed many customer sponsored R&D projects in embedded systems.

Tapas is a Senior Member of IEEE (USA) and holds a Ph.D. (Science) from Jadavpur University. He has over 120 publications in various international journals & conferences and holds around 40 patent grants in geographies like USA, EU and Australia.



Avik Ghose

Avik Ghose is a Senior Scientist at TCS Research and Innovation. He has more than 19 years of experience in embedded systems and platforms. He is currently leading a research program on connected digital health, which focuses on providing instrumented continuum of care. His passion is to innovate around human sensing for wellness and behaviour modelling. Avik received his M.Sc. in the year 1999. He has more than 20 papers published on the subject and has also several granted patents in the field.

Prasant Misra

Parama Pal

Arpan Pal

The Who, What, and How of Voice

Preparing for the 10x increase in audio interfaces

Sunil Kumar Kopparapu

IN BRIEF

When analyzing voice, three dimensions can offer business uses—identifying the person behind the voice (who), the linguistic content in the voice (what), and the expressed emotion (how). Due to advances in machine learning (ML) and artificial intelligence (AI), automatic speech recognition today can offer good clues toward all three dimensions.

At TCS Research and Innovation, we have been concentrating on low-resource natural conversational languages in noisy conditions to push the limits of robustness with our patented voice analysis tools and algorithms. In this article, we discuss some voice-related problems that we are trying to solve. We also look at several interesting trends emerging in the area of voice and audio technologies.

The idea is to do a "Just Sufficient Analysis" to garner insights into "the who", "the what", and "the how" of speech

The most natural thing that I do when I have a problem is to speak with someone face to face and seek an empathetic answer which can solve my problem. With the societal infrastructure turning digital, there is a paradigm shift; I continue to speak, except that now, I speak to a networked mobile device or into a cellphone, with either a person or, more often, an intelligent machine on the other side!

An audio signal primarily has three dimensions associated with it in the form of "The Who," "The What," and "The How." "The Who" dimension looks at the person behind the voice, while "The What" dimension looks at the linguistic content in the voice. Finally, "The How" dimension looks at the expressed emotion in the voice. Thanks to rapid advances in machine intelligence (via AI and ML), the machine of today is able to decipher "what" is being spoken (automatic speech recognition) to a certain extent, "who" is speaking (speaker recognition) quite reliably in a constrained scenario, and "how" it is spoken (emotional state) to a lesser extent.

There are two co-existing schools of thought, both aiming to decipher "the who," "the what," and "the how" of the voice independent of the spoken language, dialect, accent, age, gender, and the calling environment of the speaker. They are

Fact File

TCS Research: Data and Decision Sciences

Outcomes: Spontaneous Audio Emotion Recognition, Noise, Robust Speaker and Speech Recognition, Real-Time Speaking Rate Monitoring, Automatic Speech Quality Assessment

Principal Investigator: Dr. Sunil Kumar Kopparapu

Academic Partners: IIT Bombay, IIT Madras

Techniques Used: Speech Signal Processing, Deep Machine Learning, Innovative Algorithms

Industries benefited: All

Patents: 30 Granted

Papers: 25

as follows:

1. Crunch huge volumes of audio data, employing deep neural network (DNN) architectures and allow machines to gain "on their own" insights.

2. Understand the way human cognition works and mimic the complex human communication process that involves the way humans interact, converse, and understand each other in real life.

Both routes are here to stay. In the short term, the trend of trying to identify the optimal set of features or using an optimal DNN architecture will see an emphasis. Sooner or later, the emphasis will be on xAI (explainable AI) and not just on how accurate the machine is able to perform a task.

As Speech Research Scientists and Engineers in TCS Research and Innovation, we have been concentrating on low-resource natural conversational languages in noisy conditions to push the limits of robustness with our patented voice analysis tools and algorithms. The idea is to do a "Just Sufficient Analysis" to garner insights into "the who," "the what," and "the how" of speech. Some specific activities in the lab are discussed here.

Voice Bots, Smart Speakers, and Speech Language Translation

A bot that can engage in small talk is considered to be "intelligent." Our focus is on building voice bots and smart speakers that can assist an enterprise to allow its customers to help themselves effortlessly. We plan to adopt variants of recurrent neural network architecture to train such a system with raw speech rather than go the conventional way of converting speech to text and then using a text bot.

Using the same raw speech data, we plan to build a usable speech language translation system that can be used by large geo-separated teams within an enterprise to communicate and work efficiently and effectively.

> The real-time identification of emotion in spoken audio not only allows for an enterprise to enhance its customer satisfaction index but also to upsell and retain customers

Robust Speech and Speaker Identification

In an Indian context, robustness to noise is crucial because of the degree or intensity of the noise levels in public spaces. We have made significant progress in terms of building novel techniques and algorithms (based on our IP) that are able to do better than deep learning-based general-purpose speech recognition engines. We look at fine tuning these techniques so that they can be used in a practical enterprise setting where speaker recognition for authentication purposes and speech recognition to enable human–machine interaction are desired. This is especially required when the speaker(s) is in a noisy environment.

Speech Analysis for Assistive Technology or Pathological Speech Processing

Societal inclusion and a growing elderly population necessitate assistive speech analysis technology for differently abled populations. We have been looking at addressing abnormal speech related to neurological and articulatory disorders. We also plan to look at addressing speech and language disorders resulting from stroke from the dual perspective of therapy and rehabilitation. We plan to build a mechanism for early detection of neuro motor disorders (such as Parkinson's and Alzheimer's) using speech analysis. From an enterprise perspective, we plan to use speech analysis to assist employees improve on their articulation and presentation skills.

Audio Emotion Analysis

We focus on identifying emotion in spontaneous speech especially in a dialogue setting, common in a call center scenario. The real-time identification of emotion in spoken audio not only allows for an enterprise to enhance its customer satisfaction index but also to upsell and retain customers. We are evolving a framework that allows robust identification of emotion.

Automatic speech quality assessment (ASQA) is a homegrown tool with powerful visualization capable of providing a drilled down analysis of spoken speech. It can be used as an unbiased systematic analysis of spoken speech for training and assessing staff who "voice face" customers (such as call center employees). It analyzes pronunciation fidelity, speaking style, and speed. It is also able to determine the stress level in the speech and the likability of voice based on the intonation and the pauses in speech. The tool has language learning applications as well. In countries where spoken English is a sought skill, this tool will allow those with limited language skills to learn the nuances of spoken English. We have built this tool as a usable, in-demand solution with current, rough-and-ready technological capabilities, rather than wait eternally for technologies to evolve to high maturity.

Voice Trends of the Future

Voice or audio processing will open up plenty of opportunities especially in the use of natural language conversational speech for human–machine interaction (HMI).

Automatic Speech Quality Assessment (ASQA) is a homegrown tool with powerful visualization capable of providing a drilled down analysis of spoken speech.

Some voice trends to be expected are as follows:

- Adoption of speech technology for enterprise self-help, especially in banks and financial institutions. Voice biometrics with "continuous verification for the entire duration of the call" will get nonintrusive and user-friendly, personalized, and be available 24*7. Enterprise self-care applications will leverage cloud computing to accelerate and automate L1 support. Environment agnostic, mixed language recognition will progress.
- Better and faster on-device and on-chip personalized speech analysis including speaker and speech recognition will see an increase. On-device processing and edge computing will enhance security and also help in restricting the possibility of voice theft.
- Voice bots and smart speakers which identify the mood and emotion of the person speaking will be in actual use and enhance user experience. The automobile sector will adopt speech analysis for in-vehicle controls and passenger safety. Local language speech processing will see an emphasis, boosting resource-starved languages.
- Real-time usable speech-to-speech translation systems for a very large set of language pairs will be available as application programming interfaces (APIs) for personalization.
- Audio surveillance to identify the mood of a crowd for security and policing of large areas will rise. Signals below 20 Hz and above 20 kHz will see an exploitation for surveillance and assisted living for elderly.
- Voice-based gaming will gain popularity.
- Speech processing on-chip for the differently abled will see progress in "text-to-speech" (TTS) aspects. Newer sensors will make "thought to speech" and "intent to speech" possible. Work in "speech for the deaf" and "speech for the mute" will be explored.
- Applications in the finance domain such as blockchain and voice for audit purposes will be integrated into speech understanding systems. Do-it-yourself learning in the form of speech therapy, pronunciation learning, and voice articulation will gain popularity. Identifying the person behind the voice will not be restricted to use of voice print alone but will evoke the use of voice temporal dynamics as an additional cue.
- One of the problems of using voice for transactional activity is the lack of privacy, as one can be heard by people around. This might lead to the use of throat microphones, wearables, and sensors which will enable "thought-based processing." However, the use of inaudible sensors can result in legal issues especially if the transaction results in an undesirable outcome due to inaudible sensors. The balance between privacy and legality must be maintained.

The future of man–machine interaction will certainly depend on the three dimensions of voice but will make use of other orthogonal modalities such as image, video, touch, smell, and hopefully some more modalities that can be captured by sensors that are yet to be invented.

> Voice or audio processing will open up plenty of opportunities especially in the use of natural language, conversational speech for human-machine interaction (HMI)

Automatic Audit of Call Center Conversations

Voice-based contact centers are often used by customers when they are in need of any information about a service or a product that they have purchased. The law of the land or the need for an enterprise to maintain a certain level of customer service requires that these call conversations be audited.

Typically, manual audits address only ~5% of the total volume of calls coming to the service desk because of the cost involved and fatigue that sets in when humans perform audits. (Auditors will have to listen to calls that are sometimes 20–30 minutes long.) Manual audit also leads to random call sampling where important calls can be missed. Human assessments are prone to human bias resulting in nonuniformity in the audit process.

Using TCS' ability to process speech signals along multiple dimensions, we have built an industry-grade and configurable call audit solution that can process a large number of calls without any human intervention. The solution also provides better insights to reliably identify customer concerns in addition to analyzing the affective content (emotion, sentiment and other factors) expressed in the call. We explain the process with a detailed flowchart in Figure 1. This solution is currently being used by a TCS customer from a pharmaceutical industry.

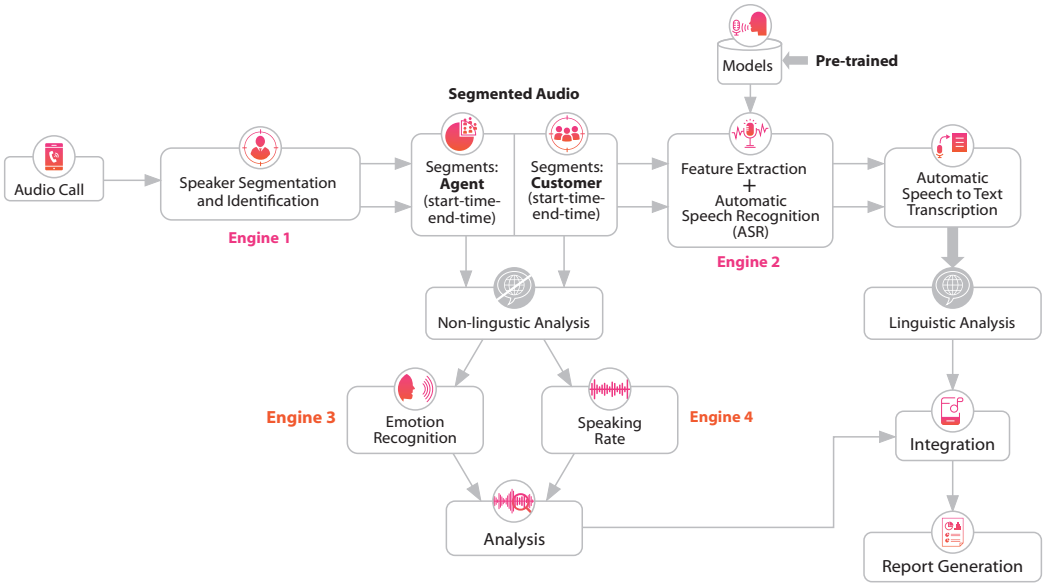


Figure 1: *TCS' Call Audit Solution Process*

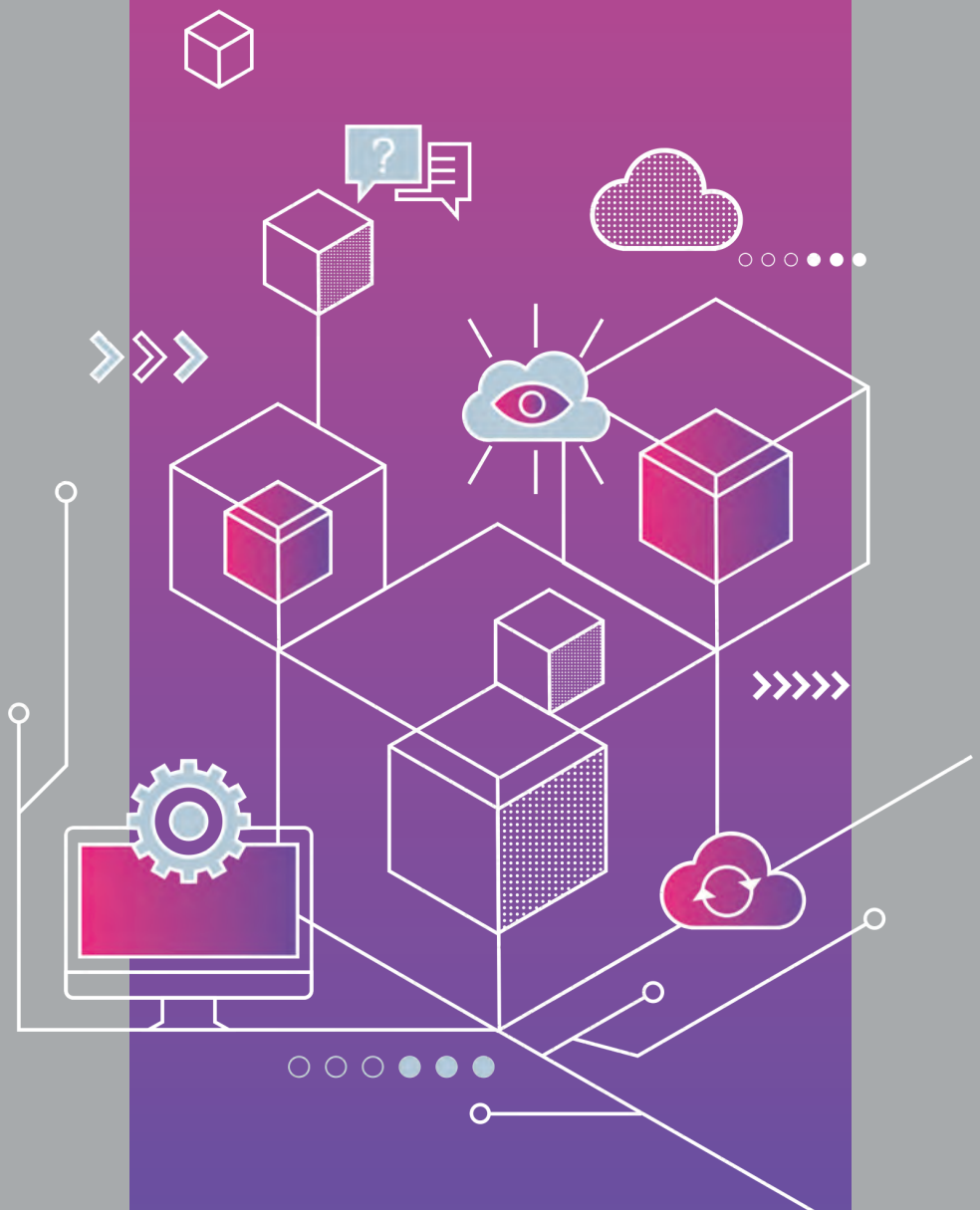
Sunil Kumar Kopparapu

Dr. Sunil Kumar Kopparapu is a Principal Scientist with TCS Research and Innovation. He works in the areas of speech, script, image, and natural language processing with a focus on building usable systems for mass use in Indian conditions.

He has authored several books. Some noteworthy ones include '*Friendly User Interfaces* (2018), *Analyzing Emotions in Spontaneous Speech* (2017)', '*Non-linguistic Analysis of Call Center Conversation* (2014), and *Bayesian Approach to Image Interpretation* (2001)'. He has several patents as well as journal and conference publications to his credit.

Sunil obtained his doctoral degree in Electrical Engineering from the Indian Institute of Technology, Bombay. He is a Senior Member of IEEE and of ACM, India.

Transformations and Computing








Research Reception

The Research Reception was a special event held to felicitate the TCS Research community — the founding fathers of Research at Tata Research Development & Design Centre (TRDDC) and all Researchers and Scientists who continue to enrich TCS Research. The event brought insights on how Research began in TCS, its contribution to TCS' business growth and development. The book *Reimagining Research,* Volume I was launched by TCS' CEO & MD Rajesh Gopinathan on this occasion .

Transformations and Computing

Arun Bahulkar

We look at the computing arena across a wide spectrum—from the "here and now" to a new land full of promise.

Covid-19 pandemic has had a dramatic impact on businesses. Digital transformations that help in adopting agile methods, collaboration, distributed and scaled up operations, areas related to overcoming changes, reliability and trust assume greater significance.

The world of DevOps has been undergoing a transformation in many ways. We are seeing organizations adopt agile methods to plan and build systems. Development and operations streams are getting better connected. System operations are now seeing increasingly more self-detection and correction of faults, and much better predictions of time and resource requirements. This has led to streamlining the operations of systems to a new level.

The article "Jile™ - To Adopt Agile, Scale, and Transform", highlights how organizations are successfully adopting and implementing various agile practices to help them transform to the new digital paradigm using a business value-driven approach.

We have all read a lot about how artificial intelligence has been used in various business scenarios. Here we present how it has been used to transform the space of IT operations. The article "AI OPs: Closed-Loop Intelligence Engineered for Change" walks us through the steps—from recognition to reasoning, and finally to the appropriate action.

The use of artificial intelligence is increasing and quickly becoming business critical. So, it is no surprise that performance, reliability, and trust have started to become major issues. The article "High Performance Computing – Building Performance Assurance Tools and Exploring Heterogeneous Architectures" presents the issues involved and our contribution to this space.

Today, AI has become mainstream and newer applications are being targeted every day. A measure of "intelligence", as represented by AI techniques, is to benchmark it against human expertise. Many high skill mind games such as chess, Go, and bridge have been the subject of AI research. As we all know, chess and Go are considered "conquered", but the card game of bridge remains elusive. Read more about it in the article "Bridge Playing Software – Vehicle to Further our Understanding of AI".

And finally, in this section we discuss the new promised land of quantum computing. All of us have questions like: What is it really? What new things will it enable? Where is it today? When will we really see it? The article "Preparing for Quantum Computing" attempts to answer these questions.

Arun Bahulkar

Arun Bahulkar is a Research Advisor at TCS. Earlier, he headed the TCS Software Engineering Research group and was involved in building several tools that have been extensively used as TCS products. He is currently eyeing the next World Computer Bridge Championship and using it as a vehicle to explore the world of knowledge-intensive systems.

Arun received his post graduate degree in Computer Science from IIT Madras in 1980 and has been with TCS since then.

Jile™

To be agile, to scale, and transform

Arumugham Prabhu

IN BRIEF

Computing is not just about hardware, software, and networks. It is also about the last mile, where effective computing applications are delivered to the user. In a fast-changing world, with consumers adopting digital technologies at a high rate, computing applications are considered to be in a "perpetual beta" stage. This implies that features and upgrades will be delivered continually. This necessitates Agile ways of working to deliver software.

Jile™ meets this challenge in a unique way. The innovation behind Jile™ is in its business model. A Business 4.0 offering from TCS, it enables user organizations to drive their transformative programs with agility and at scale.

The 2020 pandemic added prescience to the Jile™ team's vision. About three years ago, we had observed that a growing number of our customer teams were distributed. In fact, many teams wanted to "work from home" well before the pandemic. Agile methodologies worked well for small, closely knit, co-located teams. But with multiple teams which were geographically distributed, managing Agile at a program level was challenging. Jile™, an Agile DevOps platform, was born with the vision to serve distributed teams, offer flexibility and scale, and make Agile easy to adopt.

Shift to a Digital Landscape

In the last decade, organizations worldwide have seen rapid changes in the way they deliver goods and services to customers. Technology led by software and sensory devices is at the forefront of this "digital transformation". Software, being at the center of this transformation, has also seen agility in the way it is developed, released, and deployed. New software delivery methodologies were created to support such demands, and organizations are increasingly adopting the Agile way of developing software to keep up with the rapid changes in delivering to market needs.

Fact File

Patents: 6 Filed

Papers: 6

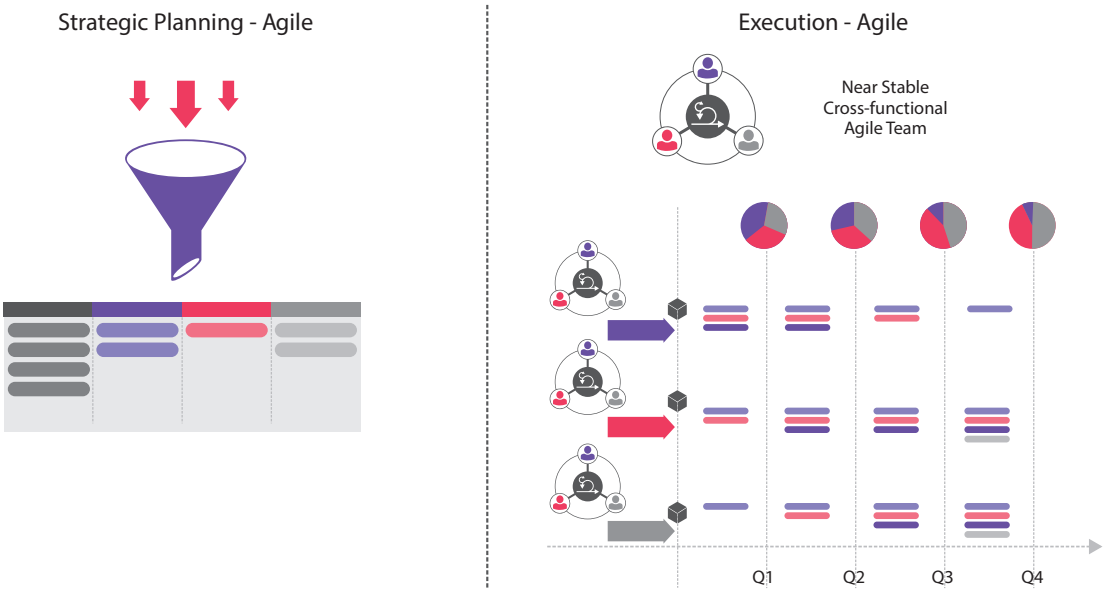


Figure 1: *Product-Centric Delivery Operating Model*

Made for the Business 4.0 World

Based on its study of industrial evolution and business response, TCS' thought leadership framework (termed "Business 4.0"), reveals that leading businesses deliver mass personalization and exponential value while embracing risk and leveraging ecosystems. Such organizations rely on cloud and intelligent automation and believe in agility.

The solution we built, Jile™, is a cloud-based platform that catalyzes agility. In terms of technology, Jile™ is founded on TCS' expertise and experience delivering large and complex projects globally.

Business Model Innovation

We observed that Agile project managers needed a solution that is self-manageable, easy to adopt, and seamless while scaling.

Based on research conducted by IDC and others, we found that existing products in the market were either positioned at the high end, which necessarily required consulting support, or at the low end, which just about satisfied team-level agility. But it was not merely the technology. We felt that innovation was needed in making the product easy to try, easy to buy, and easy to use, while delivering true Agile principles.

We conceived of Jile™ in a way that would appeal to enterprises of all sizes, irrespective of what stage of

Jile™ enables enterprises, in any stage of their Agile transformation journey, to grow and consistently evolve by making it easier to plan, develop, and deliver high-quality software, thus accelerating value delivery.

Agile practices they were in. The uniqueness of Jile™ is in its ability to allow an organization to adopt and define its own way of working. It allows it to scale as an enterprise matures in its Agile journey.

Jile™ supports product-centric delivery which is a major shift from the current projects-based delivery approach.

There was innovation in pricing too. We made Jile™ available as a SaaS product that can be procured by any enterprise on a retail (pay-per-use) basis or on an annual bulk usage basis. Further, Jile™ is the first and only product to offer a monthly units-based pricing model, which is essentially a model that is flexible in terms of usage and consumption of licenses beyond a monthly limit.

Jile™ has a full range of features. It offers a single integrated platform to define a strategic vision, set a roadmap, and realize business value through quantifiable objectives and operational measures. It supports multiple leading Agile and DevOps frameworks with a single go-to-market version in production and with continuous releases. It provides multi-level planning through hierarchical backlog management and prioritization, based on business needs and value. While enabling a real-time Agile work environment and continuous deployment, it facilitates Integrated Agile quality management.

All for Agility

Jile™ is an example of TCS' commitment to agility. The company's count of Agile projects just crossed the 10,000 mark. TCS believes in not just Enterprise Agile, but also in location-independent Agile. TCS helps its customers to embrace agility. The adoption of Jile™ is testament to that.

With inputs from Alexandra Huepel and the Jile™ team.

Arumugham Prabhu

Prabhu is the product head for Jile™. He is a Certified Scrum Master and SAFe® Program Consultant with over 20 years of experience in building innovative solutions across different industries.

His expertise is in product management, strategy, design, engineering, and deployment support. He is passionate about Agile methodology and has pioneered product development in the areas of test management, test automation, application lifecycle management, and DevOps.

Prabhu has several patents filed and granted in categories of Agile planning test automation and management. He has a Bachelor of Engineering degree in Computer Science from College of Engineering, Chennai.

AI OPs

Closed-loop intelligence engineered for change

Maitreya Natu,
Vaishali Sadaphal,
et al.

IN BRIEF

Most enterprises today strive for systematic continuous improvement of their IT systems. The effort has largely been time-consuming and ineffective, resulting in IT systems that are often found to be inefficient, expensive, and brittle.

IT systems need a real-time closed-loop solution that can augment human intelligence and are engineered for change. Incorporating intelligence in IT plants can be a game changer. Intelligence, however, needs to be of three types: recognition, reasoning, and operative. Systems need to discover hidden patterns; reason and derive inferences; and learn and adapt based on experience. On top of all these, systems should be trustworthy. We present here ignio™ as a case for intelligent transformation of IT systems.

An enterprise's business health is governed by the health of its IT. The IT organizations of all enterprises have a common objective—to drive continuous and significant improvements in the quality, cost, stability, and agility of IT-enabled business functions. However, organizations today are struggling to meet this objective. IT systems today suffer from the following key pain areas:

- IT-related outages exist and are expensive. Most Fortune 500 companies experience a minimum of 1 to 2 hours of downtime per week. The average revenue cost of an application outage in such companies is often estimated to be in millions of dollars.
- IT systems often fail to cope with continuously evolving business and technological changes. Introducing any change requires many days of planning and still carries the risk of being expensive, brittle, and inefficient.
- Operations teams lack the foresight to predict the future behavior of the system. This leads to inherently reactive operations, with most of the time spent in firefighting.
- The management of IT systems is an art form, which relies heavily on the ability of domain experts to use their past learning and domain knowledge. While an

Fact File

TCS Research: Cognitive Automation

Outcomes: ignio™, a Cognitive Automation Product, Deployed by 70+ Customers

Principal Investigators: Maitreya Natu, Vaishali Sadaphal, Rahul Kelkar

Techniques Used: Machine Learning, Artificial Intelligence (AI), and Model-driven Engineering

Industries Benefited: Retail, Banking, Finance, Telecom, Supply Chain Management, and More

Patents: 30 Filed

Papers: 15

expert-based approach has the advantage of generating usable insights and crisp recommendations, this approach cannot scale. Furthermore, such an approach is limited by the knowledge base of an expert and stumbles in new environments and unforeseen scenarios.

These pain areas can be attributed primarily to the following reasons:

- **Scale, diversity, and complexity:** An IT environment that supports a business function generally consists of thousands of different types of interdependent components, including business applications; databases; compute, communication, and storage devices, as well as people. Because of this, most organizations don't have one single person who has complete visibility, understanding, or even intuition about the end-to-end operation of a business function.
- **Slio-based approach:** Today's analysis approach is silo-based where several dimensions such as servers, networks, storage, SQLs, and applications are analyzed in isolation. Such analysis lacks a holistic understanding of the as-is state of the enterprise's IT and fails to connect business processes to applications to infrastructure.
- **Continuous evolution:** Enterprise IT systems evolve continuously to accommodate changes in business requirements and technologies. This makes the understanding of the system obsolete in a relatively small amount of time.

Because of these challenges, driving systematic, continuous improvement is time-consuming and ineffective. As a result, IT systems are often found to be inefficient, expensive, and brittle.

IT systems need a real-time closed-loop solution that can augment human intelligence and is engineered for change. This solution has to:

- Provide comprehensive visibility across different layers of business, application, and infrastructure
- Adapt to the continuous change in business as well as technology
- Possess foresight and bring about

IT systems need a real-time closed-loop solution that can augment human intelligence and is engineered for change

a paradigm shift from reactive to proactive operations

- Offer end-to-end automation to react quickly and avoid problems due to delayed implementations

Intelligence as a Game Changer

Incorporating intelligence in IT plants can be a game changer. There are three types of intelligence: recognition, reasoning, and operative (Figure 1).

Recognition intelligence: This intelligence refers to the ability to construct a comprehensive context of the system connecting various layers from business to application to infrastructure. It has the ability to describe the "as-is" state of the system. The construction of the context involves two key tasks: model construction and the profiling of normal behavior.

Model construction: It requires the ability to model entities, relationships, and attributes by tapping into various structured and unstructured data sources such as technology manuals, event logs, inventory repositories, monitoring logs, among others. Various challenges such as incomplete data sources, inconsistency in the data sources, and continuously changing environments need to be addressed while forming the context.

Normal behavior profiling: To manage any system, it is important to understand the normal behavior of each of its components. An understanding of this normal behavior creates other capabilities such as anomaly detection, root cause analysis, prediction of future behavior, among many others.

Incorporating intelligence in IT plants can be a game changer

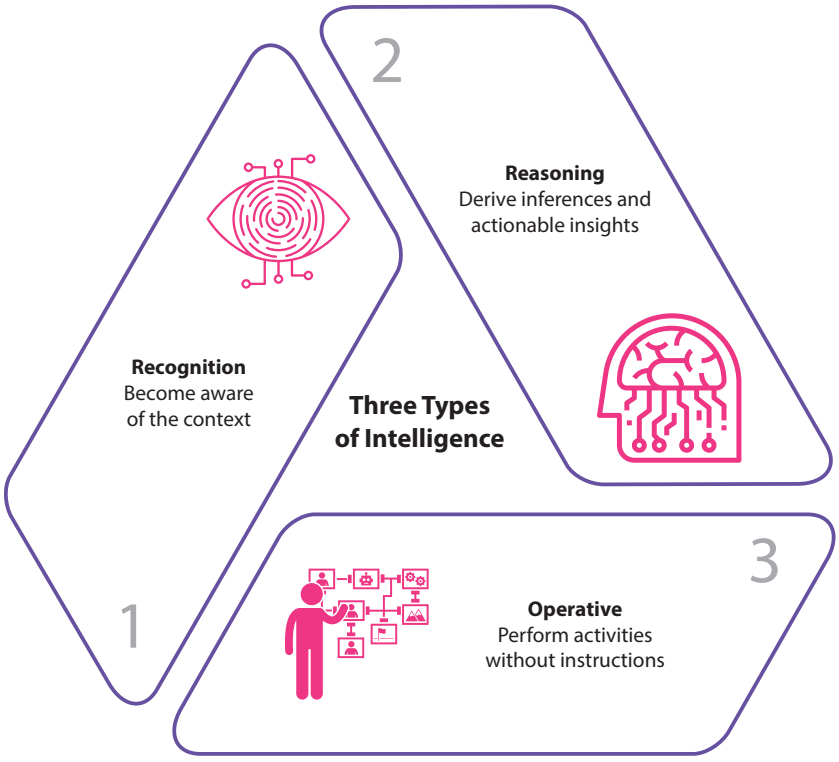


Figure 1: *Three types of intelligence*

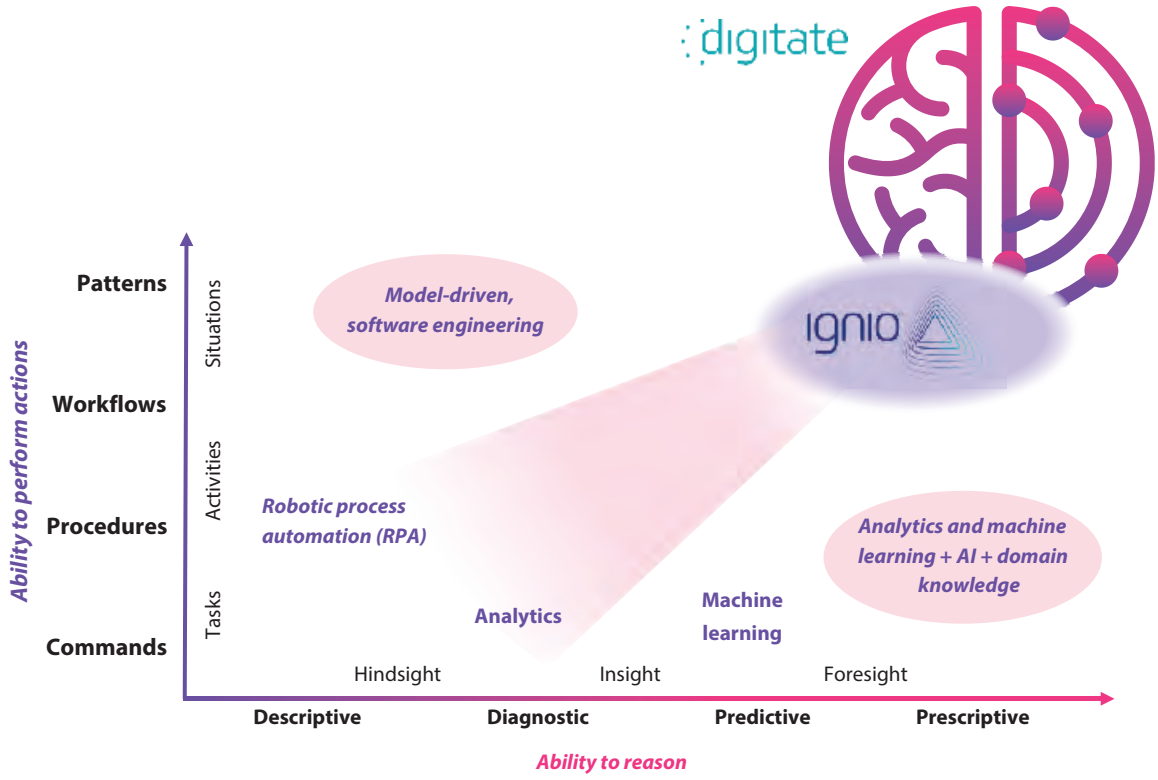


Figure 2: *Inside Cognitive Automation*

Some examples of recognition intelligence include the extraction of entities and relationships from technology manuals, or the derivation of trends and temporal patterns from historical behavior (Figure 2).

Reasoning intelligence: The focus of reasoning intelligence is to derive inferences and reason about the system. Most analytics tools provide only statistical observations and essentially answer the question: "What has happened?" Reasoning intelligence translates these observations into meaningful insights and actionable recommendations. It has complex capabilities and is able to:

- Diagnose the likely cause behind a certain behavior and answer the question: "Why did it happen?"
- Predict future behavior to provide foresight and answer the question: "What will happen?"
- Explore the possible solution space to provide actionable recommendations and answer the question: "What should be done about any described, diagnosed, or predicted insight?"

Operative intelligence: While the recognition and reasoning intelligence provide all the required capabilities to infer the what, when, and why about the system, there is still a need to take timely actions. Operative intelligence provides the ability to act without giving explicit instructions. It uses concepts of model-driven systems engineering to separate the process logic of orchestration from the execution of atomic tasks. Furthermore, it infers the orchestration logic using context and patterns built through recognition and reasoning intelligence. Operative intelligence thus makes a cognitive solution complete by closing the loop. This

The focus of reasoning intelligence is to derive inferences and reason about the system

is done by taking actions on the derived recommendations. Some examples of operative intelligence include intelligent alert suppression and aggregation, and auto-triage and self-healing of an incident.

Most products in the market are attempting to achieve higher levels of maturity along only one or two types of intelligence. For sustainable benefits, a cognitive system must be designed to achieve a high level of maturity simultaneously along all three types of intelligence. The development of such systems requires the combined power of artificial intelligence (AI) and automation. With this combined power, we can go far beyond simply reducing human effort or eliminating the need for intervention altogether. We can go further and unlock a whole realm of new, previously unimagined possibilities! This concept has the power to transform virtually all technological devices into intelligent, thinking machines.

The Technological Foundation of the Solution

There is considerable overlap in various terms used in the AI/ML space, and quite often, different people use different terms interchangeably. However, some clear distinctions can be made in terms of functionality. Intelligent systems require the ability to:

Discover hidden patterns: One of the fundamental requirements for any data-driven reasoning is to discover unknown properties in the data to explain some phenomena. These hidden patterns help build intuition about what is really happening in the data and form a foundation for sophisticated reasoning patterns. This functionality is mostly powered by statistics and data mining. Classical statistics techniques such as probability theory, optimization techniques, and regression models focus on using mathematical tools to quantify data and derive relevant data properties. Data-mining techniques such as clustering, classification, and association rule mining use statistics and other programming methods to discover unknown properties in the data.

Reason and derive inferences: This functional element uses the inferred properties of the system to reason about the world. This gives rise to intelligent behavior such as the ability to play a game of chess or drive a car. Various reasoning techniques such as rule-based, model-based, or case-based reasoning can be used to assess the impact of different actions and then choose the best action to achieve a goal.

Learn and adapt based on experience: This function ensures that the derived inferences adapt to various planned and unplanned changes in the system. Machine learning (ML) techniques use known properties to build models of what is happening behind some data and use it for a well-defined objective (usually prediction). Deep learning achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones.

Building Trustworthy Solutions

Recent years have witnessed the rapid evolution of technologies to design intelligent and autonomous

> Operative intelligence makes a cognitive solution complete by closing the loop

systems. This advancement can be attributed to many factors such as increased digitization, improved data availability, improved hardware, as well as a plethora of analytics libraries. They have allowed AI solutions to get out of research labs and into real-world deployments. We see advances in many critical areas such as medicine, banking, finance, telecom, and supply chain management. But, as these applications penetrate deeper into the society, they bring attention to trust-related issues. The real-world deployments of these systems have been hampered by a fundamental challenge: How can one make intelligent and autonomous systems "trustworthy"?

Various guidelines are being developed to ensure trustworthiness on multiple dimensions such as fairness, transparency, privacy, explainability, and auditability. For instance, the European Commission has introduced ethics guidelines

Driving Business Agility with ignio™

We present the ignio™ journey of one of the fastest growing financial services institutions in the UK. The company aspired to digitally reimagine its business.

Their objective was to drive rapid growth while ensuring high-quality customer experience. However, introducing any business or technology change was a slow and fragile process. It required days of planning among experts from different verticals to make any change in their batch systems.

Some of the key challenges were:

(a) Scale and complexity: The batch estate consisted of 52 business units, 50,000+ unique batch jobs, and multiple and legacy technology stacks.

(b) Pace of change: The organization frequently offers new business offerings and experiences rapid customer growth. Stringent regulatory compliance, and increasing service quality expectations, further added to the challenges.

(c) Constraints: Any plan for growth and change relied heavily on people and tacit knowledge to assess the impact of change.

ignio™ transformed their batch systems by bringing about a paradigm shift from "reactive to proactive" and "manual to automated." It first created a unified, enterprise-wide blueprint of batch jobs which covered 52 business units and spanned diverse batch schedulers and technologies such as Control-M, OpCon schedulers, SAP, Unisys, and legacy jobs. It then self-learnt normal behavior by profiling and modeling trends, patterns, and relationships. Finally, it helped the financial services firm for future business by predicting emergent violations of business commitments and providing prescriptive recommendations for handling situations.

The result: ignio™ delivered predictions with more than 90% accuracy and reduced the time required to plan for change from days to minutes. Today, ignio™ is operational in their production environment and is used in many interesting ways, ranging from anomaly detection, noise suppression, to predicting future behavior, and performing change-impact analysis.

for trustworthy AI that focus on ensuring that the AI system is lawful, ethical, and robust. General Data Protection Regulation (GDPR) also includes the right to explanation.

While there are different initiatives to ensure trustworthiness, these criteria are hard to formalize and quantify. We believe that to earn trust, intelligent and autonomous systems must exhibit two key characteristics: explainability and trust.

First, these systems must be capable of explaining or justifying every decision they take in every situation. There is no unique definition of explainability, and no common agreement on the approaches to ensure explainability. However, there is a common agreement that explainability refers to the ability to explain or present the decision's rationale in terms that are human-understandable.

Different approaches provide different levels of transparency to enable such explanations. On one hand, there are transparent methods, such as decision trees, classification and regression trees, and Bayesian classifiers, that can be completely traced and explained. While, on the other hand, there are black box models, such as artificial neural networks (ANNs) and support vector machines (SVMs), whose complexity makes their logic difficult to explain. One way to bring explainability in such approaches is to use a transparent method to get an approximate explanation of the output of a black box approach. For instance, researchers have derived decision trees or rules from SVMs. Similarly, Bayesian reasoning has been used to explain the decisions of image classification performed by deep neural networks (DNNs).

Creative text and visualization techniques are also becoming popular for producing explanations. For instance, visualization has been effectively used to explain feature importance in a classifier or to explain the intermediate layers of neural networks. Similarly, justifications for decision trees are given textually, or textual evidences are provided for image classifications performed by DNNs.

Second, these systems must be safe. AI solutions provide unparalleled flexibility in dealing with continuously evolving systems. However, the AI solutions themselves can be a target for attack. Creative solutions are required to ensure that an AI is not "mistrained" or manipulated. It should be possible to prevent the system from doing something bad or undesired. It needs solutions such as an "action firewall" that can mitigate both inadvertent and intentional safety risks.

All data scientists have been in a situation where they think an ML algorithm will do a good job of predicting something, but in production, it doesn't perform as expected. The consequences could vary from annoying to dire. Hence, it is important to establish a process to ensure that the ML algorithms are qualified and validated. Commonly used techniques include k-fold cross-validation, hold-out, leave-one-out cross-validation, random subsampling, and bootstrapping.

Another crucial task is to ensure that the data sources used for training the ML algorithms are secured in integrity. Techniques such as regularization and randomization are often used to ensure this. Game theoretic techniques are also effective

AI solutions provide unparalleled flexibility in dealing with continuously evolving systems

to follow a learn–adapt–relearn cycle to ensure the integrity of training data sets and the derived models.

Looking Ahead

AI has been a consistent buzzword in the global IT industry for several years. But the industry trend is now changing to "become digital" rather than "doing digital." The increasing focus on monitoring and instrumentation has further opened more possibilities for AI. AI systems will help the modern tech-savvy marketer become digital at breakneck speed. However, as these systems permeate every aspect of our lives, they need to address various trust issues in order to gain acceptance among end users.

Maitreya Natu

Dr. Maitreya Natu is a scientist at Digitate and is the product owner of ignio™ AI. Workload Management. He has a Ph.D. in Computer and Information Sciences and works in the space of systems research. Dr. Natu specializes in building cognitive solutions for managing complex systems.

Vaishali Sadaphal

Dr. Vaishali Sadaphal is a scientist with TCS Research and Innovation. She did her Ph.D. from the Indian Institute of Technology, Delhi. She has over 20 years of experience and specializes in the fields of sensor networks and data sciences.

Rahul Kelkar

Harrick Vin

High-Performance Computing

Building performance assurance tools and exploring heterogeneous architectures

Rekha Singhal

IN BRIEF

Performance Assurance of an application during its lifecycle faces new challenges. One of these is the exponential growth of data volume in the production cycle.

The conventional modeling approach is limited and too complex. Especially when deployments have to be agile. We, at TCS Computing Systems Research, are working on methodologies and tools to address this problem so that different big data processing frameworks can be modeled for large data sizes to accurately predict the performance of big data analytic workloads even during development.

Motivation

The world is hungrier than ever for more computing power. The world of science has had its taste of new discoveries in high-performance computing: unravelling the genome, deciphering astronomical data, running large-scale simulation of materials, or even insights into human behavior. Businesses crunch vast volumes of data and run complex cloud architectures to remain competitive. The entertainment industry produces more immersive experiences with gaming. The new possibilities with AI, better infrastructure promised by 5G, and the consumer demand with the growing edge (devices) are asking more of computing. There is so much happening under the hood that the twenties promise to explode with innovation and disruption.

The Emergence of Heterogeneous Architectures

Moore's Law is slowing down. Dennard scaling is hitting the power wall. The general-purpose CPU, which has itself evolved from a processor to a SoC, seems pushed to the limit. Therefore, the industry is now moving toward domain-specific heterogeneous architectures e.g., FPGAs (field-programmable gate array), CGRAs (coarse-grained reconfigurable architecture), and GPUs (graphics processing unit) that trade the flexibility of CPUs (central processing units) for improved efficiency. These

Heterogeneous architectures work on low clock frequency and data-flow architecture without any control unit, which leads to a higher performance per watt.

Fact File

TCS Research: Computing Systems Software

Outcomes: High-performance Frameworks and Systems for ML/DL, Performance Prediction, Analysis, Tuning models, and Tools for Big Data Analytics Workloads and Systems

Principal Investigator: Rekha Singhal

Academic Partners: Professors at Stanford University, IISc., Bangalore

Techniques Used: Mathematical Models, Simulation, Benchmarks, Optimization of Frameworks and Algorithms

Industries Benefited: All Industries

Patents: 13

Papers: 50+

The need for agile deployments and growing data sizes in the production environment has led to the challenge of assuring an application's performance during its life cycle.

specialized hardware work on a low clock frequency and data-flow architecture without any control unit, which leads to a higher performance per watt. For example, Microsoft Brainwave and Google's Tensor Processing Unit (TPU) are designed to accelerate deep-learning workloads.

Challenges Introduced by Heterogeneous Accelerators

We envision the next-generation scalable architecture platform for accelerating complex analytic applications, on heterogeneous architectures, accessing a variety of data models (such as relational, unstructured, spatial, temporal, etc.) processed by their corresponding specialized engines. Now, the middleware will have to identify and offload components of an application that are amenable to acceleration to specialized hardware.

Building such middleware is challenging and introduces new research problems motivated by the use of heterogeneous accelerators such as CPU for scalar, GPU for vector, TPU for matrix, and FPGA for spatial operations. This heterogeneity motivates the use of both parallelism and pipeline executions to improve performance. This leads to challenges for a compiler to translate a given workload across heterogeneous data engines and hardware accelerators for its efficient execution.

To be more specific, questions arise on how to:

- Capture user intent by facilitating the development of clearly specified (or clarity-optimized) heterogeneous programs for both expert and laymen users. That is, exploring different modes of conveying intent such as natural language, speech, and the development environment itself
- Convert heterogeneous programs into extensible unified representations
- Identify patterns in workloads to accelerate processing

- Optimize execution plans across heterogeneous computing units (e.g., CPU-based data-processing engines and hardware accelerators), a multiobjective optimization
- Generate optimized code specific to a computing target (e.g., FPGA and GPU)
- Execute the workloads efficiently across heterogeneous architecture by utilizing both software and hardware parallelism and pipelining

The Performance Assurance Problem

While middleware has a set of problems to address, high-performance computers have to deal with the issue of performance assurance of growing data sizes.

The need for agile deployments and growing data sizes in the production environment has led to the challenge of assuring an application's performance during its lifecycle. The standard approach to performance assurance is by conducting iterative testing of an application for different workload and data sizes, which can elongate the testing cycle time. This will not be acceptable when business agility demands speedy time to market. This motivates the building of predictive performance models for big data analytics.

The conventional performance modeling approach is limited to Queuing Networks, Markov Chains, Petri nets, and Simulation, which do not take in to account the intricacies of query processing engines, data management layers, and heterogeneous architectures. Moreover, these techniques are far too complex for developers to adapt to their daily project work.

Performance Assurance Tools and Methods

We, at TCS Computing Systems Research, are working on methodologies and tools to address this problem so that different big data processing frameworks can be modeled for large data sizes to accurately predict the performance of big data analytic workloads during development without the need for designers and developers to understand such techniques. These tools further help developers to tune workloads and the underlying system configuration parameters for optimal performance in the production environment.

Our key innovations are in building models and tools to estimate SQL query performance on both conventional RDBMS and newer data management systems such as Spark, for projected larger data sizes and workloads. We use the gray box approach to build these models. These models and tools also extend to next-generation architectures. They are augmented with optimization techniques to recommend tuning parameters for the system.

For instance, the performance prediction tool was used to estimate the performance of a reporting application executing on an Oracle database for large data sets (in terabytes) of a large international insurance company. The tool built a prediction model using statistics collected on application execution on small data sets in the development environment. The model was used to optimally design the database, tune the data engine parameters, and provision the hardware for optimal performance of the reporting

Our key innovations are in building models and tools to estimate SQL query performance on both conventional RDBMS and newer data management systems

Experiments with CGRA

Coarse-grained reconfigurable architecture (CGRA) has gained traction in recent years as a high-throughput, low-latency, and energy-efficient accelerator. CGRA is a spatial accelerator, which is configured statically and yields energy efficiency by avoiding dynamism introduced due to control unit and cache hierarchy in general purpose systems.

In contrast to field-programmable gate arrays (FPGAs), CGRAs are reconfigurable at the word level, a higher level as opposed to the bit level. This rigidness in CGRA reduces routing overheads and improves clock frequency, compute density, and energy efficiency compared to FPGAs.

There is a race amongst different types of heterogeneous architectures to mature to mass commercial deployment. Working hands-on on CGRA for a common problem that we observe as we service customers gives TCS an immensely valuable experience to meet the emerging future.

Modern data analytics applications, such as data exploration, often require joining of more than two relations. During my stint at Stanford, I had the opportunity to design a new algorithm for "Multiway Join" to join multiple relations in parallel using Plasticine, a spatially reconfigurable CGRA developed at Stanford for building deep-learning pipelines.

Joining of multiple relations in parallel, that is, multiway joins, can execute faster than the traditional way of joining them in a sequence of binary joins, when the intermediate output of a binary join is larger than the

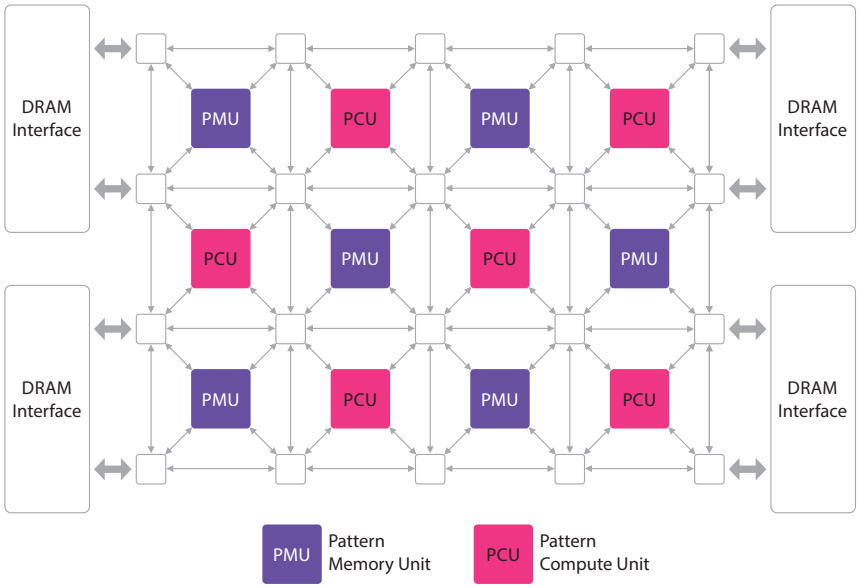


Figure 1 : *Plasticine Architecture – A CGRA Research Prototype Developed at Stanford for Ml/DL Workloads*

required replication of input data in the multiway join and the final output is relatively small.

This is the case in many examples such as finding cliques of people in Facebook or in the video streaming business scenario. (Example: finding the customer profiles of those who watched the movie *Parasite* for the second time last Thursday in Michigan) This may need joining of customer, product, viewing time, or other tables. Multiway joins could be accelerated on specialized hardware such as Plasticine, facilitating inexpensive data replication. Plasticine-like architectures have amenable features for executing multiway joins efficiently with: (1) peak high compute power of 12.3 FLOPS throughput and (2) high-bandwidth static network supporting cheap broadcast and hence inexpensive data replication.

We studied algorithms to efficiently perform multiway joins on a Plasticine-like accelerator.

We proved a performance gain for executing a traditional sequence of binary joins on a Plasticine-like accelerator over general purpose systems. Further, we showed additional improvement in performance using a three-way (a multiway join) over traditional joins leading to 4500X speedup over a general purpose-based Postgres database system.

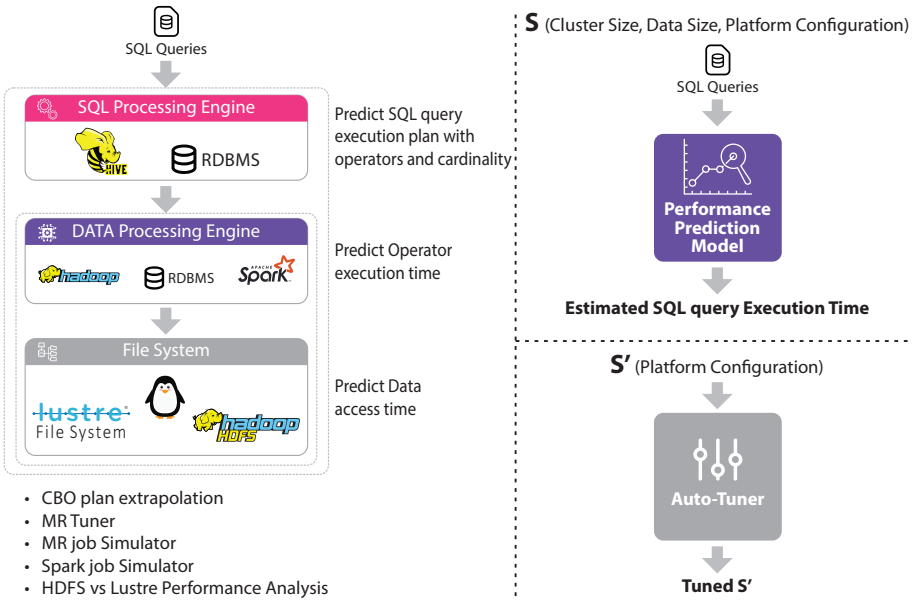


Figure 2 : *Predicting SQL Query Execution Time for Big Data Systems*

application. Similar engagements were carried out for Spark and Postgres data processing engines.

A Look at the Future Future

We envision modern business applications to be data driven and deployed on heterogeneous architectures. These applications will experience the desired high performance while being transparent to change in the business environment.

In the future, we expect a library of performance-optimized algorithms and performance prediction models to be used during the complete lifecycle of application development, right from the request for proposal (RFP) stage to deployment. Our frameworks shall recommend possible options for next-generation architectures for an application deployment meeting the business objectives. Further, these frameworks can be used to build optimal deployment with capacity sizing and auto-tuning and/or auto-scaling of the production system while meeting the performance guarantee.

We live in interesting times, particularly from the perspective of computing architecture. There are several experiments happening in the space of Von Neumann computing and non-Von Neumann machines. The race to market maturity is on. The best way for IT services providers to prepare for the next disruption is to keep watching and experimenting with partners in this space.

Rekha Singhal

Rekha Singhal is a Senior Scientist and Head Computing System (Software) Research Area. She focuses on accelerating development and deployment of enterprise applications in data-driven programming environments. Her research interests are heterogeneous architectures for accelerating ML pipelines, high-performance data analytics systems, big data performance analysis, query optimization, storage area networks, and distributed systems. Rekha has 13 granted patents in international territories. She also has several publications/presentations in international and national conferences, workshops, and journals. Rekha received her M.Tech. and Ph.D. in Computer Science from IIT Delhi, and has been a visiting researcher at Stanford University, United States.

Bridge-Playing Software

A vehicle to further our understanding of AI

Arun Bahulkar,
Venkatesh R, et al.

IN BRIEF

Mind games like Chess, Go, and Bridge have fascinated AI researchers for several decades now. While Chess and Go have been conquered, Bridge continues to remain elusive, since the handling of incomplete information and randomness of card distribution pose a different kind of challenge. We, at TCS Research, believe that we can harness the power of advanced machine learning techniques, along with simulation and logic to go beyond human expertise in this regard. Our early experience with this indicates that this is still a hard problem but is possibly within reach.

It is well known that computers have outplayed humans at complex mind games like Chess and Go—representing major strides in the artificial intelligence (AI) and machine learning (ML) space. However, the game of Bridge continues to present challenges. The best of Bridge programs today stand no chance against human experts. The primary reason for this is the nature of the game, where every deal is different (a new set of hands are dealt out each time), the information available to a player is partial and uncertain and a very small space is available to communicate a large amount of information.

TCS has taken up the challenge to build a Bridge program that can take on human experts. Our initial focus is on the bidding aspect of the game since that is where the major expertise gap is.

An annual competition[1] for Bridge-playing programs is organized under the aegis of the World Bridge Federation providing a platform to benchmark progress in the field. The Federation has mandated that the programs need to run entirely on standard laptops and the game has to be played within an allotted time (about twice as fast as humans). This has consequentially made research in the area even more interesting.

TCS has taken up the challenge to build a Bridge playing program that can take on human experts

[1] Details of the World Computer Bridge Championships are available at https://bridgebotchampionship.com/

Fact File

TCS Research: TCS Bridge Bot

Outcomes: 6th place in the 2019 World Computer Bridge Championships

Principal Investigators: R Venkatesh, Rajendra Gokhale, Arun Bahulkar

Academic Partners: Prof Shirish Shevade, IISc Bangalore

Techniques Used: Formal Methods, Machine Learning

The present set of programs leverages handcrafted rule bases for bidding, and simulation for card play. Some of them also use simulations during the bidding phase to choose a final contract. These techniques have been fine-tuned over a couple of decades now but new breakthroughs are needed to take the programs to the next level.

It comes as no surprise that people are now attempting to apply ML techniques to resolve bid ambiguity. Researchers are attempting to replace the handcrafted rules with those derived using machine learning. There are two major challenges here. Firstly, the programs need to bid using a published bidding system that opponents and partners understand in the same way. Secondly, the bidding has to be optimized globally rather than locally, considering all possible situations that could ever occur. Any change in a bidding rule at an earlier stage of bidding can impact the later stages, thus, exploding the space in which learning has to be done.

Using rule bases for bidding has inherent limitations such as the inability to capture "tacit" knowledge as "explicit" knowledge. As any experienced Bridge player will testify, bidding knowledge is not fully expressible using rules. This has led to further experiments where the bidding system is encoded as a rule base and neural networks are used for finer "judgment" (to ensure that we meet the "explainability" requirement). No good results are available yet on this front and TCS Research is pursuing this line of approach. Results have been reported by others to learn bidding completely from scratch, though in a limited context. These results look

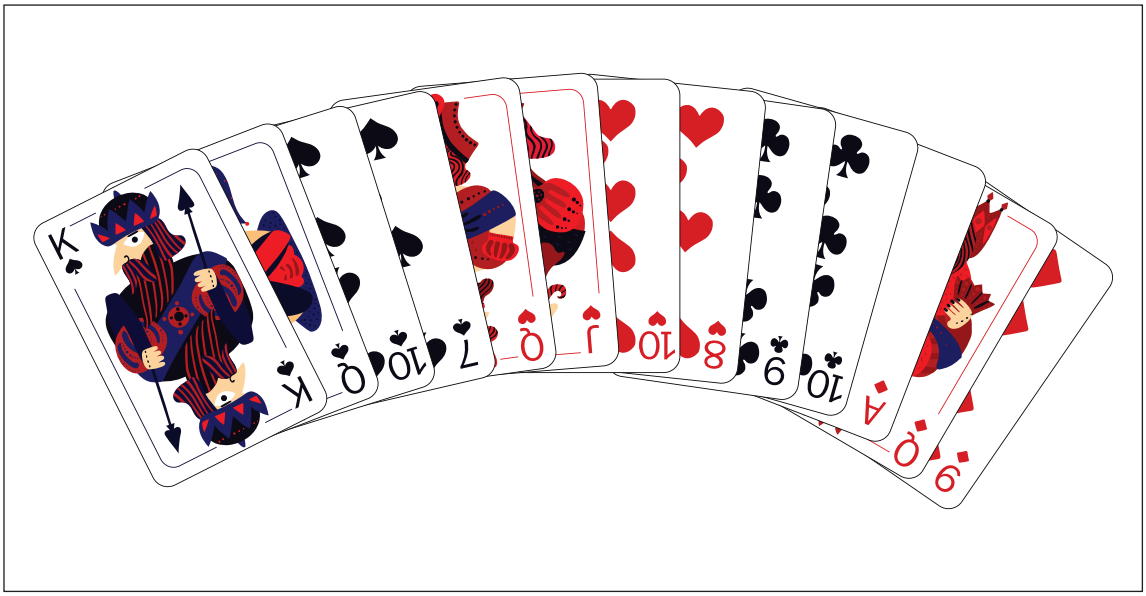


Figure 1: *An example of an Opening Hand (ACOL Contract Bridge)*

encouraging but do not meet the "explainability" requirement.

Despite the promise that neural networks hold, it is still believed that simulation has a crucial role to play, especially during the final stages of bidding. So, part of the research challenge is also to see how to combine rule bases, neural networks, simulation, and various methods of machine learning to optimize bidding.

There is scope to do a lot of research in the play engine as well. There are no good techniques to draw inferences from a card played, or to induce opponents to make errors by hiding information—and this is something experts are really good at.

We look forward to the day when media will announce the victory of the machine over the human in Bridge. And that will certainly imply a lot of progress in the development and application of AI. From then on, we expect that machines will be able to help people become better players.

For a quick introduction to the game of Bridge, you can visit: https://www.acbl.org/learn_page/play-bridge/

It comes as no surprise that people are now attempting to apply ML techniques to resolve bid ambiguity

Arun Bahulkar

Arun Bahulkar is a Research Advisor at TCS. Earlier, he headed the TCS Software Engineering Research group and has been involved in building several tools that have been extensively used as TCS products. He is currently eyeing the next World Computer Bridge Championship and using it as a vehicle to explore the world of knowledge-intensive systems. Arun received his postgraduate degree in Computer Science from IIT Madras in 1980 and has been with TCS since then.

Venkatesh R

Venkatesh is a Chief Scientist at TCS Research and heads its Foundations of Computing Research Area. He has been with TCS for more than 30 years, working primarily for sectors like software engineering, formal methods, and verification. During this tenure, he has led several tool development projects, including TCS ECA, a static analysis tool that is sold commercially by the company. Other tools include MasterCraft and, more recently, a formal specification notation called EDT. Venkatesh holds an M.Sc. in CS from Pune University.

Rajendra Gokhale

A Computer Science graduate of IIT Bombay, Rajendra Gokhale has worked in organizations like AT&T Bell Labs, Microsoft, and Cloudera, creating leading edge technologies. He is an eminent Bridge player in India and has many national-level wins to his credit.

Preparing for Quantum Computing

Insights from TCS' Academic Research Partner TIFR

R Jaikumar,
R Vijayraghavan

TCS has established a partnership with the Tata Institute of Fundamental Research (TIFR), a premier research institute in India, to collaborate and co-create new solutions on the quantum computing platform. TIFR has successfully built the first superconducting circuit-based quantum computer in India with three qubits. TIFR is on course to build a seven-qubit quantum computer and TCS has partnered with TIFR to enable cloud access to this quantum computer. This interview reinforces some quantum computing facts, enriched by TIFR team's experimentation.

What gives quantum computing some unique capabilities?

Dr. R Vijayaraghavan: A classical computer uses a bit as the fundamental unit for storage and processing. It can take one of the two possible states at any time, that is, 0 or 1. On the other hand, a quantum computer uses a quantum bit or qubit as the building block. The laws of quantum mechanics allow the system to use a combination of both states, that is, 0 and 1. When extended to many qubits, the number of possible states that can exist simultaneously becomes enormous and gives a quantum computer some unique capabilities.

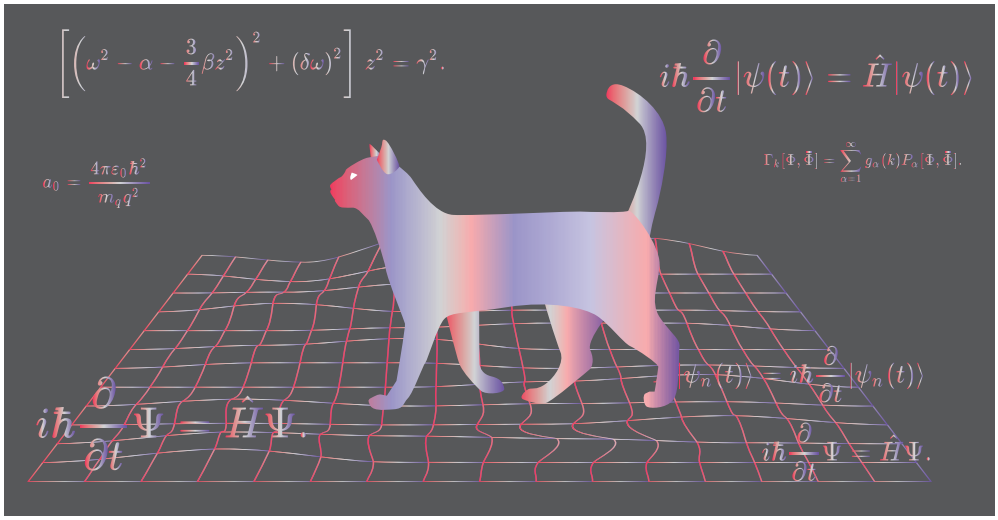
Dr. Jaikumar: In quantum computation, the state of the computer at any point is modeled as a superposition of basis states $|0\rangle$ and $|1\rangle$. The superposition assigns to each basis state a complex amplitude α and β, respectively, which behaves roughly as the square root of the probability. The subsequent evolution of the state is modeled as unitary operations (for example, rotations, reflections, etc.) on the state, which preserve the sum of the squares of the amplitudes ($|\alpha|^2 + |\beta|^2 = 1$) associated with the basis states. Unitary operations allow for cancellation of amplitudes, which gives quantum computation its additional power.

For example, in classical computation, the state of a bit is a probability distribution over the set {0, 1}, so it is determined by a probability vector of the form (p_0, p_1) where $p_0 + p_1 = 1$. A typical evolution might replace the bit by a random bit if it is 1, and leave it as it is if it is 0. This would modify the state (p_0, p_1) to $(p_0 + p_1/2, p_1/2)$. This operation corresponds to the application of the stochastic matrix:

$$\begin{pmatrix} 1 & 0.5 \\ 0 & 0.5 \end{pmatrix}$$

In quantum computation, the state of a qubit is described by a unit vector of the form $\alpha|0\rangle + \beta|1\rangle$, with complex amplitudes α and β such

Unitary operations allow for cancellation of amplitudes, which gives quantum computation its additional power



© Shutterstock

Figure 1 : *Artist's visualisation of Erwin Schrödinger's Cat Explaining a State Known as Quantum Superposition*

Understanding of linear algebra and some basic probability is usually sufficient to work in the area of quantum algorithms

that $|\alpha|^2 + |\beta|^2 = 1$. For example, a unitary operation can change this to $((\alpha+\beta)/\sqrt{2}, (\alpha-\beta)/\sqrt{2})$, which corresponds to the unitary matrix:

$$1/\sqrt{2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$$

This matrix corresponds to the single-qubit Hadamard gate.

Can quantum computing (QC) perform search functions more efficiently?

Dr. J: If we are given access to a table with "n" entries and are required to locate where a particular item "x" is, a quantum algorithm, namely, Grover's algorithm, would (with high probability) locate the element in about sqrt{n} steps. Whereas, it can be shown that every classical algorithm would need around "n" queries for this problem.

One must, however, be cautious in interpreting this claim. For Grover's quantum algorithm to work, we must have the ability to probe the table with qubits in superposition. We cannot just build the quantum algorithm over the current implementations of relational database management systems (RDBMS). The quantum algorithm promises a reduction in the number of queries made to the database. Between two queries, there is some quantum processing that is done in the algorithm. In an ideal quantum computer, this will account for algorithmic overhead.

If the RDBMS allows quantum access, then in principle, one can locate an element in a table of four elements using just one query, assuming each entry appears only once. The exact number of queries needed for table with "n" elements is complicated, but it grows as (pi/4)sqrt{n}. So for a table of 100 entries, one would need about eight quantum queries.

How can teams build and operate their own quantum computers? What are the difficulties?

Dr. V: There are several competing quantum hardware platforms, leveraging trapped ions, superconducting circuits, or others. Each has its own

challenges. However, the most important challenge is to build a sufficient number of coherent qubits with high fidelity; that is, qubits and quantum gates with low error rates. Then, implement error correction, and build a computer with enough qubits to have access to a large enough computational space (such as an "n"-qubit Hilbert space). The problem is that quantum systems are fragile and very difficult to control, manipulate, and measure.

Is there a simple way to understand some basic concepts of quantum computing, such as superposition and entanglement, without resorting to mathematics?

Dr. J: It is possible to compare superposition and entanglement with classical notions of random variables and dependence, respectively. Such qualitative understanding, however, has its limitations. Some use of basic linear algebra is unavoidable to properly appreciate the significance of superposition and entanglement.

What difference will a software engineer experience when writing programs for a quantum computer?

Dr. J: Quantum computation has not developed sufficiently where one can write a program in a high-level language. It is difficult to predict the precise form quantum computers will take.

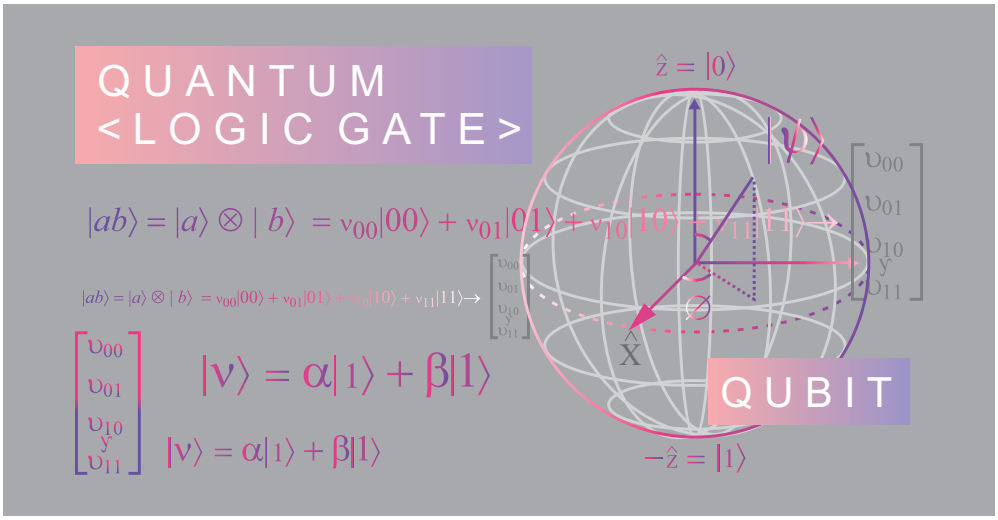
For a general programmer, access to quantum computers might well be provided in the form of libraries. When a call is made to a sub-routine responsible for some specialized task, perhaps the task will be delegated to the quantum hardware to accomplish, and the results will be conveyed to the classical program. If quantum computers do develop, most software engineers will use quantum computing through such an interface, without having to deal with quantum nitty-gritties.

Is the knowledge of quantum mechanics essential to design algorithms for the quantum computer?

Dr. V: Yes, a basic knowledge of quantum mechanics is needed but the necessary rules are easy and straightforward to grasp. The tricky part is to figure out how to exploit these new rules to solve a particular problem.

Dr. J: The basic algorithms in quantum computing, as they are approached by theoretical computer scientists, are based on a small number of principles. Understanding of linear algebra and some basic probability is usually sufficient to work in the area of quantum algorithms. Knowing quantum mechanics helps in many ways. Many of the computations involved in quantum algorithms can be seen as discrete manifestations of physical phenomena, for example, the two-slit experiment and the Deutsch–Jozsa algorithm. Furthermore, understanding advanced ideas, such as quantum random walks, shed light on what might be possible using quantum algorithms. Conversely, narrowing the study to the discrete setting often allows one to understand quantum phenomena in a simple way, that otherwise appear rather inaccessible to an average computer scientist.

Quantum computation has not developed sufficiently where one can write a program in a high-level language



Can we write a program in C or Python and expect it to run on a quantum computer, assuming there is a reliable one with 10,000 qubits?

Dr. J: A quantum computer when deployed in a regular computer is likely to serve only as an accelerator for some specialized tasks (just as a GPU). 10,000 qubits might well be able to speed up some tasks by a large factor, but one must account for the additional time taken in interfacing with quantum hardware, etc.

Dr. V: The exact programming language doesn't matter in the sense that the high-level language could be anything. However, the software environment would need to have some compiler which would take the code and convert it into instructions that can be executed on a quantum processor. If C or Python developers don't include the relevant functions in their software platforms, then one cannot write a program for a quantum computer. If the question is whether a regular classical program written in C/Python would run on a quantum computer, the answer is no; only quantum algorithms should run on a quantum computer.

Are there any limits on the length of a quantum program? How long can a quantum program run today? Do I have to run a program many times or will a one-time run suffice?

Dr. V: In principle, there is no limit to the length of a quantum program. However, today's quantum computers do not use error correction and hence are prone to errors. The longer the program, the higher the chances of an error. Qubits with superconducting processors have coherence times of about 100 microseconds. Any algorithm which takes significantly longer than that time will be prone to errors. Running a program once or many times will depend on the nature of the algorithm.

What would one not be likely to use a quantum computer for?

Dr. V: Well not for most of the things! Not every problem can be solved faster using a quantum computer. For example, I can't imagine needing a quantum computer for sending e-mails or word processing.

In what ways does quantum computing influence cryptography?

Dr. J: Cryptography, especially the kind based on public keys, often relies on assumptions that certain mapping that can be computed efficiently cannot be inverted. It might appear that since quantum algorithms are capable of inverting more efficiently than classical algorithms, such assumptions are violated if the adversary is in possession of a quantum computer. This is true only to a small extent and is misleading. General-purpose inversion algorithms (for example, Grover's algorithm) provide only minor improvements over classical algorithms, and do not pose a threat to cryptography.

The real threat emanates from the remarkable tool of Quantum Fourier Transform (QFT). Many cryptographic algorithms based on number theoretic functions often have an underlying periodic structure. By learning this periodic structure, one can mount and attack such cryptographic protocols. QFT can unravel this periodic structure rather efficiently. So, if a quantum algorithm for factoring numbers is implemented, we will no longer be able to use the Rivest–Shamir–Adleman (RSA) algorithm.

Anticipating such threats, several new proposals exist for basing cryptography on functions and tasks that are expected to be hard even for quantum algorithms.

Why is it that quantum chemistry is touted to be the application most likely to be speeded up by early quantum computers?

Dr. V: One of the earliest applications envisaged for quantum computers was to solve problems in quantum mechanics which are very inefficient to solve on regular computers. Quantum chemistry involves solving the Schrödinger equation to understand the properties of molecules. Scientists have figured out various ways to map problems in quantum chemistry to quantum computers. We are still far away from solving quantum chemistry problems of practical relevance, but the progress is promising.

Many cryptographic algorithms based on number theoretic functions often have an underlying periodic structure. By learning this periodic structure, one can mount and attack such cryptographic protocols. QFT can unravel this periodic structure rather efficiently.

Quantum
– Here and Now

While hardware development has been impeded by the challenges of qubit coherence and errors, there is much that can be done today:

Algorithms: Several new quantum algorithms have been discovered. These algorithms have been validated, thanks to the availability of quantum simulators (on classical computers) which can simulate systems as large as 49 qubits.

Today, there is a need to study and evaluate the current quantum algorithms and conceive an architecture on which the available QC platforms will fit in to solve existing or near future applications. It is then necessary to estimate the number of qubits and the error rates required to solve the problems of industry scale with quantum supremacy. This could help project the likely year when quantum computing could be a business value proposition.

Quantum on the cloud: Given the costs of owning and maintaining a quantum computer, it is very likely that most enterprises will prefer to use a quantum computer on the cloud. Enterprises could just start using cloud services to create value for themselves and their customers.

But how can we validate the results returned by the quantum computer on the cloud? These computations will be difficult to compute with classical servers. There have been some recent breakthroughs in this regard. One solution involves computing something similar to a hash which can be computed only by a quantum computer. The correctness of the hash can be verified easily by a classical computing platform.

While a reliable and practical quantum hardware will take many years to come, many alternative solutions are coming up. Examples: mem-compute [1], hardware-based spiking neural networks [2], coherent Ising machines [3], and digital annealers [4], especially when one looks at applications like optimization and its closely related fields like machine learning. More such technologies are expected to evolve soon.

As we write this, we know that not all quantum computing applications will be able to run efficiently from day one. Nevertheless, there is a tremendous opportunity for IP creation and, thereafter, exploitation in the search of newer quantum algorithms and use cases. However, this will require heavy investment and therefore needs careful evaluation.

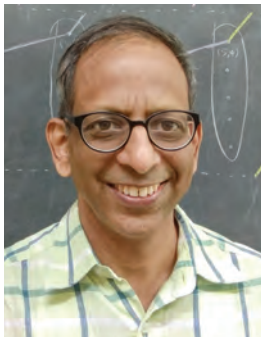
Manoj Nambiar

[1] https://www.memcpu.com/

[2] https://en.wikipedia.org/wiki/Spiking_neural_network

[3] https://www.nature.com/articles/s41534-017-0048-9

[4] https://spectrum.ieee.org/tech-talk/computing/hardware/fujitsus-cmos-digital-annealer-produces-quantum-computer-speeds



Jaikumar Radhakrishnan

Dr. Jaikumar Radhakrishnan is a Senior Professor with the School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai. He obtained his B.Tech. (Computer Science and Engineering) from IIT Kharagpur in 1985, and his Ph.D. (Computer Science) from Rutgers University in 1991.

His research interests include randomness and computation, combinatorial and algebraic methods in complexity theory, information theory, and quantum computing.

Rajamani Vijayraghavan

Dr. Rajamani Vijayaraghavan is an Associate Professor in the Dept. of Condensed Matter Physics and Materials Science at TIFR, Mumbai, since December 2012. He received his Ph.D. in Applied Physics from Yale University and carried out postdoctoral research at the University of California, Berkeley. His research group at TIFR called The Quantum Measurement and Control Laboratory (QuMaC) investigates quantum phenomena in superconducting circuits. Some key highlights of the group's work include the development of a broadband ultra-low noise amplifier for quantum measurements and a novel multi-qubit processor design for applications in quantum computing.

Glossary

Coherence—The merit of a particular quantum computer is determined not only by its qubit counts but also by the "coherence" of these qubits, which means how long these qubits stay in their superposition states to process information, before being decohered by external (quantum) noise. The decoherence time is usually very small, in orders of microseconds (quantum dots) to milliseconds (in electron spin qubits).

Entanglement—Entanglement is a quantum mechanical phenomenon in which the quantum states of two or more objects have to be described with reference to each other, even though the individual objects may be spatially separated.

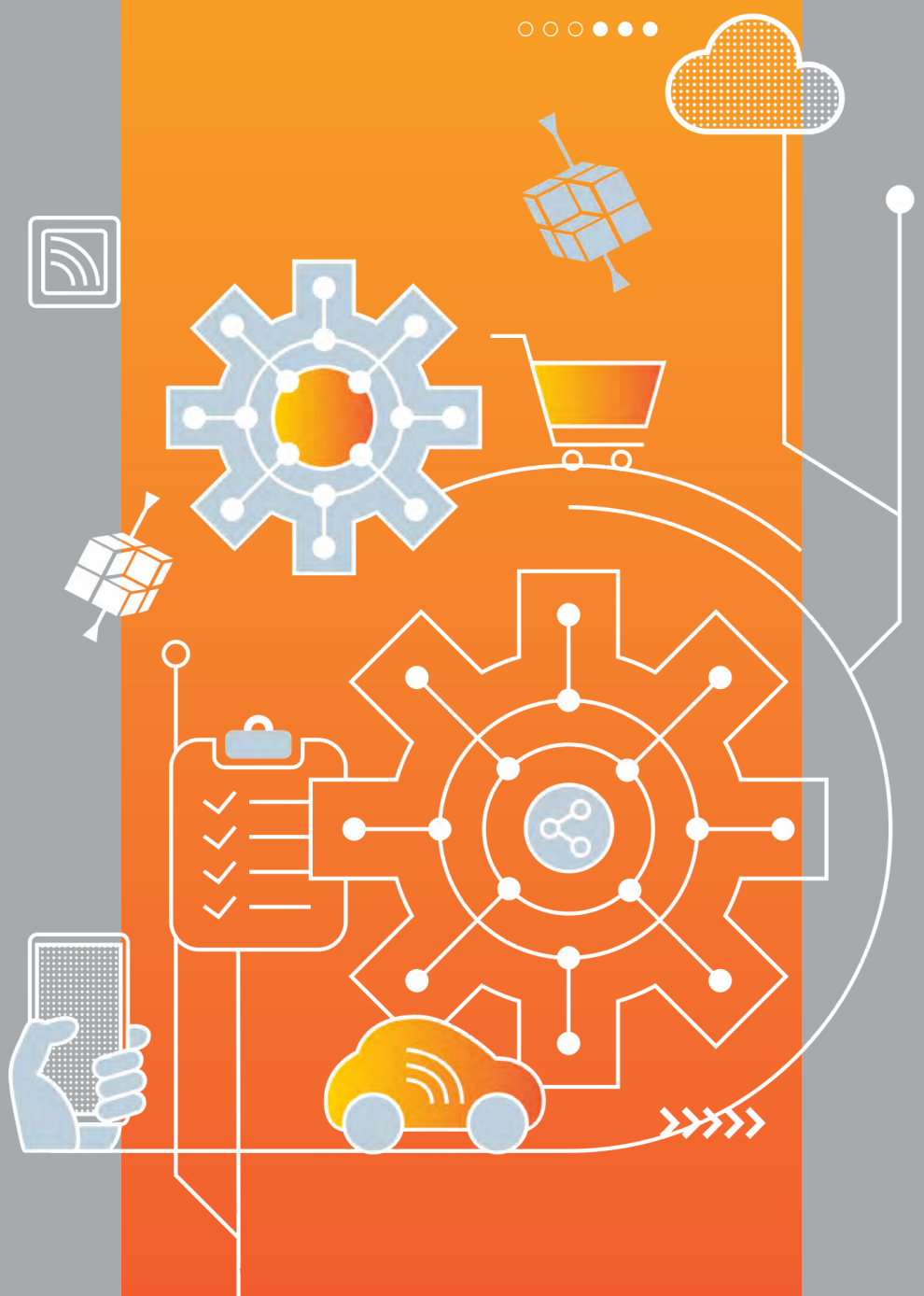
Hilbert space—A Hilbert space is an abstract vector space possessing the structure of an inner product that allows length and angle to be measured.

QFT—Quantum Fourier Transform (QFT) is the mathematical and conceptual framework for contemporary elementary particle physics.

RSA algorithm—RSA (Rivest–Shamir–Adleman) is an algorithm used by modern computers to encrypt and decrypt messages.

Superposition—Superposition is the ability of a quantum system to be in multiple states at the same time until it is measured.

Digitization of Industry



InnoVista

TCS InnoVista and Tata InnoVista are platforms for recognizing and celebrating innovations within TCS and at the Tata group level. TCS encourages and motivates teams to leverage these platforms and participate in InnoVista, promoting a culture of collaborative innovation, learning and sharing.

TCS teams have continuously participated in Tata InnoVista and won awards every year.

TCS InnoVista 2020 had more than 6400 entries. 16 Teams from TCS InnoVista reached the finals of Tata InnoVista 2020.

Digitization of Industry

Sankha Som

Across industries, the move is from atoms to bits. In his 1995 book titled *Being Digital*, the then-MIT Media Lab director Nicholas Negroponte posited that the key indicator that an industry was undergoing digitalization was when the core products and services that the industry delivers move from the realm of physics (atoms) to the realm of information (bits). So, bits standing in for atoms was the indicator that an industry was undergoing irreversible digitalization. The phenomenon did play out for many industries, such as music, print media, banking, and more; many others such as manufacturing, logistics, and pharmaceuticals, to name a few, still depend on making or moving atoms. And that seems to hold true for the foreseeable future at least, until we reach a "Beam me up, Scottie" era. As long as human existence stands rooted in atoms, there will be industries where bits won't stand in for atoms. However, today bits profoundly influence the movement of atoms. In fact, this phenomenon has accelerated during and in the recovery phases of the global pandemic where entire industries and societies have started to increasingly rely on the promise of digital technologies.

At TCS Research, we are harnessing the power of data, computation, and connectivity to not only usher in the next wave of digitalization in information-led industries, but also to lead the first wave of digitalization in industries still reliant on the physical world. Our research in digital twins for the enterprise is helping highly digitized industries, such as Telecom, to build purposive models of their enterprises and use these models for data-driven decision making. For industries such as transportation and logistics, which are in the midst of a digitalization wave, we are using sophisticated algorithms and hyper-connectivity to optimize commutes in urban environments that can improve people flow across multimodal journeys. Our research in supply chains is helping optimize material flow by automating decision making and ensuring that supply chains are more predictable. And for manufacturers that are experiencing the first wave of digitalization, we use a combination of data-driven modelling and simulations to create platforms that help design functional materials with desired properties for a wide range of applications, including energy storage, drug delivery, and nano-materials.

To bring the benefits of advances in science and technology to our customers, it is not enough to merely invest in path-breaking research. The scalability of these technologies in our customers' context is also extremely important. Most novel technologies go through incremental cycles comprising early experimentation, pivots, and pilots before they truly deliver value for an enterprise in a scalable manner. Today, TCS Research has validated delivery mechanisms to partner with our customers through these lifecycle stages. We allow customers to partner with us and our ecosystem in early experimentation through our Co-Innovation Network (COIN™). Our globally spread Pace Ports are physical facilities where our customers can collaboratively shape and pivot our intellectual assets to their specific needs. Our agile, scalable innovation centers become the foundries where validated technology is shaped and readied for scalable deployment within enterprises.

At TCS Research, our ambition is to 'Innovate at Scale' with our customers and thereby become the key entity that powers their growth and transformation journeys.

Sankha Som

Sankha is a Chief Innovation Evangelist, TCS Research and Innovation and leads a team that evangelizes TCS Research and Innovation to clients globally. He also oversees the Insights, Foresights, and Marketing functions for Corporate Research and Innovation.

He has over 18 years of experience in R&D and Innovation in the Information Technology sector. He has fulfilled several roles in Research, Product Development, Project and Program Management, Business Development and Sales, Innovation Management in a customer organization, increasing the efficiency and effectiveness of R&D organizations.

Sankha is an MBA in Analytical Finance and Strategic Marketing from the Indian School of Business (ISB), Hyderabad, India. He also holds an undergraduate degree in Computing Sciences.

Seamless Customer Journeys

An enterprise digital twin for CSPs

Senthilvelan Natarajan,
Kaustav Bhattacharya, et al.

IN BRIEF

While incurring huge infrastructure investments, communications service providers (CSPs) have to regularly offer new services and keep operations efficient. End-to-end enterprise visibility, operations automation, and dynamic data-driven decision-making become critical to achieve these. The business user needs to understand business-IT bottlenecks, what-if scenarios for products, services, and processes that may have to change. A purposive model of the enterprise environment which can be used for various customer journey simulation exercises will be invaluable to a CxO.

In this chapter, we outline how an Enterprise Digital Twin can help CSPs gain cross-functional visibility and insights to understand, design, and optimize customer engagements across the lifecycle.

A key success factor for CSP operations in this dynamic business environment is to ensure successful, first-time-right customer journeys

Communication service providers (CSPs) must be smiling: There are more mobile connections than people on this planet; the mobile has become a channel for hundreds of services. Technology is galloping in the area of communication.

Business must be booming? But CSPs are actually gasping. Technologies such as the Internet of Things (IoT) and 5G are introducing new complexities that are driving customers to expect a range of new services. Just offering operations support services at a personalized level for millions of customers is extremely complex. Advancements in technology are creating new business models but are also allowing plenty of free riders (such as over-the-top service providers) who seem to be getting more of the revenue pie, while using CSP networks. Investments in infrastructure are growing heavier and profit margins from traditional business lines are growing thinner for CSPs.

Limited Visibility

A key success factor for CSP operations in this dynamic business environment is to ensure successful, first-time-right customer journeys. A CSP tracks its customer's journey through five key stages: lead, order, activation, billing, and service assurance. At every stage, intervention from business groups, IT, and network operations—both human and automated resource groups—are required to orchestrate a seamless journey for the customer. However, in the current telecom landscape, there is limited visibility

Fact File

Patents: 3

Papers: 1

and correlation between enterprise divisions to do this. Organizational operating blocks are highly fragmented, working in silos without end-to-end visibility and awareness of their contributions towards customer journey outcomes. A customer journey deviation caused by performance bottlenecks of one operational block, if not identified and resolved, inevitably impacts the next block, thereby compounding the result of unresolved bottlenecks.

The Human Factor

While customer journeys are key to ensuring operational excellence, today's customers driven by retail expectations expect contextual and personalized engagements across their lifecycle. CSPs deal with petabytes of customer related data—purchase transactions, feedback, complaints, call detail records, tweets, and blog post records which can be used to construct a faithful representation of a customer's lifestyle.

As more and more customers make transactions in shopping, banking, healthcare, and travel through digital channels, CSPs have a great opportunity to understand each customer's lifestyle better at the n=1 level and therefore, customize its products and services to offer a "purposive" experience. That most CSPs are currently unable to do this is apparent from the deviation experienced from the ideal journey map charted: 25–40% of all customer journeys deviate from the "ideal" experience map, leading to lost business opportunities and low net promoter scores (NPS). On their part, customers struggle to access the support systems the CSPs have set up, such as self-service modules and constantly have to reach out to support executives to complete their tasks. This results in sub optimal outcomes for both CSPs and customers.

Models—Past and Present—to Parse Human Behavior

Various mathematical and statistical machine learning (ML) techniques were historically developed for "static problems" such as image processing and text recognition, since the 1970s. These techniques handled fairly stable signals/data in the past. Human behavior, however, is a different story. A traditional ML model trained to detect fallouts from customer connections and journeys may deteriorate in accuracy over time. It now requires constant monitoring and maintenance by a data scientist with continual infusion of new semantic knowledge. As customer behavior, markets, and competitors are subject to frequent (and constantly evolving) trends, constant retraining of the model, as well as frequent development of the new features to detect these trends (emergent behavior, complex behavioral properties that are not part of raw data) are required. These can only be done through the combined work of a semantics domain expert working alongside a data expert.

An Enterprise Digital Twin can help CSPs looking to gain cross-functional visibility and insights across their organization to understand, design, and optimize customer engagements across the lifecycle.

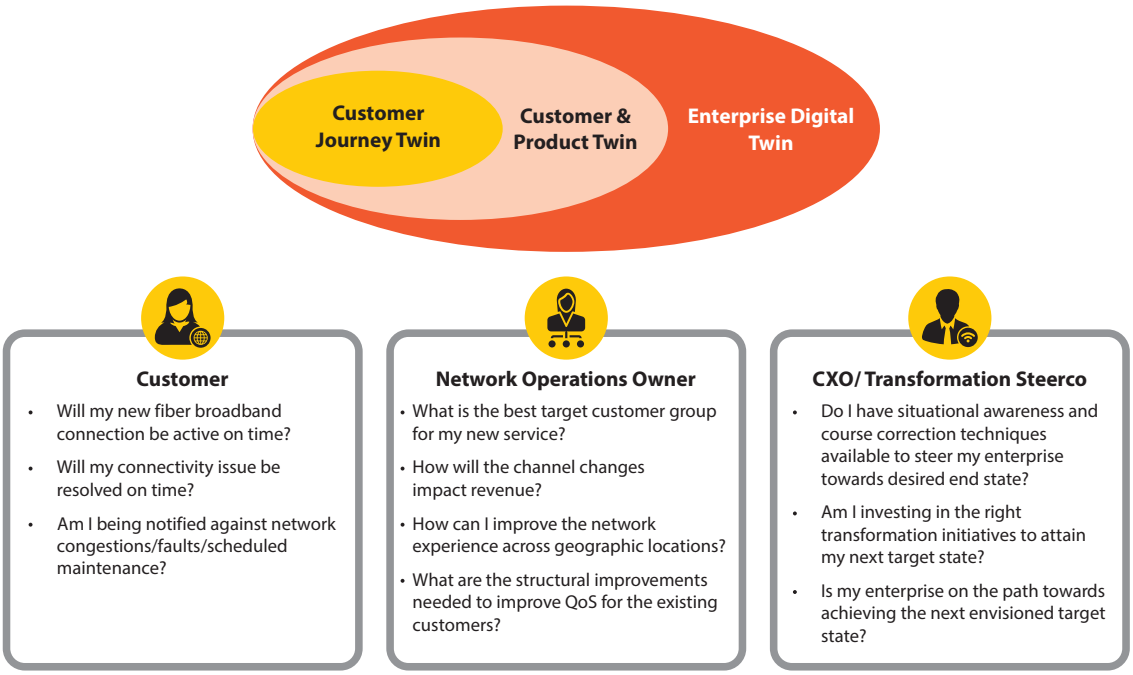


Figure 1: *Key Customer and Enterprise Issues for CSPs*

Business and technical events are aggregated and ingested into the journey correlation engine for real-time correlation of customer journeys and creating a cross-functional single view across organizational blocks

In recent years, data scientists have started to employ "heavyweight" statistical methods and ML algorithms to try and cope with this complexity. These powerful tools, including new deep learning techniques, collect data and analyze its attributes in order to be able to classify behavioral patterns, detect anomalies, and predict future trends. However, even such tools (historically developed for "static problems" such as image processing and text recognition) cannot easily cope with human behavior data and complexities.

Actor Simulation Approach

The agent simulation approach works in a completely different way. Instead of deriving patterns from input data, it is based on the discovery of entities (in other words, actors) and their behavioral relationships "social behavioral laws" mathematical relationships that emerge in an enterprise when people, process, and systems operate in the same space. These laws govern the way various statistical properties of crowd behavior evolve over time. This happens regardless of the type of data, the demographics of the users who created it, or the data size. An engine integrates these laws into its data analytics component, which efficiently extracts the underlying social attributes of all people contained in the raw data being provided as input (e.g. phone calls, taxi rides, financial investments). It supports visualization of an enterprise along three aspects—people, process and technology—with relationships across aspects as well as levels.

Building the Enterprise Digital Twin

We believe that organizational systems with all their chaos and silos can be faithfully represented in a digital format, which we call

the Enterprise Digital Twin. In this virtual environment, key entities such as customer, product, process, network, and resource are mapped and key strategic decisions (e.g. rolling out new products/channels of engagement) are tested before being piloted in the market. Through exhaustive enterprise simulation in this high-fidelity virtual representation of the enterprise, deviation risks can be identified and eliminated without disturbing any dependent entity in the real enterprise. Additionally, the digital twin of the enterprise also helps identify the optimized KPIs to align them with organizational goals and objectives.

An Enterprise Digital Twin can help CSPs looking to gain cross-functional visibility and insights across their organization to understand, design, and optimize customer engagements across the lifecycle. It envisions a simulated enterprise environment comprising three distinct blocks:

- Infrastructure (IT, networks, physical assets)
- IT application stack
- Business processes

Business and technical events are aggregated and ingested into the journey correlation engine for real-time correlation of customer journeys and creating a cross-functional single view across organizational blocks. Correlated customer journeys are further simulated and monitored using an artificial intelligence-based (AI-based) hybrid decision engine. Deviations are predicted, tracked, reported, and the next best actions are triggered toward corrective resolutions to minimize impact to the customer experience. Additionally, enterprise simulation enables simulating macro trends such as competitor actions, influence of emerging technologies, and micro factors such as inventory availability and resource performance, among others, and understanding whether these have an impact on current enterprise operations. Further, what-if and if-what analyses are supported powered by a self-learning controller which enables continuous optimization of the enterprise to address macro as well as micro trends, impacting customer journey outcomes.

A Hybrid Decision Engine

The AI and enterprise simulation-based hybrid decision engine is critical to the Enterprise Digital Twin solution, as it is responsible for the following tasks:

- Journey clustering
- Journey simulation
- Journey deviation prediction
- Dynamic deviation analysis and problem auto-discovery
- Enterprise simulation
- Next best action recommendation

The simulated environment, or the digital twin of the enterprise, is connected to the real enterprise through amorphous interfaces across the infrastructure, IT applications, and business operation blocks, enabling the recreation of an enterprise context and related customer journeys in real-time as well as producing a simulation of ideal customer journeys.

From early implementations in industrial engineering, the digital twin technology is gaining momentum in other industries including telecommunications. The success of the digital twin depends on its core algorithm, data models, and the trainability of these models periodically, based on the insights gained.

TCS TwinX™ Mapping Journeys and Preventing Fraud

TCS' business insights platform, named TCS TwinX™, brings together AI and digital twin technologies. It simulates digital twins of:

- The customer: This is in accordance with their personas and archetypes to predict their behavior and guide them to behave in a desired manner.
- The customer journey: This is mapped at a process step level by gathering context across all layers that attribute to the occurrence of a process step in real-time.
- The enterprise: This is done to preempt and allow for simulations of strategic decisions and predict their impact to the enterprise, detect deviations, and translate impact respective to each enterprise function.

The platform ingests entity data from various data sources and is feature engineered using inputs from telecom process SMEs and algorithms such as time series, NLP and RNN, persona generation, and pattern association rule mining. Correlations are used to attribute the data to milestones and are compared with statistically derived ideal states for deviation detection and labeling via anomaly detection using Isolation Forest, Local Outlier Factor, and DNN Auto-encoders.

Aggregating the entities provides macro insights using clustering algorithms such as K-means, DB Scan, and Agglomerative. The architecture remains dynamic and adaptive using self-optimization capabilities enabled by traditional algorithms such as Decision Tree and Random Forest and advanced algorithms such as xgboost, LightGBM and catboost, and deep neural nets (DNN).

Coupling context from information gathered with machine learning capabilities enables the TCS platform to deliver actionable insights at both micro (individual entity) and macro (aggregated entities) levels. The insights could be informational or action recommendations.

The platform predicts deviations at future milestones and provides preventive or course correction recommendations. In addition to entity level insights, it can use accumulated knowledge from each journey to correlate with other journeys and

form behavioral entity groupings, denoted by clusters and decision trees. Therefore, insights and action recommendations are shared for groups of entities with similar characteristics.

The TCS platform will also translate micro and macro insight derivations to predict KPI performance and offer intelligent decision aiding capabilities such as virtual simulations of business decisions. This helps to accurately predict outcomes considering the machine-learned state of current and future (predicted) affairs. The insights are presented on custom-built user interfaces and can also be shared across various notification channels. TCS TwinX™ also uses ingested data to predict the likelihood of occurrences like churn or fraud and shares prediction accuracy scores to external applications for informed decision-making. The TwinX platform for customer journey management was piloted for a South African telecom company in early 2019. The key business issues addressed were the entity fallout prediction in its customer acquisition/renewal/service assurance journeys along with fraud predictions in its SIM swap customer journey.

A SIM swap fraud (also known as SIM splitting and sim jacking) is a type of financial fraud that generally happens in two steps. A fraudster targets and cracks the two-factor authentication and two-step verification to change the SIM for the mobile number of the victim. Subsequently, mobile telephone usage fraud for international calls, data or bank transaction OTP is executed.

Benefits achieved were as follows:

- 12% reduction in average activation cycle time with 10% volume of data ingested. Activation times is expected to reduce by upto 30% as data increases
- A 4X increase in first-time-right (error free) orders
- Prediction accuracies between 80% and 90%
- Typically, 70% of SIM swap requests are fraudulent and the TCS TwinX™ platform has helped avert 62% of such requests
- Over 63,000 loyal customers have been protected from potential fraud with an average monthly customer fund protection of approximately $6M

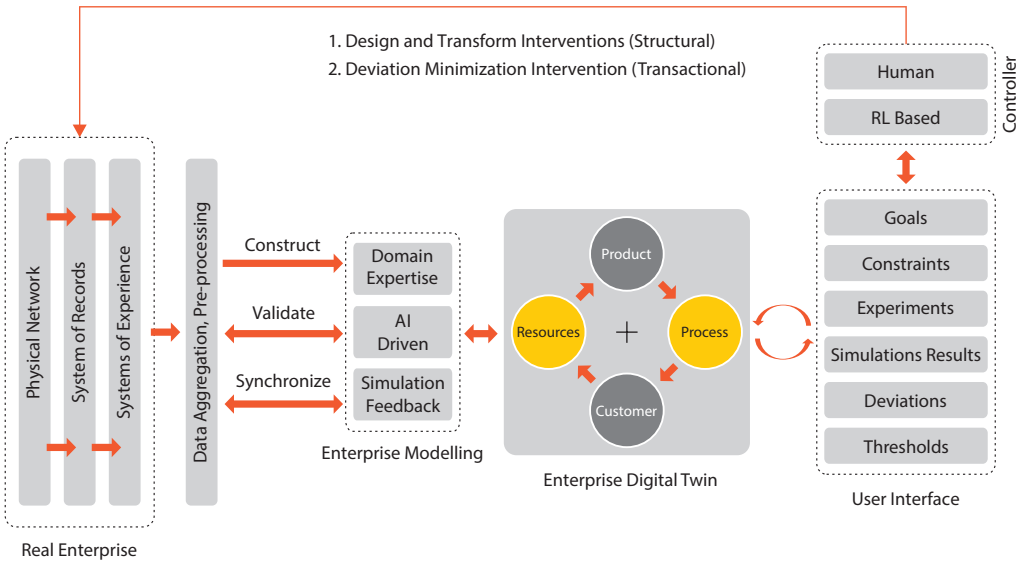


Figure 2: *TCS TwinX™ Solution*

Senthilvelan Natarajan

Senthilvelan Natarajan is the Global Head AI Platform Solutions of Communications, Media, and Information Services Business Unit. He has more than 22 years of experience in Global IT Services distributed across the US, Europe, South Africa, and APAC markets, notably in consulting and implementing transformative solutions. He works with TCS Corporate Research teams and the Industry Research teams in solving business problems with innovative technology solutions, particularly the portfolio of AI and intelligent automation solutions. He has filed patents in the area of digital twin and AI. Velan holds a bachelor's degree in engineering from the Government College of Technology, Coimbatore, India.

Kaustav Bhattacharya

Kaustav Bhattacharya is the Global Product Manager for TCS Enterprise Digital Twin. He is actively involved in TCS Research and Innovation programs for the development of new products and concepts and also conducts digital maturity assessment for Telcos across the world to help them realize real-time enterprise goals. His interest areas include intelligent automation, blockchain, and digital twin, and he brings a decade of rich consulting experience in this field. Kaustav holds a bachelor's degree in engineering from Manipal University, India and a master's degree in business administration from Cardiff School, UK.

Harmeet Kaur

Industry analyst - Telecommunications

Connected Commute

Hyperpersonalised cost efficient travel

Arvind Ramanujam, Sowmya Arasu, et al.

IN BRIEF

Two seemingly different problems are discussed here: multimodal urban commute and airline routing. Today, urban commuters face multiple challenges in the form of congestion, unpredictable travel times, and unoptimized transportation networks. City planners are under tremendous pressure to improve cost efficiency while responding to events near real-time. Similarly, in air travel, service providers operate on very thin margins while the travelers are compelled to optimize their journeys through tedious manual planning. In response to such challenges, stakeholders are leveraging the computing power of smartphones, pervasive network of sensors, and machine learning algorithms to make commute easier.

This article describes our research on development of commodity sensor-based traffic prediction and multimodel navigation systems and their extensions to the airline route optimization problem.

Rapid urbanization and population growth are exerting enormous pressure on transportation resources, be it urban commute or long-distance travel. Traditional travel industry business models are being undercut by technology-based services such as cab aggregators, and online travel portals. To enable better traffic management and ease of commute planning, our research team has looked at four problem areas: instrumentation, analytics, real-time navigation, and journey planning.

Instrumentation

Extensive instrumentation in developing countries is a necessity for city administrators to efficiently manage congestion. Unfortunately, most developing cities aren't sufficiently instrumented.

Developed countries, on the other hand, now face the complications of too much instrumentation. Cities such as London and New York have expert teams like Transport for London (TfL) and Department of Transport (DOT), respectively, to plan, monitor, and optimize their transportation resources. The

Traditional travel industry business models are being undercut by technology-based services

Fact File

TCS Research: Data and Decision Sciences (DDS) / Intelligent Transport Systems (ITS)

Outcomes: TCS Multi-Modal Trip Planner Application, Pilot Instrumentation, Prediction Module, Smart Journey Planner, Live Navigation, and Routing Platform

Principal Investigator: Arvind Ramanujam

Academic Partners: Indian Institute of Technology, Madras

Techniques Used: Machine Learning, Open Source Routing Machine, Computer Vision, Geographic Information System (GIS), Sensor Fusion

Industries Benefited: Travel, Transportation

Patents: 5 filed

Papers: 4 in Tier-1 Journals and Conferences

The accuracy of predictive analysis is high, thanks to the large amounts of instrumentation data collected over time

accuracy of predictive analysis is high, thanks to the large amounts of instrumentation data collected over time. Now, many of these hardware devices have become redundant and thus expensive to maintain. These cities are migrating toward predictive analysis tools while reducing their instrumentation footprint.

Most of the instrumentation devices such as induction loop counters, cameras, and lasers are tailored for western traffic conditions. They may not work well in developing countries where lane separations are nonexistent and vehicle types are highly heterogeneous. Thus, there is a need for a cheap, scalable instrumentation architecture that adapts well to the traffic conditions in these countries.

Prediction and Analytics

Prediction and analytics services based on instrumentation require clean and reliable data for good results. This is often not possible in developing countries. The data is noisy and unreliable. Data quality issues also arise due to ad hoc planning and changes by transportation authorities. Unplanned events increase the fragility of the transport system and pose additional challenges.

Real-time Routing and Navigation

Today, navigation app users rely on static schedules to plan their journey. These apps often provide inaccurate start and arrival times. Real-time scheduling can help users obtain accurate and latest information on temporary transport inconsistencies and delays. Real-time navigation view across all modes of transport can help the user enjoy a seamless journey experience. Therefore, data from historic trends and current traffic conditions must be fused to provide correct navigation cues.

High-performance Long-haul Journey Planning

For an airline, multimodal long-haul journey planning with the help of search engines had its share of challenges:

- With the addition of every new mode and every connection/ break in a journey, search options explode by many orders of magnitude.
- The network throughput required to service millions of queries per second, when search options explode, was very high.

Therefore, it was decided that making incremental changes to the airline's existing platform was untenable.

Solution Approaches

For each of these problem areas, we initiated research with our academic partners. Our anchor customers allowed us to work on validated datasets and test solution approaches.

Instrumentation

To monitor the traffic, we worked on an approach that uses:

- Wi-Fi-enabled MAC scanners deployed at specific points of interest (POIs) between which we can track data (MAC address and time) anonymously from the various mobile devices that cross these points
- GPS devices on public transport to capture geospatial information along with time, which can be used in predicting the arrival and travel times
- Traffic cameras, to capture the live feed information of city traffic flow patterns
- Social media platforms like Twitter to extract sentiments regarding the city traffic

Commercial-grade sensors and cameras were chosen for their simplicity, ruggedness, and ability to work in difficult conditions including unreliable power supply, dust, and irregular maintenance schedules. Their limitations in data accuracy, consistency, and quality were managed in the data processing pipeline.

Prediction and Analytics

The sensor and camera data were processed using a combination of domain-based techniques and machine learning techniques. Domain-based techniques include a flow-based model (spatial and temporal). Data-driven techniques include:

- Artificial Neural Networks (ANN), K-NN (k-nearest neighbors), SVM (Support Vector Machine)
- Image processing with deep learning and custom training sets
- Time series analysis

Data noise was reduced with domain-based preprocessing

Real-time scheduling can help users obtain accurate and latest information on temporary transport inconsistencies and delays

The state and the validity of a route under navigation was constantly evaluated against the evolving network state

techniques. Custom image processing techniques were developed to deal with heterogenous vehicle types, lane splitting, and non-uniform lighting. The state of the traffic network and public transportation system were tracked in real-time along with future projections.

Real-time Routing and Navigation

With an accurate tracking of the present and future state of the network, a custom routing algorithm was engineered fusing schedule-based (RAPTOR algorithm) and topology-based (contraction hierarchy) routing techniques. This provided multimodal capabilities to the end user. The state and the validity of a route under navigation were constantly evaluated against the evolving network state. Based on a custom scoring mechanism, alternate routes were computed and provided to the user.

High-performance Journey Planning

While in principle, the routing problem is the same for a city-scale public transport system and a global airline network, the underlying engineering challenges are different. The major differences are the dependency on unreliable third-party systems to get availability of seats and the inherent volatility of fares. In addition, the throughput requirements and the nature of the airline network warranted a different engineering of the routing algorithm. In addition to this, an extensible query mechanism was

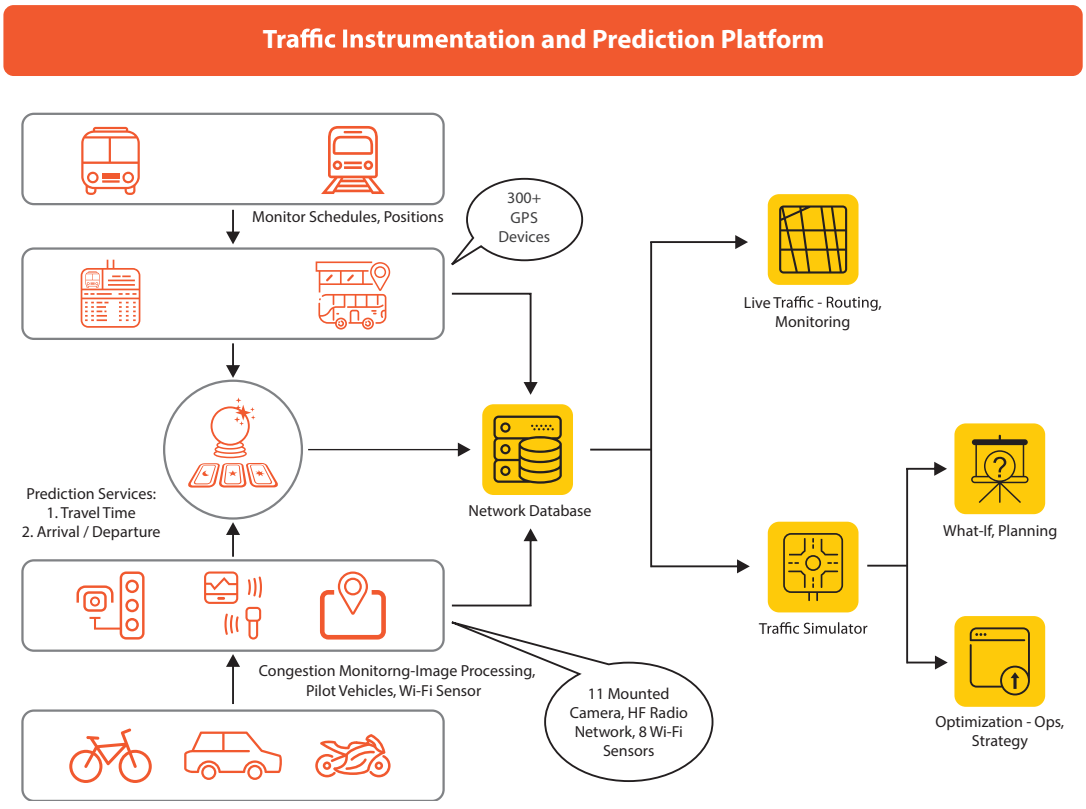


Figure 1: *Instrumentation, Prediction, Routing, and Analysis*

put in place to generate complex journey planning requirements across space, time, and user preferences.

Field Trials

Trip Planning Research Concept

About 300 GPS sensors were installed in Chennai's Metropolitan Public transportation buses, 8 Wi-Fi MAC scanners were installed on arterial roads. Initial experiments were done with data generated from 1,231 trips over a period of 45 days for GPS and 1.3 km distance of road stretch for about two weeks for the Wi-Fi sensors.

For the purposes of the field trial, we created an app that was tested for buses plying within the IIT-Madras campus. City-scale field testing was done with buses, the metro rail, overland rail, and pedestrians. Other navigation systems simply failed due to their inability to track the actual schedules.

(View also "Multimodal Travel Options, Meaningful Search Results.")

Future Work

Given the effectiveness of our pilot program at IIT-Madras, the next step will be to establish an optimal network of low-cost devices throughout the city. This will help us enhance traffic analysis and prediction.

Multimodal Travel Options, Meaningful Search Results

Travel meant moving from point A to point B with few variations in route or mode of transport. No longer so. Not just adventurers, even the mainstream traveler often wants to go down paths less taken. Today it is common for travelers to make two or three hops, stay in a hotel en route, the choice of these driven by various personal preferences. For such journeys, travelers are unable to get quick and optimal options from online portals of airlines, or travel aggregators. At the same time, travelers view the travel agent model of travel planning, where the agent orchestrates the journey, to be outdated.

Today, most travel information is received, and transactions are made, online. Web-based over-the-air (OTA) aggregators are projected to grow into a trillion-dollar business by 2022 at nearly 11% growth year over year (YoY). Every transportation major wants to get into the travel planning space and offer journeys as seamless as possible, thereby gaining market share for their core services.

Global Distribution Systems (GDS) and airline companies now work together with other service providers like cabs and hotels to plan multimodal journey

options along with virtual interlining. Some OTAs provide virtual interlining along with travel insurance to cover the risk of missed flights. However, travelers often tend to find better deals by planning travel themselves. Currently available multi-modal multi-hop travel plan aggregators are not suited to users' needs as they provide suboptimal results based on static schedules. Third-party information aggregators like Google provide journey planning options based on static schedules but cannot book tickets for the traveler. Search engines also throw up thousands of suboptimal search results which are difficult to sift through.

While in principle, the routing problem is the same for a city-scale public transport system and a global airline network, the underlying engineering challenges are different. The major differences are the dependency on unreliable third-party systems for information on the availability of seats and the inherent volatility of fares. In addition, the number of origin–destination pair combinations and the number of searches per second along with the nature of airline networks, which currently do not provide interlining of mainstream and low-cost carriers. This warranted a different sort of engineering of the routing algorithm.

A focused design thinking workshop was held with the customer (a GDS platform provider) team. Through the workshop, the problem statements were crystallized. The MVP was built in a time boxed schedule with agile interactions with the customer.

As part of the MVP, an extensive query mechanism was put in place to generate complex journey planning requirements across space, time, and user preferences for the customer.

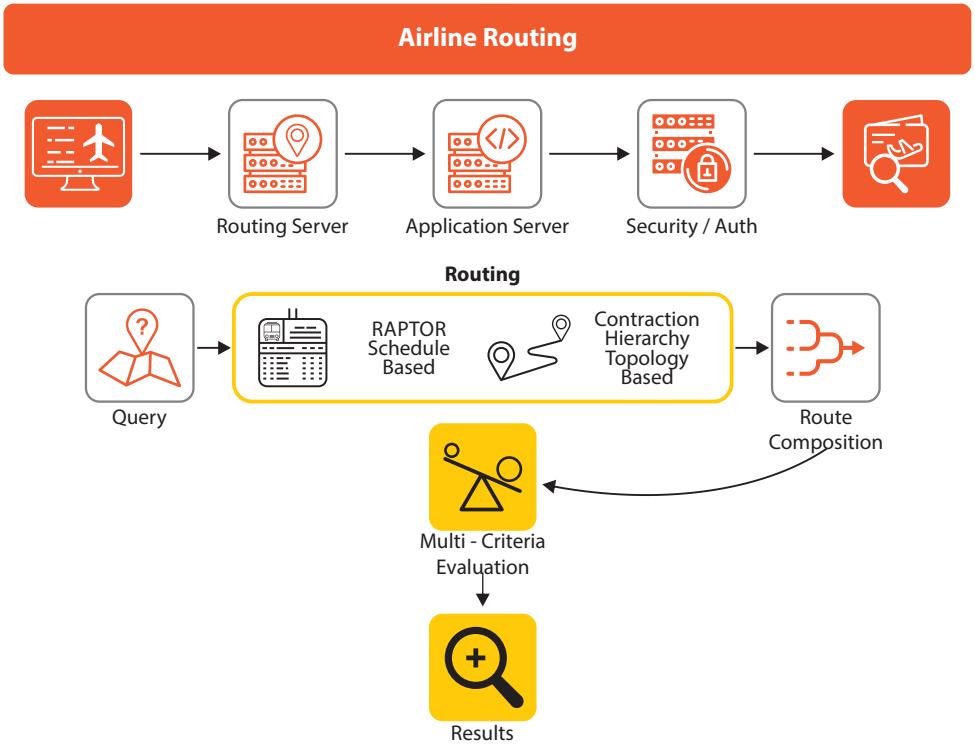


Figure 2: *Airline Routing System*

The task of optimizing the user's time, cost, and choice in getting from point A to B was undertaken. The MVP provided the following advantages to the traveler:

- Search: Relevant but limited options to help the traveler make a choice easily were provided.
- Modes: Three modes of travel air, bus, and rail were offered; all combinations of these could be tried.
- Choices: Cheapest, shortest, or earliest arrival options were highlighted; cities that could be excluded en route were all provided as searchable options.
- Hotels: Layover options could be included.
- The Golden Ticket: The best option at every minute of booking was displayed.

From an engineering perspective, the MVP was built for extensibility with:

- A normalized data model and objects with low ingestion barriers to the data
- An abstracted algorithm that can run over new sources of data without the need of modifying the internal system of the routing engine

The airline routing engine MVP was built on a variant of TCS' multi-modal trip planner routing engine. It provides features that the current GDS providers are yet to build into their core systems such as virtual interlining, multimodal transportation, journey planning with complex requirements and last-mile connectivity. The MVP worked well with the valid datasets provided by the airline customer alongside open data available for rail and bus. The engine was tuned for matching the throughput of the existing platform while consuming a fraction of its resources.

Arvind Ramanujam

Arvind Ramanujam is a Scientist at TCS Research and Innovation and is a member of the Data and Decision Sciences Research group. Arvind's areas of interest include large-scale traffic simulation, electric vehicle modeling and cross-domain simulation. Arvind's team has been working with IIT-Madras' Center for Excellence in Urban Transportation to monitor and predict traffic parameters using frugal instrumentation. He has designed and delivered network planning and management products for customers across the world.

Arvind holds a Bachelor's degree in electrical engineering from Pondicherry University. He is also a telecom network architect.

Sowmya Arasu

Sowmya is an Architect from the Travel, Transportation and Hospitality (TTH) unit. She was part of the innovation team that created a Journey Planner for the customer. She has over 12 years of experience and has worked with customers in Telecommunications and Travel Commerce domains. Her interests are in Data Science, Database design, and Query Optimization. She is a TCS Contextual Master and has participated in, & won, hackathons in TCS. Sowmya is a University Rank Holder from Anna University for Bachelors of Engineering in Computer Science.

Sudharshan Chakravarthy

Managing Uncertainty in Supply Chains

Exploring stochastic and robust optimization

Viswanath Ganesan
A P Srinivas

IN BRIEF

Managing uncertainties in supply chain is a CEO/Business concern. Today, the CEO or the Chief Supply Chain Officer is relying on IT to help, as it all hinges on good data flows, insights, strong algorithms, simulations, and efficient IT systems.

The TCS Research team believes that a lot can be done to manage uncertainties, automate several aspects of the supply chain, and drive not just efficiencies but actual monetary gain.

A proactive supply chain offers a clear competitive advantage for a Consumer Packaged Goods (CPG) company. CPG supply chains are under extreme pressure from macro-economic trends such as shifting trade alliances, fluctuating fuel prices, changes in legislation and tax regimes, strategic salvos from digital businesses and multi-channel operations, activism (including, demand for transparency) around provenance, and volatility in the retail business. Supply chains are also affected by natural disasters - such as the recent corona virus outbreak.

At an operational level, uncertainty in the flow of goods may occur due to delays at the manufacturing facility, transportation issues between echelons, damage during handling, pilferage, changing customer preferences, changes in the competition, and a range of other factors that need to be managed. These uncertainties affect various stakeholders – producers, logistics providers (trucking companies), retailers, and customers – in different ways.

Uncertainties and Disruptions in Networks

Most real-world network applications (including those in financial services) have data flows that operate under uncertainty. In fact, uncertainty is a crucial characteristic of network structures. Disruptions to networks can be classified as small-scale random disruptions, amplified random disruptions, and large-scale major disruptions.

Modeling Supply Chains and Logistics Operations

As supply chains operate with a very large number of parameters and face

> Designing supply chain and logistics models for decision support depends on the nature of data available, considering various factors such as random disturbances (or uncertainties) in the system, worst case performance needs, and disruptions

Fact File

TCS Research: Supply Chain Logistics Risk

Outcomes: Frameworks/Architecture, Decision Support Models, Proofs of Concept (PoC) with Business Units, Patents and Research Papers

Principal Investigator: Viswanath Kumar Ganesan, Ananda Padmanaban Srinivas

Techniques Used: Multi-Objective Optimization (Linear/Non-linear/ Mixed Integer Programming); Discrete Event Simulation; Network Flow Models; Machine Learning Models

Industries Benefited: Transportation, Retail, Consumer Product Goods, Manufacturing and Industrial Services, Internet of Things (IoT)

Patents: 2

Papers: 5

different types of disruptions, it is difficult to make optimal decisions using gut feel or simple calculations. Building a complete model is essential.

Designing supply chain and logistics models for decision support depends on the nature of data available, considering various factors such as random disturbances (or uncertainties) in the system, worst case performance needs, and disruptions. The decision problems can be formulated as deterministic or stochastic as well as robust optimization problems. Robust optimization is preferred when it is not possible to define specific distributions to problem parameters as like in the case of stochastic optimization. Supply chain problems can be modeled in one among the three possible ways considering the following types of uncertainties:

Completely deterministic: This is a chain where, from the point of production to the end customer, goods are assumed to flow without uncertainties.

Stochastic: Stochastic chains experience relatively low-impact disturbances, e.g. transportation delays and product damage, which do not significantly affect the flow of goods but nonetheless require robustness so that the disturbances can be absorbed.

Completely disrupted: This state of the chain arises as a reaction to extreme disturbances, such as climatic upheavals and political unrest, among other factors.

Solving Decision Problems in Supply Chains and Logistics Networks

In supply chains, uncertainty exists at multiple levels, starting with demand planning, covering inventory management, transportation and delivery, and extends till pricing policies. A de-risked framework targets the following areas:

- Enabling visibility in the flow of goods from the point of produce to the end customer

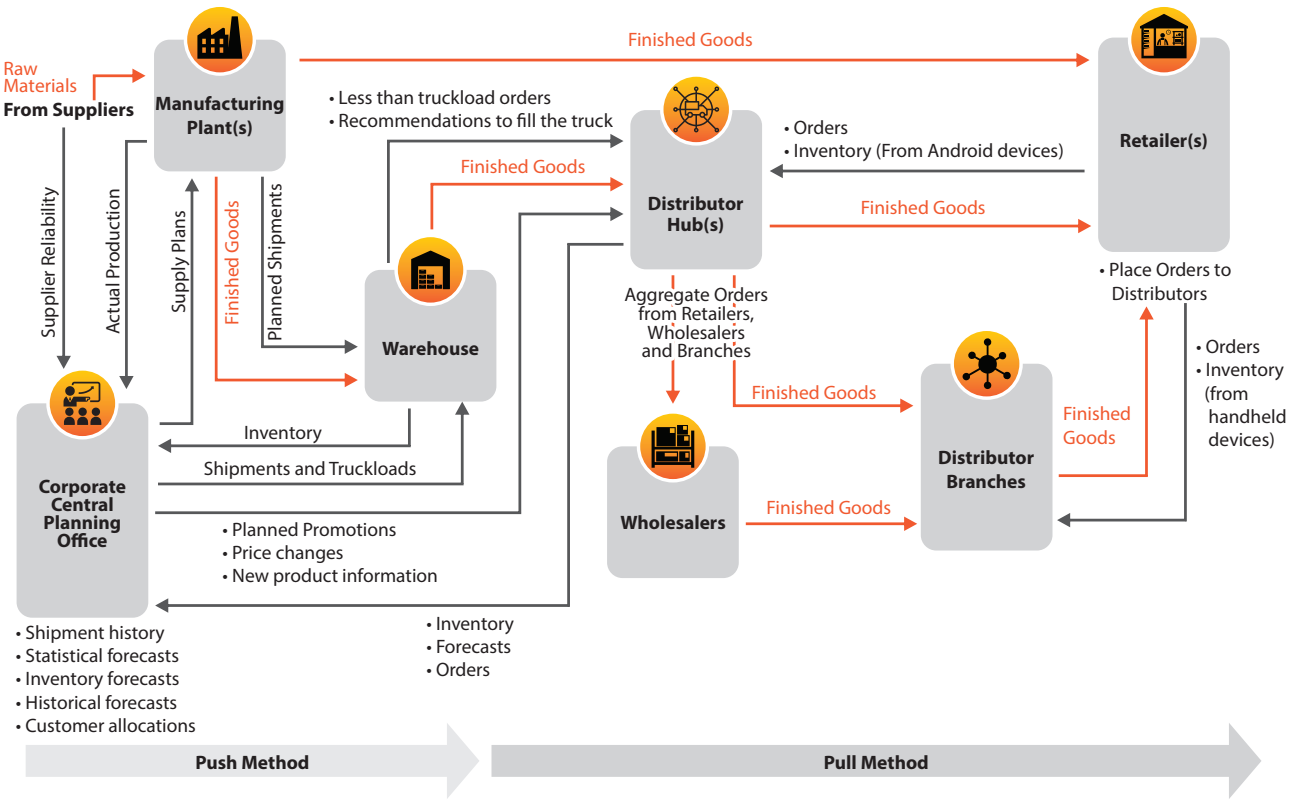


Figure 1: *Enabling New Capabilities in Supply and Logistics Networks*

- Forecasting end customer demand and aggregation of demand upstream up to production planning
- Storage and replenishment of goods at echelons and planning for safety stocks to meet uncertainties and disruptions
- Disaggregation of production volumes into allocation volumes from points-of-production to retailer echelons
- Capturing the utilization of vehicles used for transportation to understand the logistic efficiency of flow of goods between echelons

Figure 1 presents a comprehensive framework capturing new capabilities in Supply and Logistics Networks.

Thought leadership and research related to supply chain uncertainty have been directed towards supply chain integration, with the formulation of the corresponding decision problems into six broad categories.

Demand planning: The research work in our labs explores various time-series/machine learning algorithms to model the behavior of the future demand at various echelons (store, retailer, and distribution channel roll-up as well as at the production facility) in the supply chain. Demand forecast is the input to prepare tactical plans and production schedules as well as to conduct a "what-if" analysis. Autoregressive (AR) models with external inputs (e.g.: crude oil

prices and dollar value behavior) were used for a consumer product goods company to understand the necessary responses to the anticipated changes in market demand, competition behavior, and economic ups or downs for our customers in retail and transportation. Long short-term memory (LSTM) models were used to define a selective forget-remember behavior to predict the pricing in transportation and logistics capturing complex non-linear time sequences.

Inventory planning and replenishment: Inventory-related research using machine learning and artificial intelligence techniques and tools enables our customers to decide how much, when, and where to order, with minimum and maximum storage levels set for various products. Dynamic (periodic or continuous) and joint replenishment models are being used for a large multi-national CPG organization to capture the uncertainties arising out of randomness in the market demand which could have either small-scale or amplified behavior due to various factors. A joint optimization strategy has been proposed to one of our group companies for condition-based maintenance and spare part provisioning. Intermittent replenishment models have been further explored and extended to the requirements of our global customers in retail, automotive spare parts, electronic goods, as well as online B2B or B2C commerce.

Order allocation and demand fulfillment: The purchase orders placed by multiple downstream echelons (such as retail outlets) are sent to the next higher echelon (warehouse) in the chain. The warehouse requires to plan the distribution of the goods with the available inventory. We have deployed an order allocation and fulfillment system designed using stochastic mathematical modeling. As a result, we have solved the problem using a two-stage linear programming formulation at the production facility as well as the distribution centers to distribute baby care products based on production capacity, supply, and downstream demand. Deterministic mixed integer linear programming models were deployed for the repositioning of empty tank containers for a third-party container transportation provider of door-to-door services for bulk liquid and food-grade products considering the priority of customers, revenue, and margins possible.

Inbound/outbound logistics: The products or goods movement in various modes of transport (road/rail/ocean/air) could be formulated and solved as transportation problems, trans-shipment problems, and assignment problems, among others, considering uncertainties and disruptions. For one of our large European customers, we have formulated the problem of shipment of supplies of raw materials for cement manufacturing and delivering building materials, cement, and aggregates to customer orders to minimize the empty miles traversed by the trucks. The problem was solved using mathematical modeling and scenario-specific heuristics to bring in about 23% reduction in empty mile trips. As part of our ongoing research activities, we are exploring various types of routing problems

Our research explores various time-series/ machine learning algorithms to model the behavior of the future demand at various echelons (store, retailer, and distribution channel roll-up as well as at the production facility) in the supply chain

including last-mile delivery (from local distribution centers to retail outlets) for our customers in transportation and logistics operations for supply fulfillment.

Supply chain design: Location selection of production facilities, distribution centers, and warehouses considering road and rail connectivity between and within geographies, the nature of transportation carriers available, e.g. trucks and railway wagons, and the nature of the product (ambient conditions during transportation like refrigeration) under various supply chain scenarios are explored in this area. Worst-case operating conditions such as unavailability of production facilities or distribution centers due to natural calamities and disasters, among others, are formulated and modeled as mixed integer linear programming (MILP) or non-linear programming (NLP) problems for disaster relief and recovery needs.

Pricing and revenue management: This addresses categorization and classification of products and customers, thus fixing the sale price of the product, based on market needs and customer wants, so that the organization's revenues as well as its profits are maximized. We have been closely working with customers in the hospitality industry to derive the pricing definitions for hotel rooms based on room inventory availability and time-dependent market demand. Markdown price optimization models for clearance sales have been delivered to a leading art and craft retailer in North America. Extensions of our ongoing research in revenue management are related to pricing in retail chains and volume contracts for logistics service providers.

Figure 2 presents a conceptual structure for such frameworks. Replenishment planning, together

An integrated supply chain system uses analytics and key performance indicators (KPIs) to measure the performance of the flow and the storage of goods across the system

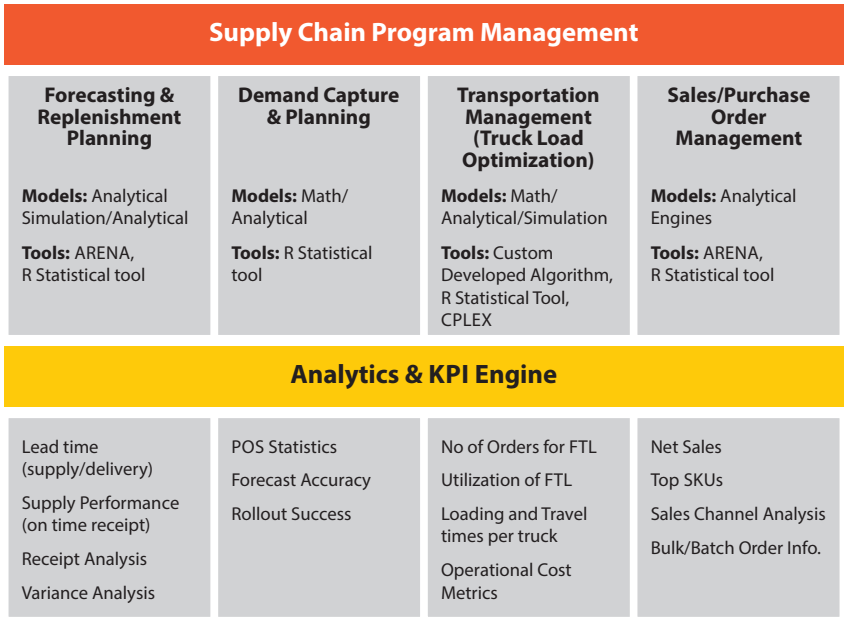


Figure 2: *Concept Diagram of a Supply Chain Framework*

with prediction or forecast models, enables supply chain managers to plan more efficiently by foreseeing future demand (typically, two to three months ahead). This helps them to decide how much inventory, and when and where to order from. An integrated supply chain system uses analytics and key performance indicators (KPIs) to measure the performance of the flow and the storage of goods across the system, to minimize the total cost of locked-up inventory, increase inventory turns, and maximize the flow of goods to the end customer, thus minimizing shortages.

The components presented in Figure 2 can be presented in an architectural form as depicted in Figure 3, with an element for business value assessment and monitoring to understand the uncertainties and, if possible, measure the risk associated with the uncertainties.

Frontiers for the Future

With select customers, we are currently exploring the following:

- Opportunities in the digital space, investing in technologies that create supply chain visibility
- Prediction and mitigation of risk, as part of our decision sciences research
- Bi-modal supply chains that enable organizations to differentiate normal from abnormal and evolution from revolution, and drive continuous improvement in managing disruption
- Enabling our core capabilities in multimodal logistics network modeling and inventory (or resource) optimization competencies, in line with Industry 4.0 trends
- Creating frameworks and artifacts linking IoT, big data, analytics, and other technologies

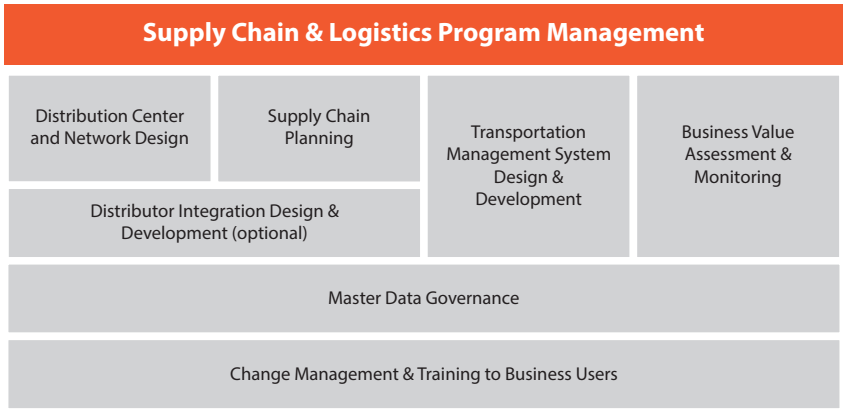


Figure 3: *The Architecture of a Supply Chain and Logistics Framework*

Supply Chain Reengineered for Better Inventory Turns

A leading multi-national consumer Packaged Goods (CPG) company manufactures and sells products ranging from detergents to nappy pads across multiple product categories such as home/fabric care and beauty. These stock keeping units (SKUs) are distributed across India to about 40 distributors in different states from factory warehouses near the manufacturing facilities. The SKUs are, in turn, further shipped to over 100 sub-distributers and modern retail hubs who then take care of demand fulfillment to about 1.4 million retailers.

The CPG company foresaw huge growth potential in India. However, one of its product category lines had been operating at a loss for the last few years. The company also saw an opportunity to revamp the distribution supply chain considering the impact of new GST norms in India from mid-2017. The customer requested the TCS Business Teams and the TCS Corporate Supply Chain Research & Innovation Team (SCR&I) to work out a proposal to re-engineer and improve delivery efficiency by enhancing the customer's supply chain and logistics distribution capabilities to ensure that the right product reaches the right place at the right time across the country.

The first phase of study involved interacting with three to five key distributors at select locations across the country to understand the existing processes and challenges in the distribution network. The following were the observations:

1. The SKU demand is captured in spreadsheets based on supply invoices at distribution centers (not based on purchase orders placed by sub-distributors) and communicated to the CPG company by email.
2. Order allocation from distributors to the next level of distributors was conducted without sensing either customer buying behavior or the market demand, using KPIs.
3. The CPG company used a "push" distribution strategy based on the demand captured from major distributors; this translated to pushing the goods to the end retailer from the sub-distributor.
4. The absence of a demand forecasting process at distributors, sub-distributors, and at retail outlets led to the entire chain relying on the CPG company's annual estimates and market data.
5. A higher bullwhip effect was clearly seen on shipments getting lumped quite frequently at CPG depots, distributors, sub-distributors, and retail outlets. The truck trip frequency increases at the end of every month for fast moving SKUs.

The major challenges that were provided as feedback from the retail outlets and customers were the absence of items (and in particular, fast-moving items) on the racks at many outlets and lower inventory turns at locations where stocks were available.

With the above inputs on the nature of the existing supply chains and scope of logistics, the TCS teams working together with the customer's supply chain team attempted to design and deploy a pilot solution quickly addressing the following aspects:

1. Introduction of an adaptable time-series forecasting process and rule-based inventory management polices at sub-distributors, major distributors, as well as with select modern retailer outlets
2. Definition of easy-to-use rule-based ordering policies with minimal use of IT tools at distributor's end to minimize accumulation of demand with the upstream supplier
3. Enhancement of visibility between the CPG company and major distributors around the on-hand inventory, on-road/booked inventory, and projected sales orders based on the company's monthly plans and projected supplies to sub-distributors
4. Classification of inventories based on inventory velocity. We wanted to ensure that the truck load utilization is improved and kept close to a level over 95%

The CPG company could see better streamlining of information flows with predictive insights into downstream orders as well as upstream orders at any given supply chain node (e.g. distributor or sub-distributor) with demand visibility to the CPG company's demand planning team. This enabled them to reach the product from the manufacturing facilities to retail outlets faster.

In the second phase, the study intended to enhance the supply chain with decision support models for automation of distribution. Here, the supply chain was to use the pull-based process involving only large or centralized distributors at each state or region. The modules that were included in the final rollout covered the following:

Demand forecasting: Use of time-series and machine learning models

Inventory planning and optimization: Use of continuous, periodic inventory review policies with safety stocks to minimize stock outs

Order management: Use of math linear programming models to allocate SKU inventories to downstream channels based on project demand and target performance considering surplus and deficit in manufacturing

Inbound/outbound logistics: Application of mixed integer programming models and customized algorithms to determine the number of trucks with trip schedules between upstream and downstream nodes to distribute SKU inventories and improve truck load utilization

The above modules were implemented in a phased manner. The company saw a sea change on replacing push-based distribution with a pull-based chain in such a short period of time. The results offered a reasonably significant increase in the inventory turns in just one year.

Viswanath Kumar Ganesan

Viswanath Kumar Ganesan, Senior Research Scientist, Solutioning & Innovative Offerings at TCS Research and Innovation, has worked in retail, CPG, transportation, telecom, human resources, and business process services on transformation in workforce planning and deployment, supply chains, transportation, and logistics. He focuses on stochastic modeling and optimization for decision support. A former Research Fellow in the Singapore-MIT Alliance at Nanyang Technological University, Ganesan is on the Board of Studies of the School of Electrical & Electronic Engineering, SASTRA University and the Coimbatore Institute of Technology.

Ananda Padmanaban Srinivas

A.P. Srinivas, Head, Solutioning & Innovative Offerings at TCS Research and Innovation, has over 30 years of experience in IT. He has designed, developed, and delivered advanced technology solutions in railroad traffic control, telecom switching, internet content management, internet banking, electronic currency and equity trading, satellite TV subscriber billing, gamified idea crowdsourcing, sentiment analysis, advanced optimization techniques for static and mobile workforce management, human capital evaluation, and technology insights gathering.

Soft and Functional Materials

Multi-scale model-driven methodology for design of materials, processes, and devices

Venkataramana Runkana, Venkata Sudheendra Buddhiraju, et al.

IN BRIEF

Materials play an important role in our day-to-day life across food, personal care, health care, or structural products. A majority of soft and functional materials are produced on an industrial scale and used routinely. However, several challenges are faced during the design and development of new products and delivery systems. We describe here TCS' research on modeling, simulation, and synthesis of soft and functional materials and their applications in energy storage and drug delivery.

Modeling, simulation, and optimization reduce time and cost to market in the development of new products

Materials that can be easily deformed by thermal stress or fluctuations at room temperature are referred to as soft materials, for example, polymers, gels, colloids, granular materials, and most soft biological materials. Materials having native properties of their own like magnetism, piezoelectricity, ferroelectricity, or energy storage functions are referred to as functional materials (for example, ceramics, metals, polymers, and some organic molecules). While a majority of these soft and functional materials are produced on an industrial scale and used routinely, several challenges are faced during the design and development of new products and their delivery systems. A few such challenges are:

- Scaling up production of nanoparticles and their suspensions
- Creating coatings that can withstand high-temperature and high-pressure environments
- Controlled and targeted delivery of active molecules like drugs, fragrances, etc.
- In silico testing of health and personal care products to avoid animal testing

Design and development of new products and delivery systems require extensive time, effort, and money. Modeling, simulation, and optimization can greatly help reduce development and testing time as well as cost to market. Since these materials are utilized in different forms in diverse applications, it becomes necessary to develop a multi-scale modeling approach, as depicted in Figure 1. For example, in the design of drug delivery devices like inhalers, it is necessary to understand the behavior of

Fact File

TCS Research: Research and Innovation for Manufacturing and Engineering

Outcomes: Modeling Frameworks for Synthesis of Nanoparticles and Design of Drug Release Devices; Experimental Protocols for Synthesis of Nanoparticles and Nanofibers

Principal Investigator: Venkataramana Runkana

Academic Partners: IIT Kanpur, IIT Bombay

Techniques Used: Multi-scale Modeling; Model-based Optimization; Bench-scale Experiments

Industries Benefited: Chemicals, Materials, Semiconductors, Life Sciences

Patents: Filed–14; Granted–7

Papers: Journal articles–15; Conference papers–40

molecules through molecular dynamics (MD) simulations. Information from these simulations is used to design particles or carriers for a drug molecule through population balance modeling (PBM). The flow of particles through the inhaler using discrete element modeling (DEM) and computational fluid dynamics (CFD) is then simulated for designing the device or the delivery system. Similarly, a multi-scale modeling approach is required not only for the design of equipment such as furnaces, but also for understanding their performance.

By synthesizing electrodes from materials that are stable in liquid electrolytes, higher energy densities can be achieved

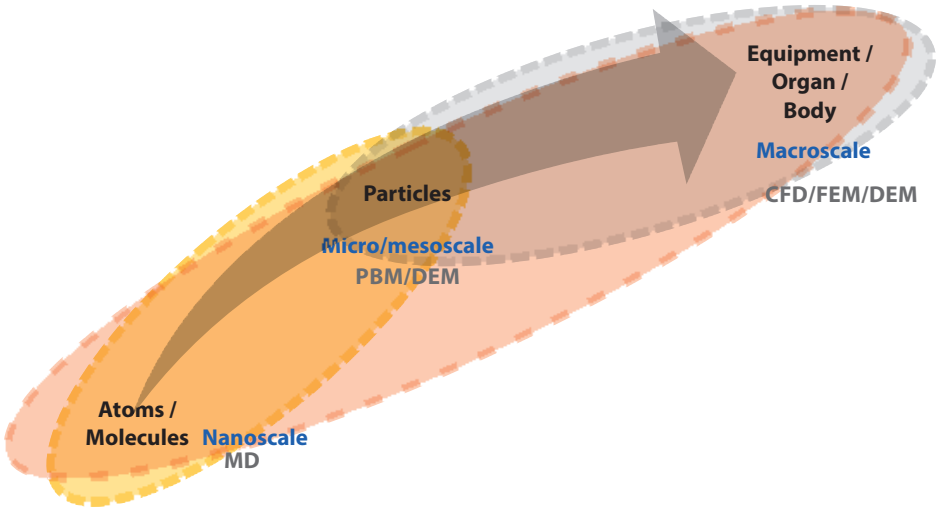


Figure 1: *Multi-scale modeling of materials, processes, and delivery systems*

Energy Storage and Functional Nanofibers

The demand for renewable energy sources calls for efficient and reliable electrochemical energy storage devices. Lithium-ion batteries and supercapacitors are the most promising candidates. Supercapacitors are preferred over lithium-ion batteries in applications like electric and fuel cell vehicles, smart electric grids, pace makers, etc. as they acquire higher power density, faster charge/discharge times, and better cycling stability. Moreover, they are safe and environmentally friendly. However, they suffer from low-energy density which hinders their usefulness. One of the challenges is the development of electrode materials for supercapacitors with high-power and energy densities compared to batteries. Nanofibers are suitable for this, and electrospinning (see Figure 2) is probably the most convenient technique for producing nanofibers continuously.

In collaboration with Indian Institute of Technology-Kanpur, we have synthesized novel hierarchical nickel molybdate nanostructures, carbon–nickel molybdate (NM) composite nanofibers and nickel–zinc ferrite (NZF)–carbon nanotube (CNT)–carbon nanofiber (CNF) composites for making electrodes for supercapacitors.

- First, we synthesized solid, porous, and hollow nanofibers and microplates of NM. An asymmetric supercapacitor fabricated from hollow NM nanofibers had a high specific capacity of 175 C/g (at current density of 1 A/g) and a capacity retention of 97% after 5000 cycles. At this current density, the device delivers an energy density of 39 W-h/kg and a power density of 798 W/kg.
- Second, we synthesized four different electrospun composite nanostructures, with different ratios of NM to carbon. An asymmetric supercapacitor fabricated from a composite with 50 % wt NM had a specific capacity of 135 C/g (at current density of 1 A/gm) and a capacity retention of 92% after 3000 cycles.
- We have also synthesized a ternary nanocomposite consisting of NZF

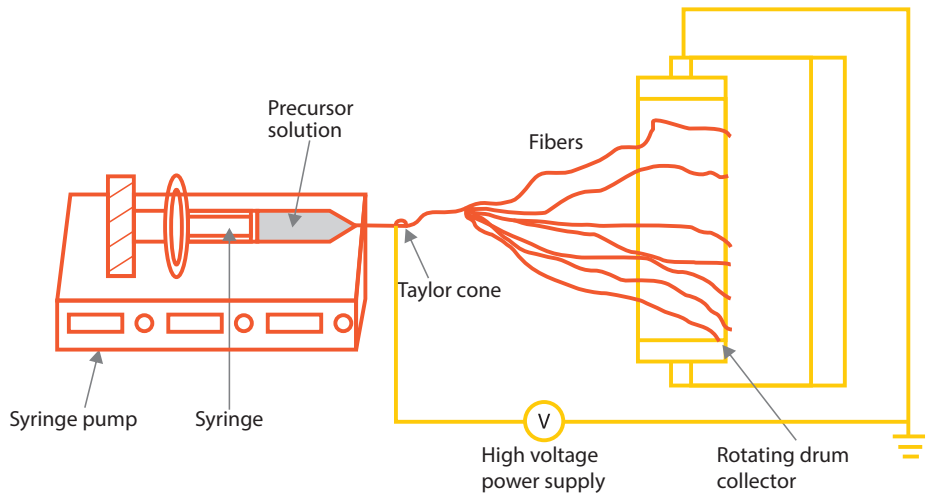


Figure 2: *Schematic of synthesis of nanofibers by electrospinning*

(Ni0.5Zn0.5Fe2O4) nanoparticles and multi-walled carbon nanotubes (CNT) dispersed uniformly in hollow carbon nanofibers (HCNF). These NZF–HCNF–CNT composite nanofibers (with 30% weight of NZF) exhibited a specific capacity of 335 C/gm (at 1 A/g), much greater than pure NZF fibers (180 C/g) and hollow carbon nanofibers (96 C/gm). A capacity retention of 90% after 3000 cycles was achieved with these fibers, while pure NZF fibers had only 62%.

The practical application of our 50-NiMo-HCNF//AC asymmetric supercapacitor was shown by arranging three as-prepared supercapacitor cells in series and illuminating a red light emitting diode (LED). The superior performance of all these materials could be attributed to their mesoporous surface with a high specific surface area (~105 m^2/g), the hollow core, and the interconnectedness of the fibers.

These electrodes can be used in electric vehicles for quick charging, engine start-up, and power bursts. The future challenge is to synthesize materials with improved surface properties so that energy density can be improved further. Generally, materials with abundance of mesopores (pore size ~2–50 nm) and nanopores (pore size ~2 nm) exhibit high specific capacity. However, controlling the pore size distribution is a challenge. An alternate way to increase energy density is to increase the working voltage of the electrolyte. We employed aqueous electrolytes which produced a maximum voltage of 1 V. Organic electrolytes and ionic liquid electrolytes can operate at higher voltages of 2 V and 3 V, respectively. By synthesizing electrodes from materials that are stable in these electrolytes, higher energy densities can be achieved. Hybrid storage devices made from supercapacitors and lithium-ion batteries are also promising since they produce higher levels of energy and power densities. Such devices can be assembled from electrodes that are fabricated from high-capacity battery materials and high-rate capacitive materials.

Nanomaterials: Model-based Production Processes

Nanostructured materials such as carbon black, metal oxides (titanium dioxide, fumed silica, ceria, zinc oxide, and alumina), and other high-purity materials (advanced ceramics, semiconductors, super alloys, and thin films) have diverse industrial applications in pigments, catalysts, and optical waveguides. The final product particle properties (size, shape, and distribution) determine the end application of a product. For example, nearly-uniform titanium dioxide particles are used for making paints and pigments, whereas aggregates of large nanoparticles are required for optical fibers. Despite their widespread application, control and optimization of the final product particle properties and large-scale production of nanoparticles is still a major challenge. Both liquid phase (sol-gel, coprecipitation, microemulsions, hydrothermal synthesis, etc.) and gas phase (aerosol flame synthesis, laser ablation, and electrospray synthesis) methods have been successfully demonstrated on a laboratory scale to produce diverse nanomaterials. However, these methods are not yet completely viable on a commercial scale owing to the difficulty in scaling up these processes.

We have established bench-scale facilities in our laboratories to produce nanoparticles of titania, ceria, silica, etc. through both liquid phase (hydrothermal) and gas phase (flame aerosol) synthesis methods, and developed a comprehensive modeling framework, combining computational fluid dynamics, population balance modeling, thermodynamics, and reaction kinetics, for process design and for scale-up (see Figure 3).

Development of novel materials such as core-shell metal complexes (ceria/titania, ceria/silica, Pt/titania, Pt/silica, etc) with relatively inexpensive precursors; design and development of novel methods to produce these materials on a large scale are key challenges in this field. Model-based design and development, in conjunction with data from laboratory experimental studies, plays an important role in enabling the production of novel nanomaterials.

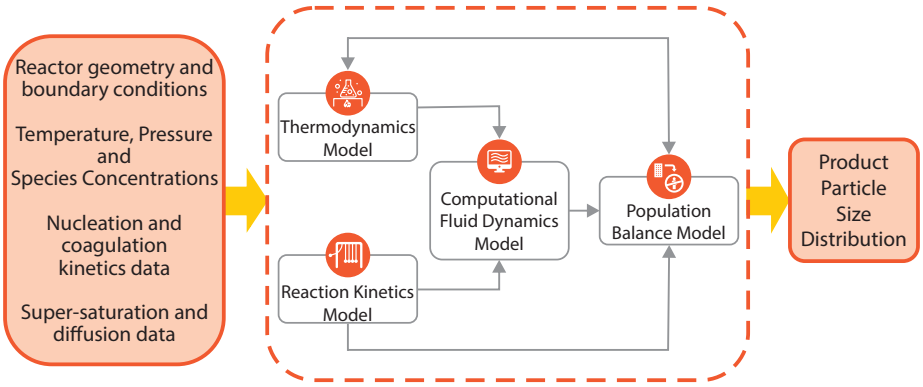


Figure 3: *Comprehensive mathematical modeling framework for nanoparticle production*

Drug Delivery Systems

The design and development of smart materials is an important and emerging field that is expected to have far-reaching implications in health care, especially as drug carriers, medical implants, artificial organs, and so on. Traditional carriers of drug molecules result in limited drug efficacy, undesirable temporal changes in drug concentration, and patient noncompliance. Polymeric systems such as biodegradable particles, stimuli-responsive hydrogels, and fibers have been identified to possess desirable properties that can provide targeted and controlled drug release. Designing these novel drug carriers is a challenge since selecting the correct polymer or polymer blend, functional agent, quantity, in addition to the carrier's shape and dimension, is critical.

Oral drug delivery is considered the most preferred method of patient compliance. We have devised a model-driven approach to design pH-sensitive hydrogels for

oral delivery as illustrated in Figure 4. Our mechanistic model estimates the charge distribution of the hydrogel, its deformation in response to swelling pressure and elastic forces exerted by the elastic nature of cross-linking molecules, and migration of ions and drugs in response to changes in environmental conditions. The model was parameterized in terms of experimentally tunable formulation parameters such as cross-linking ratio, dimensions of hydrogel, and polymer concentration to enable its use as an in silico experimentation tool. The strategy was also extended for design of glucose-sensitive hydrogels.

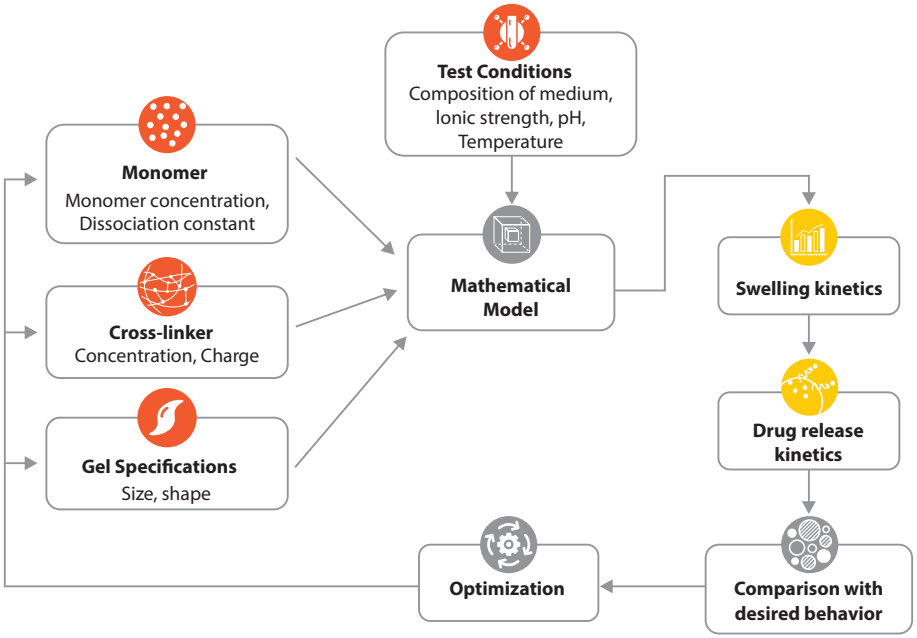


Figure 4: *Model driven methodology to design stimuli-sensitive polymer hydrogels or particles*

Looking into the Future

The long-term need is the development of an integrated knowledge-based platform that could be employed by researchers as well as practitioners for designing and developing new processes, functional materials, and novel delivery systems for applications in health care, foods, cosmetics, and agrochemicals. Here, we articulate our thoughts and ideas on the development of such a platform, for example, the design and development of materials and delivery systems for health care.

A major challenge for the pharmaceutical industry is designing delivery systems for proteins and peptides and their corresponding drug formulation. The drug pre-formulation process requires the characterization of a drug's physical, chemical, and mechanical properties so as to choose what other ingredients (known as excipients) should be used in the preparation. In protein pre-formulation, the important aspect is to understand the solution behavior of a given protein under a wide range of stress conditions such as freeze/thaw, temperature, shear stress, mechanisms of degradation, and

therefore its mitigation. It is also necessary to consider factors such as particle size, polymorphism, pH, and solubility, as all of these can influence bioavailability and hence the activity of a drug. The drug formulation and its carrier depend on the drug and its properties, properties of the potential carrier, delivery route, and disease, among other factors.

While design of formulations is carried out using software tools such as Materials Studio, computational fluid dynamics software tools such as Fluent and COMSOL are utilized for undertaking release simulations. However, these activities are carried out independently and are not solved in a coupled or integrated manner. Simulations related to drug formulation, drug release, and in vitro or in vivo testing require modeling at multiple levels of scale, starting all the way from atomistic simulations at the nanoscale to drug transport at the macro-scale. The framework for such an integrated knowledge-based platform with multi-scale modeling capabilities is depicted in Figure 5, which addresses some of the challenges mentioned above, taking into account the necessary inputs, for the design and development of soft and functional materials and their delivery systems or devices.

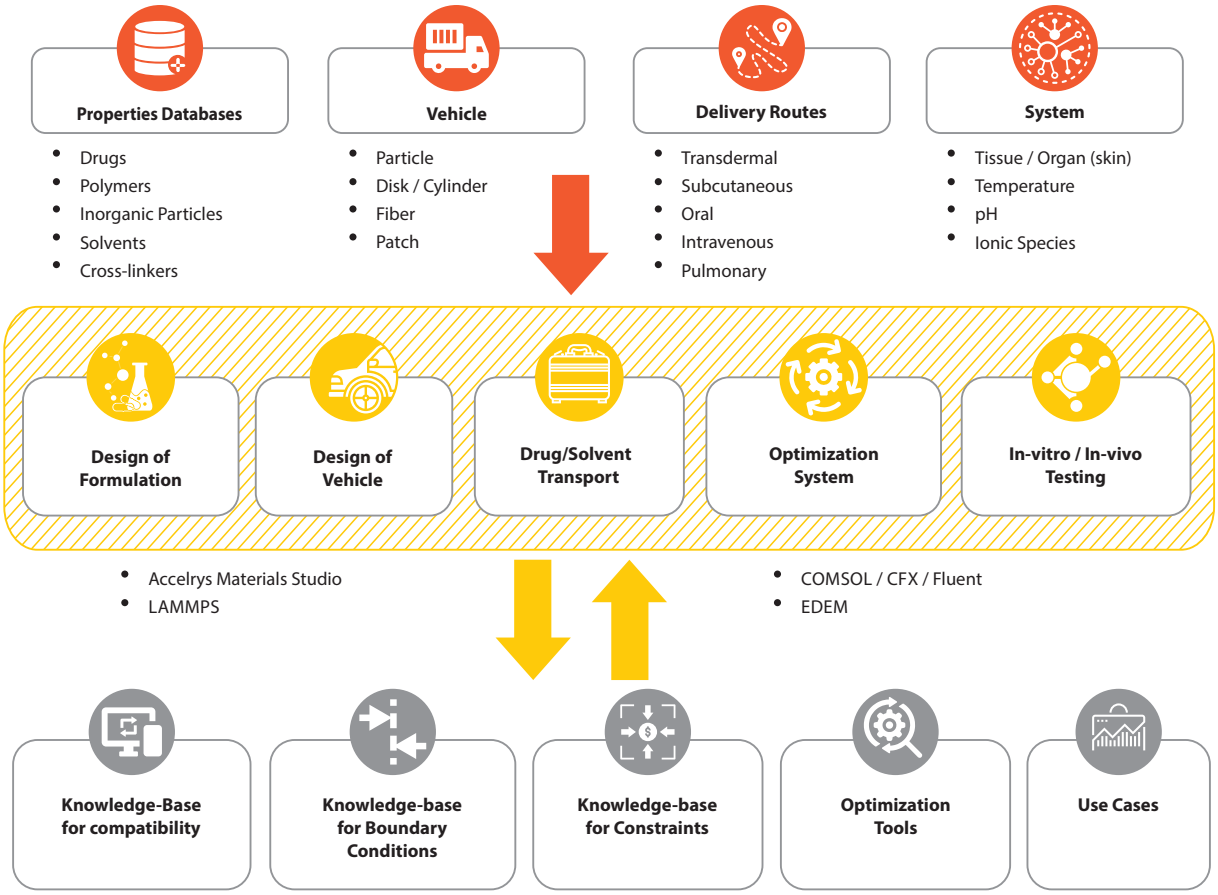


Figure 5: *Framework for an integrated, knowledge-based design and development platform for creating soft, functional materials and their delivery systems in healthcare*

Venkataramana Runkana

Venkataramana Runkana is a Chief Scientist at TCS Research and Innovation and heads the Research Programme for Manufacturing & Engineering and Industrial Services Business Units in TCS. He has more than 27 years of experience in process modeling, simulation, and optimization, advanced data analytics and digital twins, process development, scale-up and design, nanomaterials and drug delivery systems. Venkat received the TCS Distinguished Scientist Award in 2014 and was an AICTE-INAE Distinguished Visiting Professor at IIT Kanpur during 2013–2018. He is a chemical engineer and holds a Ph.D. from Columbia University, New York.

Venkata Sudheendra Buddhiraju

Venkata Sudheendra Buddhiraju is a Scientist at TCS Research and Innovation and is part of the Research Programme for Manufacturing & Engineering and Industrial Services Business Unit in TCS. He has more than 11 years of experience in process modeling and simulation, computational fluid dynamics, discrete element method modeling, population balance models, nanostructure synthesis and applications. He has presented publications in international journals and conferences. He is a chemical engineer and holds a Ph.D. from IIT Kanpur.

Nagaravi kumar Varma Nadimpalli

Aditya Pareek

Researcher

Digitization in the Enterprise

Innovation Culture

TCS has built a vibrant "Innovation Culture". It has created an environment where anyone can contribute to innovation. Through crowd sourcing, TCS is able to mine the cognitive surplus of associates spread far and wide, to find the most creative solutions to solve business problems. Each week, a hackathon or an ideathon is conducted to solve open problems that teams face. Some of these events are directly sponsored by customers and rewarded by them. Innovation events offer employees a chance to do exciting things not related to their daily tasks.







Digitization in the Enterprise

K. Padmanabhan

One of the most important areas in enterprise management is managing risk. Broadly speaking, there are two big areas of risk: those in "running the business" and those in "changing the business".

The former is called operations risk. In TCS, a large fraction of the business consists of executing projects for customers (there are over a 1000 of them at any one time) and developing and implementing products for customers. Each of these has considerable risks and managing these presents considerable challenges. This is especially true when project teams are distributed around the world and the expertise needed to solve a problem is not locally available.

TCS' approach to this is to adapt solutions from various technologies. The Japanese principles of employee participation, reducing variability, ensuring high visibility, and self-help play a big part in this process. Other principles include building institutional knowledge, learning from other people's experiences, and the rapid escalation of persistent risks.

There is inherent variability among projects. So, we use our expertise to categorize projects accordingly. Our adherence to the Capability Maturity Model ensures that we define processes for each type of project. This reduces the variability between projects and ensures that the experience of one project is readily applicable to many other projects.

TCS invests heavily in Knowledge Management (KM), a system which captures the unique experience of each project. By accessing the KM system, different projects learn the adage that "there is nothing new in this world". They are able to adopt the successful interventions of other projects to mitigate the risks in their projects.

We also have secure methods of communicating within the enterprise using adaptations of social media. This enables us to mitigate risks by rapidly finding an expert in a particular area, and tapping his/her organizational expertise.

Visibility is another key component of our strategy. We use technology to make the risks visible, and the suggested mitigations available. People are thus "nudged" into taking the required action. This system also enables rapid escalation of aging risks, and enables supervisors to act as fire hydrants to douse minor fires before they become threatening infernos.

Another big focus for risk management is the area of "changing the business". The issue here is introducing changes, not only to the company, but to its environment. For a company like TCS, this means introducing new products and services, while being able to scale them up when required. In addition, TCS needs to be able to create an environment where innovative culture exists, so that new offerings are accepted easily.

It is necessary to keep pace with technology and regularly introduce new offerings – a fact that is especially true for a technology company like TCS. To do this, TCS needs to both increase its research portfolio, and make the offering introduction timely. TCS' Co-Innovation Network (COIN) provides an answer to both these issues. Though a part of COIN (its tech arm) deals with leveraging technology startups to improve time to market, its academic arm, leverages the research capability of academia to address difficult problems, as discussed in this chapter.

TCS recognizes that if there are too many unsuccessful offerings, its reputation could suffer. To determine if an idea is ready to be offered to customers, and whether it should be a new product or a new service, TCS has created a New Products and Services Development process. This adapts the industry's best practices of

assessing technology readiness, listening to the market, testing the market, and scaling up. It also provides objective measures so that the governance of this process is possible even if several of these evaluations are happening simultaneously.

Another serious problem is the selection and scaling up of ideas. The selection of viable ideas that can be implemented using emerging technology requires using considerable expertise and knowledge of the technologies. It also requires a number of related capabilities like marketing, legal, and market testing. Scaling up often requires a very agile adaptation of the product, often pivoting in a new direction, when necessary.

To facilitate the selection of ideas, and the implementation and scale-up till it is viable, TCS has set up an incubation team to design a set of structures and processes. This team also churns out entrepreneurs using a magic potion of training, mentorship, and infrastructural support. This team has evolved a mature process with a high success rate.

To facilitate the development of an environment where innovation is widespread, TCS has evolved a mentorship program, Digital Impact Square (DISQ), to encourage social innovation using digital technologies. This provides a nucleus of bubbling innovation that can, in a self-sustaining manner, transform societies. This culture also provides a multiplying factor to the utility of the offerings that TCS brings to market.

A severe test of our entire risk management processes has been posed by the ongoing pandemic and the resulting lack of co-location. Though we need to have more time to evaluate our success in effectively managing the risk under such stressful situations, our preliminary assessment is that the risk management processes appear robust even under these conditions. We need, however, to monitor the processes, and adapt them as necessary.

We are continuing to evolve in our journey to apply technology and experience to effectively manage risk. The current snapshot of our evolution is on display in this section.

K. Padmanabhan

K. Padmanabhan (known as K. Paddy) is a consulting advisor, formerly Vice President, TCS. K. Paddy is a highly innovative leader and has been responsible for a number of innovations introduced in TCS, including many in the quality and tools areas. He has been closely associated with TCS' first Research Center, TRDDC, for many years.

He is currently responsible for the interface between the Technology organization and the various business units in TCS. He is especially involved in evolving Cloud Computing as a strategic initiative for TCS.

K. Paddy has a Master's degree in Computer Science from the Indian Institute of Technology-Madras. He has been with TCS for about forty years.

Mitigating Risks

New tools to manage operations and finance risk

K Subra,
A K Balaji, et al.

IN BRIEF

Any modern enterprise faces a complex set of risks that need effective identification and management. The larger and more global the enterprise, the more likely it is for the range and variety of risks to increase exponentially. This occurs due to the locality of the operations in each market and the limits of organizational attention to balance between growth and risks. Further, the risks span a range of domains such as regulatory, technologies, market factors, and resources. In this essay, we focus on newly developed tools that can manage risks in operations and finance.

Currently, there are approximately 800+ contextual risks in the Risk Repository, over 1400 recommendations in the Recommendations Repository, and more than 1600 questions in the Self-Assessment Questions Repository

Operations Risk Management

TCS provides production support services for 19 of the top 25 financial firms in the world. It also provides business operations support to 11 of them, handling volumes of more than 750 million transactions per day. It has similar operations stories across other industries. Even a single, unplanned disruption for a few minutes will result in multibillion dollar losses and lost opportunities. Hence, it is critical to master the science and art of managing operations risk. This is an area that is ripe for major innovations.

Sources of risks in operations arise out of interactions among factors that are inherently multidimensional—diversity of players, organizational factors, work practices, technologies of components, social factors, human factors, and interfaces across a myriad of systems and protocols. While standards such as ITIL have focused on best practices in operations, they have fallen short of recognizing the scale and complexity of modern-day operations, and embracing the opportunities that open up through digital technologies.

Through its RiO (Rigor in Operations) initiative that began in 2014, TCS has created an intense focus on continuously improving operations and bringing in new ideas and practices to reduce and eliminate risks, both at the customer

engagement level and at an enterprise level. In this essay, we will sample a few key ideas that have shaped risk management in our operations services.

Areas of innovation in RiO

- Operations risk assessment and benchmarking
- Visualization and gamification to enhance human attention
- Automated health check and alert monitoring
- Cognitive automation and self-healing

Operations Assessment for Holistic Risk Management and Benchmarking

To enable continuous risk identification and mitigation in production support, an end-to-end digitized assessment platform called the "RiO Workbench" has been conceptualized and developed.

The Workbench consists of the following features:

- An intelligent algorithm to analyze the inputs provided to the platform as part of the self-assessment by the production support team
- The ability to identify contextual operations risks
- Recommendations to mitigate risks
- The ability to automatically assess the Operations Excellence level and provide a comprehensive assessment report without the need for any external assessor
- Assistance in capturing the organizational knowledge in a structured way

A patent has been applied for the RiO Workbench assessment framework and another for the processing algorithm. RiO Workbench is being continually enriched with emerging practices, recommendations, and operations risks and kept updated each year. Currently, there are approximately 800+ contextual risks in the Risk Repository, over 1400 recommendations in the Recommendations Repository, and more than 1600 questions in the Self-Assessment Questions Repository. The repositories hold the content for 14 distinct process areas: Event, Incident, Problem, Service Request, Batch Operations, Change, Release and Deployment, Data, Access, Knowledge, Shift, Communication and Escalation, Competency, and Business Continuity Management.

RiO Workbench enables production support engagements to enhance their operations through the following steps:

- The engagement self-assesses based on answers to a questionnaire for all or any of the 14 process areas.
- An intelligent algorithm then automatically analyzes the answers.
- The algorithm generates an assessment report along with an Operations Excellence summary chart, subprocess area level findings, recommendations, and contextual operations risks.
- The outcome of the assessment enables the creation of a comprehensive risk register related to inherent IT operations risks and the adoption of best practices from TCS' experience with global customers.
- The evaluation of the Operations Excellence level (as Basic, Standard, Advanced, Best in

As part of the RiO best practices implementation drive, operations support teams have implemented visual demarcation techniques to prevent manual errors

TCS is leveraging its Machine First Delivery Model (MFDM™) to transform manual, error-ridden operations to error-free, efficient operations

Class) for each of the assessed process areas is conducted, based on process, practices, and KPI performance.

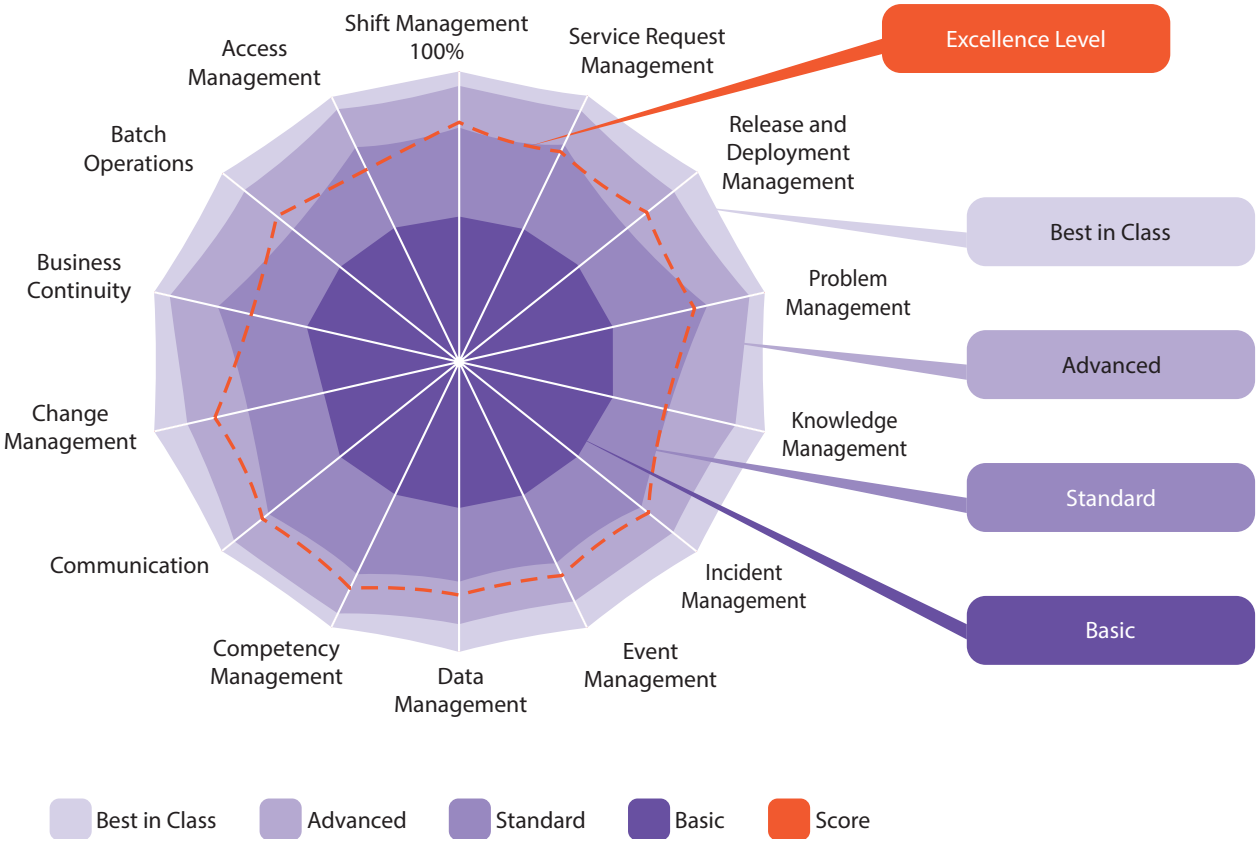
- This is followed by the creation of an improvement plan and tracking implementation.
- Account performance is benchmarked against TCS or/and industry best practices.
- Reassessment to understand the causative factors contributing to the Operations Excellence level is conducted. This exercise involves identifying the contextual operations risks and continually improving until Best-in-Class operations are achieved.

The benefits witnessed across production support engagements in TCS include:

- Mitigation of over 2500 contextual risks by implementing over 3500 recommendations across accounts
- Improvement of overall operations CSI from 91.3% (FY 15) to 93.5% (FY 19)
- Enhancement of the risk management attribute score from 87.4% (FY 15) to 90.6% (FY 19)

As a result, TCS is named as the Best Supplier for *"Reduced Risk and Disruption to Business"* by one of its major customers.

Figure 1: *RiO, A Digitized Platform to Assess Production Support Excellence*



Finance Risk Assessment Automation

A US $46 million deal with a North American customer is to be signed on Thursday this week. The contract draft that the business team has put together after negotiations with the customer will reach the finance team at close of business on Monday. The CEO expects a financial risk assessment and some recommendations for mitigation by Wednesday evening.

The issues with this contract are several. It is not in a single machine-readable format. The contract, running to 110 pages, is a set of documents consisting of word documents, a spreadsheet and two PDFs. One of the Word documents has a .jpg image of a delivery schedule. There are many regulations and compliance-related clauses in the contract that have financial implications; on average, there are eight clauses per page that require immediate attention. Many terms need computation, such as "the sum of all SLA percentages may not exceed 250% of the At Risk Amount." The document has to be read back and forth: "SLA percentages as set forth in sections 8.4.1 and 8.4.2, and subject to modifications as in section 9.3.1" and so on.

Another important factor is that many of these clauses impact many teams. For instance, there is a clause (in layman's terms) which states that "any associate who has served on this project, when moved out of the project, must not join a 'competitor' project without a cool-off period of six months, or else incur a breach of trust, and pay compensation." This risk has more to do with HR and the business team, while finance is in FYI.

It is possible that the business team, in its eagerness to close the deal, may have overlooked some risky clauses. The finance team cannot afford to do so. It has to read the documents and collate the high risk clauses at the very least, and recommend mitigation measures within two days. Understandably, this puts pressure on the finance team. As a result, while conducting manual assessments, the team is likely to miss an obligatory clause or two, which will prove costly in the future.

All legal, finance, and IP teams are familiar with scenarios similar to the one stated above. Contracts have become increasingly complex today and hardly correspond to any single template. Ever since the 1990s, the number of regulations and compliances have gone up by the dozen, across domains. These include HIPAA, SoX, GDPR, TUPE, and many more. Regulations and compliances may relate to data storage, encryption, privacy, IP, and other areas. Customers would like the contract to be explicit about their rights and obligations on matters relating to regulations, compliances, and terms of work. There are also several sections to a contract, with some sections relating to multiple teams. Ideally, a central team tracks the various rights and obligations of the company and alerts the business delivery, IT, IP, Legal, Security, HR, and other teams on the relevant clauses of the contract. This ensures that there are no lapses due to gaps in knowledge. There are off-the-shelf products for contract management, but they do not fit every scenario that a company might face. Many are cloud-based applications and, for the most part, companies are wary of using this method.

TCS is in the process of building a platform that can be used by one central team to track contract clauses across teams and timelines. The platform is slated to accommodate the following:

- A document reader
- A database of clauses and meta clauses
- A capability to extract relevant clauses
- An alert triggering capability for risky obligations and important rights
- An interaction and tracking mechanism to work with multiple teams
- A machine learning algorithm to reduce manual effort

Meanwhile, the TCS Corporate Finance team has already set to work on an effective mechanism for identifying the individual commitments and promises that are being made in contracts to ensure effective financial risk management, both for internal business units as well as external performance.

The team has come up with standard templates, checklists, and dos and don'ts which enable businesses to quickly respond to customer requirements on contract clauses as well as on the overall red lining of contracts.

The Contract Finance Risk framework systematically captures contracts using abridged standard sentences called "meta clauses." The framework has nine different attributes for a meta clause to slice and dice data. The framework, when deployed centrally, provides a quick view of risky clauses and allows for speedy action to be taken as a result. Currently, work is in progress to come up with a Master Services Agreement (MSA) checklist that will be designated to provide guidance to businesses on what clauses are must-haves, what need to be avoided, and what level of approval is required for exceptions. The framework will also define the best practices and benchmark clauses which can be leveraged for a given meta clause.

The robustness of the framework has grown with the assessment of over a few thousand contracts, and MSAs, sifting of few thousands of financial clauses and meta clauses, across categories and subcategories of rights and obligations.

With inputs from—R Vennimalai and S. Vasundhara
Contract Finance Management Team

Visualization and Gamification to Enhance Human Attention

Digital Kanban

The Kanban method, a practice invented in the manufacturing industry, has been adopted in the IT industry as an agile practice. This helps to visualize workflows at a glance, improve turnaround time by identifying impediments earlier, and improve productivity by limiting the work-in-progress nature of projects. The Kanban method uses the Kanban Board and Kanban Cards to display the prioritized work items waiting in queue (To Do), work items that are

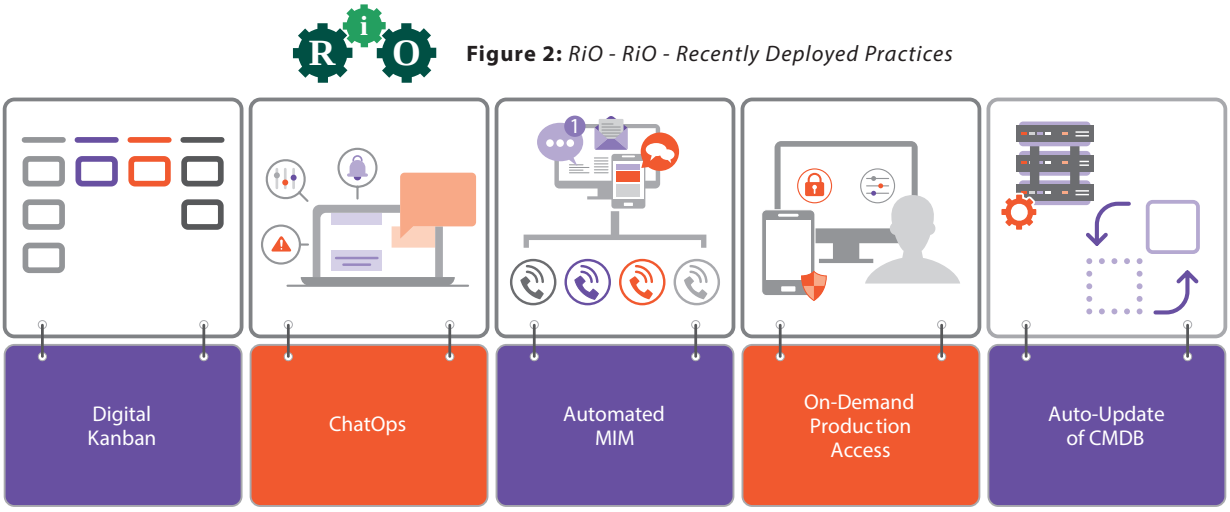


Figure 2: *RiO - RiO - Recently Deployed Practices*

in progress (Work in Progress), and the work items that are completed (Done). Kanban Cards are depicted in different color codes for specific types of work along with visual signs indicating priority. Kanban can also be achieved digitally as some of the IT service management (ITSM) tools provide this feature as an out-of-the-box product (Figure 2).

Many TCS production support engagements use digital Kanban as a best practice to improve productivity in resolving problem tickets and low-priority incidents. This has led to improved turnaround time and enhanced customer satisfaction. One of the support engagements, through the adoption of digital Kanban, was able to achieve a 30% increase in productivity.

Visual Demarcation

As part of the RiO best practices implementation drive, operations support teams have implemented visual demarcation to prevent manual errors. They have adopted the following methods to achieve visual demarcation of production and development environments:

- Using different backgrounds as well as foreground colors, fonts, and naming conventions
- Making the production server name visible in the command prompt
- Attaching prefix/suffix environments with user IDs

Automated Health Checks and Alert Monitoring

Automated Ready for Business (RFB) Checks

RFB checks are performed by the support team to ensure that the system is ready and available for business users. Such checks involve the execution of a standard set of steps to identify system health. However, these checks are time-consuming when performed manually. Many of the operations engagements in TCS have automated these standard checks. Automation helps to eliminate any

manual errors and repetitions in these RFB checks multiple times in a day.

These RFB checks are combined with alert and monitoring dashboards where they continuously display the system health information. Any issues reported in RFB checks are immediately acted on by the in-house operations team and fixed before they impact the business users. This helps improve user satisfaction.

Automated Monitoring and Auto Ticketing

Monitoring is a key aspect for the support team to detect issues and fix them before they impact business users. The various components that make up the entire IT ecosystem should be monitored. This includes application monitoring, log monitoring, batch job monitoring, infrastructure health monitoring, and networks, among others. Monitoring tools are configured to monitor production events and trigger alerts to the support team if an issue is detected. In operations engagements, command center teams monitor these alerts using centralized monitoring dashboards and proactively act on them.

For actionable critical alerts, the troubleshooting and recovery needs to be faster before it impacts business users. ChatOps is a mechanism that enables support teams to get notifications of actionable critical alerts in a chat tool for immediate attention. As part of the RiO Emerging practices implementation, operations engagements have implemented ChatOps. In an engagement with a large retail customer, TCS used ChatOps to achieve a reduction of up to 80% of the monitoring team. Clearly, ChatOps can be successfully used to free up valuable human resources from mundane, technical tasks so that they can focus on insight-driven, strategic work.

There are many engagements in TCS where monitoring tools interface with ITSM ticketing tools to automatically create an incident with appropriate severity. This incident is then resolved either automatically by the machine or assigned to the support group for manual resolution. This effectively removes the need for a monitoring team. It also accelerates the turnaround time and avoids misses of critical alerts which can occur through manual monitoring.

Cognitive Automation and Self-healing

Support teams spend most of their time in firefighting activities such as addressing critical issues, or mundane, repetitive activities such as health checks, resolving standard service requests, and repeat alerts. With advancements in technology and cognitive automation tools like ignio™, it is possible to automate the resolution of repetitive issues and service requests, and self-heal alerts as soon as they are detected. TCS is leveraging its Machine First Delivery Model (MFDM™) to transform manual, error-ridden operations to error-free, efficient operations.

The prerequisite for automation is the creation of knowledge articles (KAs) with well-defined steps on action or resolution for any repeat activity. Hence, knowledge management plays a crucial role for automation. Operations engagements in TCS create a knowledge base and a known error database through KAs. The appropriate KA is pre-tagged automatically at the time of ticket creation for the support executive to immediately leverage for resolution. One of TCS' largest engagements resolved 95% of tickets which were tagged with KAs. If the KA is not pre-tagged, a support associate tags the relevant KA manually when the ticket is resolved. KAs that are highly utilized naturally become candidates for elimination or automation.

ignio™ helps eliminate noise, undertake correlation, prioritize alerts based on business impact, and resolve alerts autonomously. For actionable alerts that require human intervention, it creates a ticket in the ITSM tool or creates collaboration channels with the support teams.

Cognitive automation applies artificial intelligence (AI) and machine learning (ML) techniques to analyze relevant information by connecting to the environment. Consequently, it is able to predict and prevent issues, thereby enabling a more stable environment.

TCS currently has multiple engagements where, in collaboration with its customers, cognitive automation solutions have been adopted to auto-resolve, improve mean time to resolve (MTTR), avoid human errors, be first-time-right, and improve user satisfaction. One of TCS' largest banking customer implemented MFDM™, leveraging orchestrator-driven automation that reduced their command center team size to zero.

K Subramanian

K. Subramanian (Subra) is the Global Head of Delivery Excellence at TCS. In his role, he owns Enterprise Quality, Customer Advocacy, Processes & Methodologies, Enterprise Agile, Knowledge Management, Estimation and Continuous Performance Improvement. Subra joined TCS in 1987. Over the past 32 years, he has played diverse roles in Software Delivery, Application Operations, Sales & Marketing, Quality, Program Management, Consulting, and Process Excellence. His experience includes several customer engagements with global IT operations. He has been a crucial member of the Rigor in Operations initiative of TCS.

Balaji A.K.

Balaji AK is Lead, Delivery Governance, TCS. A Production Support SME, Balaji has over two decades of expertise in this area, having worked across various roles such as Support Executive, Support Manager, Head of Application Operations, and Global Head for Application & Infrastructure operations including Service Desk.

He also leads innovation for Rigor in Operations Workbench. In his current role, Balaji is responsible for Application Operations, Delivery Governance, and driving Rigor in Operations for Production Support engagements across the enterprise.

M G Duraibabu

Senior Consultant

Co-Innovation with Academia

New models of collaboration

Raju Goteti

An early adopter of collaborative and open innovation, TCS has decades of learning in creating research partnerships with a variety of institutions. TCS' Co-Innovation Network (COIN™) links to multiple entities. One of COIN's arms expands TCS' research and innovation capabilities by linking to the best academic institutions and researchers from around the world. Through this program, TCS has built expertise in technical aspects. Processes such as transparency in handling IP, speeding to business, and providing the best atmosphere for fostering creative work thus creating win–win outcomes related to resource planning, funding and joint work. Academic COIN has four levels of collaboration with institutes, moving from transactional connects to long term multi -project MoUs. Academic COIN has experimented different models with academia—sponsored research, joint labs, research advisory, consultancy, research sabbatical programs, and internships. In the last two years, COIN has forged some novel alliances, which reflect its maturity in orchestrating complex networks. Here, we have described examples of some of these new approaches.

Multidisciplinary Work: TCS Research has broad-based agreements in a number of areas with esteemed universities. Going beyond the narrow focus of one-on-one partnerships, these programs facilitate rich interactions and output.

Consortia and Clusters: TCS is a part of many government-led industry consortia, where a number of research, industry, and government agencies are involved in solving specific real-world problems. A good example of this is IIT Kharagpur's DHI Manufacturing Consortium that involves Tata Sons, Tata Motors, Tata Steel, Bharat Heavy Electricals Ltd., and Heavy Engineering Corporation, along with TCS. The consortium focuses on problems in smart machines manufacture in the capital goods sector. Another example is the project on stroke rehabilitation under the Uchhatar Avishkar Yojana (UAY) of the Ministry of Human Resources, Govt. of India. It has TCS collaborating with a cluster of institutes such as IIT Madras, IIIT Hyderabad, and NIMHANS in Bangalore.

TCS Research has broad-based agreements in a number of areas with esteemed universities.

TCS Pace Ports bring an academic partner, a COIN accelerator and TCS Research to work together in a physical space to enable agility in Innovation

Pace Ports: Pace Ports are physical spaces strategically chosen close to academic research campuses. Pace Ports create a unique synergy among the co-located academic research centre, a COIN Accelerator, TCS Customers and TCS Research. The idea is to bring all parties to work together in a physical space which is equipped with design thinking areas, agile work spaces and innovation show cases. TCS Pace Port New York is housed in the Tata Innovation Center at Cornell Tech and has worked on some problems relating to the retail industry.

As apparent from essays in this volume and Volume I, we have had academic research collaborations across technologies and industry segments. Some of these have involved engineering innovation and prototyping too. These collaborations have vastly enriched TCS' efforts in solving various industry challenges (Amazon Robotic Challenge, Vol 1, page 36), addressing customer problems (Digital Twins, Vol 1 page 97), responding to societal challenges (The Second Genome page 161), and executing long-term research (Preparing for Quantum Computing page: 58 The Mutant Gene, Vol I page 171).

With inputs from Sujit Guha.

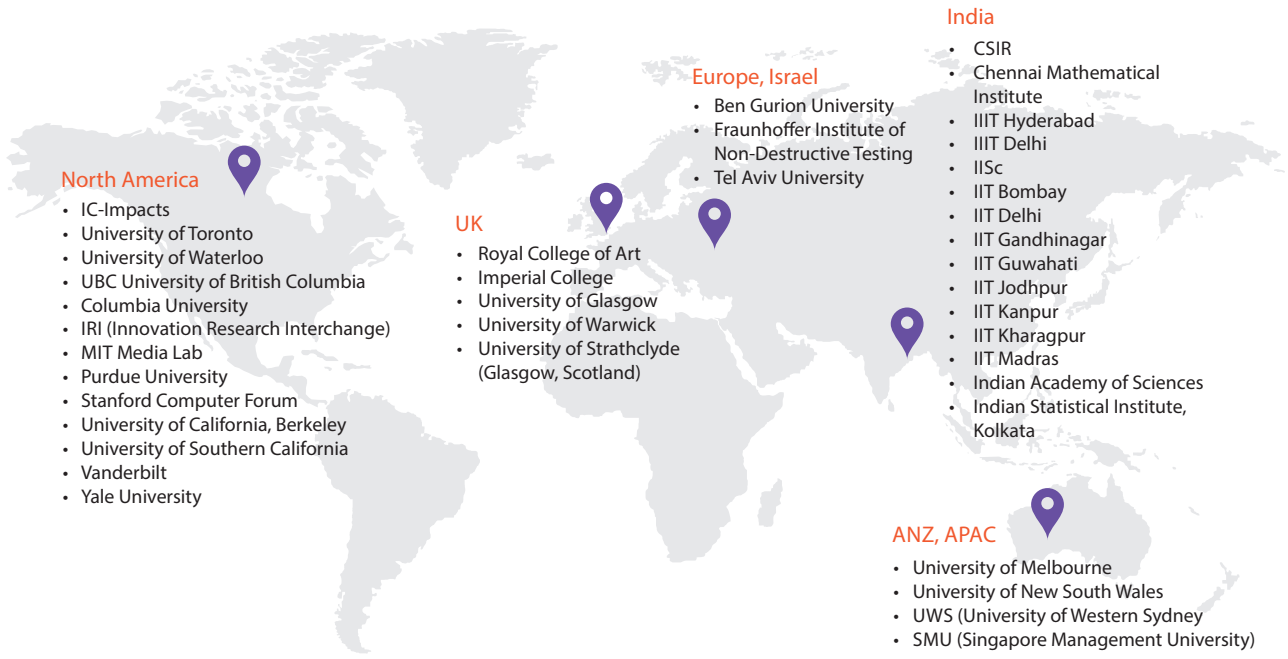


Figure 1: *Global Footprint of TCS' Academic CON Program*

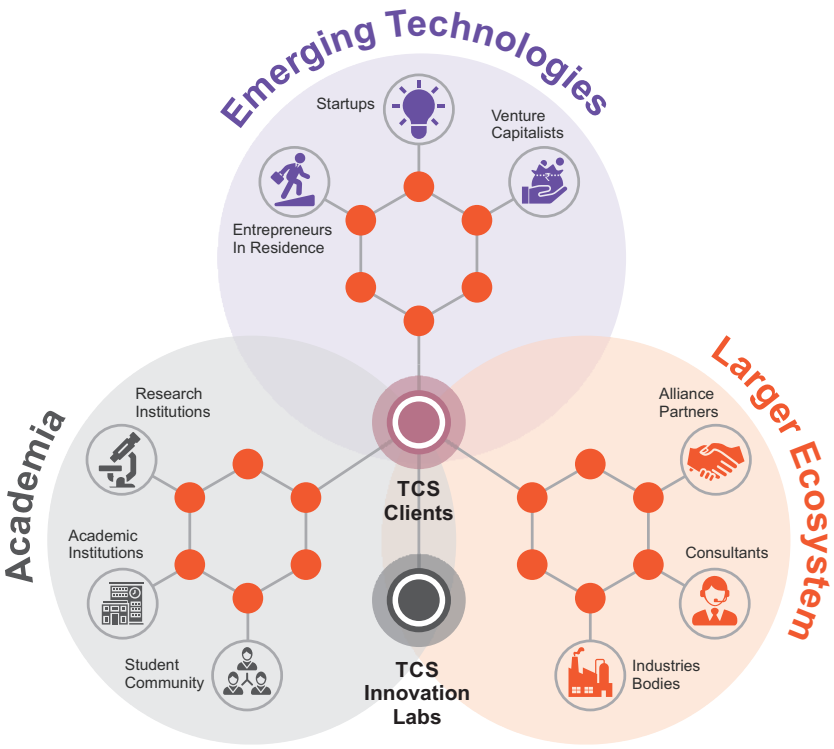


Figure 2: *TCS COIN™*

The Academic COIN Program pays special attention to research students by encouraging lab visits, sabbaticals, poster presentations and internships

Scholars Scorecard

The Academic COIN Program pays special attention to research students, by encouraging lab visits, sabbaticals, poster presentations and internships. Here are some initiatives:

TCS Research Scholarship Program: This program supports Ph.D. Researchers in India with a stipend and other benefits. In its 10th year, it has supported 300+ Ph.D. scholars, across 40+ institutes across India.

Student Mentorship – North America: In the last two years, more than 100 students in North America have been involved in our projects where they have been mentored by TCS on real-world business challenges.

TCS Research Cafes: TCS researchers have informal meet-ups with students from premier institutes to inspire a career in research.

With inputs from – Flt Lt Venkatesh Madihalli

Raju Goteti

Raju Goteti is Vice President and Head of TCS' Co-Innovation Network. He is responsible for creating capacity and driving programs along with academic institutions , while leveraging emerging technology startups to bring the newest innovations to clients globally.

Raju is responsible for managing TCS' partnerships with leading emerging technology companies. He has co-founded TCS' COIN Accelerator to engage with startups in a client's context and develop new solutions.

Raju has postgraduate degrees in Social Innovation from the University of Waterloo, Canada and Ocean Engineering from Memorial University, Canada. He has completed his Bachelors in Mechanical Engineering from Indian Institute of Technology, Delhi.

Taking Research to Business

TCS' New Products and Services Development (NPSD) framework

Chakravarthi Sathyanarayana

TCS Research aims to solve real-world problems and create value for TCS' business and to society. TCS now places a greater emphasis on patents, platforms, and products. This has created a strategic focus on Research and Innovation (R&I). To build rigor into IP-based assets, TCS R&I created the New Products and Services Development (NPSD) Framework. NPSD functions as a governance framework to assist and review the conceptualization, assetization, and commercialization phases of all innovative products in TCS.

A Matter of Scale

TCS has a dedicated R&I unit and has been investing in it consistently for four decades. Since 2015, it has expanded the R&I footprint in each of its industry units—it crowdsources ideas; it encourages entrepreneurs in residence; it co-innovates with its ecosystem technology partners and customers; it incubates research projects and high-impact business ideas. This scaling of innovation across the organization ensures a steady stream of new ideas and experiments. Candidate ideas with strong business potential have to pass through the NPSD gateways (Figure 1) to attain market maturity. The NPSD framework brings research and business stakeholders together in a stronger and more purposeful journey towards new business offerings.

The Three Lenses

As research should lead to a business prototype, the NPSD framework asks the right questions at the right time. It examines the offering using multiple lenses, chief of which have been succinctly termed "Desirability, Feasibility, and Viability" by Ideo. Among other things, the framework forces researchers to answer the following questions about their work:

- Who needs this?
- Is it possible to do this effectively?
- Will this create value?

The framework's deep dive on each of these aspects compels projects to examine technical robustness and also provide proof

NPSD functions as a governance framework to assist and review the conceptualization, assetization, and commercialization phases of all innovative products in TCS

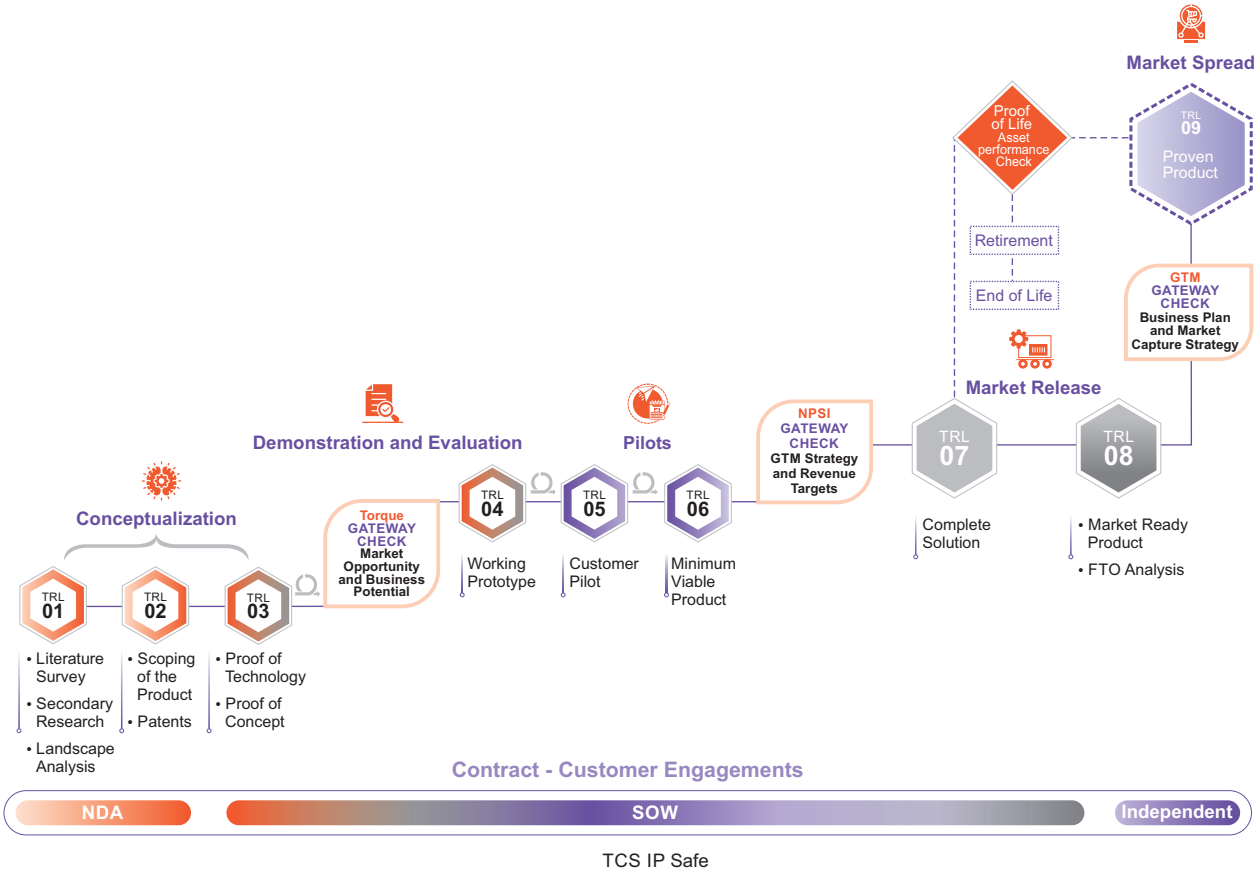


Figure 1: *The NPSD framework*

The NPSD framework asks the right questions at the right time, examining the offering using multiple lenses

points, study failure hypotheses, and prepare better for the market.

A Balancing Act

TCS' NPSD framework attempts a delicate balancing act to provide the required freedom that researchers need to explore and to assure the business of a robust and novel solution. It assesses maturity along a "Technology Readiness Level" (TRL) scale that graduates from 1 to 9. (Figure 1). Stage 1 (TRL 1–3) provides plenty of freedom for research to explore new technologies. Stage 2 (TRL 4–6) examines the desirability of the proposed offering and the technical feasibility in real-world scenarios. By Stage 3 (TRL 7–9), the offering must have gained in technology robustness and shaped itself for business viability.

A Business Thinking Tool

The NPSD framework prepares research for the long haul. If the potential research-based offering does not seem capable of technical efficacy or have market viability, it is sent back to the drawing board. TCS invests in blue sky research projects. But there are several

projects that work on real-world problems too. The NPSD framework enables these projects to understand business expectations. It also invites business to take an interest in research as it progresses. Thus, the NPSD framework nudges the business and technology teams to work closely as the new offering moves down the pipeline. It leverages synergies in the organization, strengthening the collective understanding of the business potential of new technologies.

As a well-defined process, the NPSD framework brings clarity not only to research and business teams but also to aspects such as engineering and functions, such as legal and marketing. Most importantly, NPSD enables leadership to look at the funnel of options and prioritize accordingly.

NPSD leverages synergies in the organization, strengthening the collective understanding of the business potential of new technologies

Chakravarthi Sathyanarayana

Sathya is a Principal Consultant with TCS Research and Innovation. He works in the area of Innovation Management. He has a Master's degree in Computer Science from the Indian Institute of Technology, Madras.

Incubation: Scaling Innovation for Impact

Anita Nanadikar

Incubation is responsible for creating impactful new business by fostering ideas that are at the intersection of evolving customer needs, technology discontinuities, and emerging market opportunities

New ideas typically have a high infant mortality rate. Anita Nanadikar has been charged with the responsibility of heading an incubation group for scaling up new business ideas (applying TCS R&D output to business problems) through their infancy until they form new business streams.

We bring you excerpts from an interview with Anita Nanadikar, who has led the group for 11 years.

On the Initial Mandate

Incubation is responsible for creating impactful new business by fostering select ideas that are at the intersection of evolving customer needs, technology discontinuities, and emerging market opportunities. Typically, these are ideas applicable to multiple industries that show promise of providing sustained value to our customers and thereby can be built into a new business.

Examples:

We created a completely new market with digital solutions and services for a CMO at a time when IT companies were still targeting CIOs.

We built an IoT business for TCS drawing from our expertise and our IP, i.e., our IoT platform, TCUP, that came from TCS Research.

We also mastered emerging technologies such as drones and AI vision and used this skill to build business-ready solutions.

With these, TCS could grow and win business in new areas such as digital marketing, IoT, drones, AI, cognitive, and Next Gen operations and create a substantial financial impact as well.

On Identifying Promising Ideas

Through constant customer interactions, extended co-innovation connections, and technology landscape scans, TCS Incubation is sensitized to white spaces in the market and the capabilities of technology. Then, the most promising ideas are chosen from different sources: formal research, incremental innovations in projects, ideathons, process improvement initiatives, and proactive ideas from associates.

Fact File

Patents: 53 granted

Awards: 17

With the rich innovation culture spread across the organization, at any given time, more than 3000 ideas are evaluated by us to select the ones that will lead to creating a high-impact new business or contribute to an existing one. We leverage our contextual knowledge (and known resource capability in advanced technologies) to identify the right ideas. At times, the market is not ready for the idea, so the idea is shaped to address the existing market. Often, the shaping is so extensive that the final product is very different from the idea that it started with (View example—Symbiosis: Drones and Machine Vision).

On the Magic Potion to Create an Entrepreneur for the Market

TCS Incubation is like an incubator + startup builder. The first step is to select the right disruptive ideas with potential for high impact (after concretizing fuzzy ideas and shaping them to address market needs). This is followed by mentoring entrepreneurs to build cutting-edge solutions by providing emerging technology talent and expertise, access to market, research, IP, and legal support. To jump start the entrepreneur, we provide a core shared team and infrastructure to support activities linked to business development and presales. We also provide an ecosystem of co-innovation and alliances, any market research needed, and support for marketing, rapid development and engineering using emerging tech, as well as productizing and IP expertise.

With our broad knowledge and understanding of new technologies, markets, and industries, we also help to identify adjacencies and additional value streams for the idea. Having an entrepreneurial mindset, applied research capability, and iterative shaping of ideas based on the market feedback also help in acceleration toward creating exponential value and impact. To ensure consistency and repeatability, a proven and systematic process has also been implemented to handhold the entrepreneur through this journey. We have taken a conscious decision to focus on the intellectual property (IP)-driven approach to create value.

On Ideas that did not Create Value

While statistics vary, it is common knowledge that the success rate

TCS Incubation is like an incubator + startup builder

An Entrepreneurial mindset, applied research capability, and iterative shaping of ideas based on market feedback also help in acceleration toward creating exponential value and impact

of startups is less than two digits. At TCS, we have managed to flip the success and failure rates of startups, i.e., our failure rate is in one digit, so far. We also work much faster than the average startups. While most startups take up to 7–10 years to mature, many of our programs have matured in half that time.

Even in the case of a couple of projects where we may have not scaled out to business lines, we have reused the new skills learned and pivoted to adjacent areas. We believe that we have generated value even in those cases.

Meeting Resource Challenges

Since TCS is a large and diverse organization, the Incubation team has access to research, engineering, and industry expertise. However, since technologies are new, the teams require a lot of learning. We spend a lot of time to get things off the ground.

Each Incubation program based on a selected idea is backed by a well-thought-out business plan that covers a design and development roadmap, competency building, target markets, and potential customers. For technology competencies, we adopt a multipronged approach that includes intensive, often on-the-job training on emerging technologies, leveraging of TCS research, and access to our partner ecosystem. Additionally, there is an Incubation lab where teams are constantly experimenting with new technologies and evaluating their fitment to each program. Shared teams provide sales, marketing, finance, operations, and legal competencies required by each program.

For example, the entrepreneur and the team formed for the drones program were new to the technology and went through a rapid learning cycle by co-locating with the research team. This was further complemented by physical experiments in the lab and access to partners and customers for field tests.

On Learning and Building Talent

Most entrepreneurs come with ideas and maybe with one or two skills. Incubation helps develop a well-rounded startup team with a combination of skills relating to creativity, problem solving, technology and engineering, business development, marketing, customer engagement, and working with, and in, an ecosystem.

On 'Rapid Labs'

The Rapid Lab, a recent initiative, is an excellent example of honing young, often fresh-out-of-school talent. It is gaining traction with customers' C-level executives, innovation leadership teams, and business and IT stakeholders.

The Rapid Lab is a physical space equipped with cutting-edge technology. Bright young TCSers' fresh thinking is used to rapidly prototype innovative solutions for customers and for addressing social challenges.

Since this lab is air gapped from the enterprise network, with an open infrastructure and ecosystem support, they can try multiple options very quickly and are free to learn from failures. They have access to senior mentors and an extended innovation ecosystem of local innovation labs, government bodies, academia, startups, and partners.

Some of the challenges successfully solved by the lab include the following:

- A lounge concierge robot designed and built ground-up, including the hardware fabrication. This has been deployed and is servicing customers 24/7 at a leading airline's airport lounge
- Electronic noses deployed in TCS washrooms powered by AI and sensors to detect odors
- An AR/VR-based product

Each Incubation program based on a selected idea is backed by a well-thought-out business plan that covers a design and development roadmap, competency building, target markets, and potential customers

configurator for a high-quality audio device company. It has allowed them to save significant time and effort

Individual solutions developed by the lab have also come together like pieces of a jigsaw puzzle to create new solutions. For example, AI solutions developed by the lab, such as AI-powered call denoising, language-based call routing, emotion identification and sentiment analyses, and AI-generated cartoon strips for agent training, have helped in creating a new solution set for AI-powered contact centers or back-office value chains.

On Transition to Business

Having nurtured and grown a program/solution, it is imperative to identify the ideal time for commercializing a venture. Incubation follows TCS' New Products and Services Development (NPSD) process, where it obtains business unit buy-in early on and hands over the program once all NPSD requirements are met.

Today, technologies are so complex that when an idea is taken up, many of the established business units may not have the skills for it. Over a couple of years, business interest picks up but Incubation does the handholding in the interim.

We ensure that outputs are well-rounded products and solutions and engineer a smooth transition to the identified business units.

On the Future

The future seems exciting for us. While we have been working as a central team incubating and passing on mature solutions to the business units, this is set to change. As requested by the CEO and CTO, we are working with business groups on incubating their strategic initiatives. This will help TCS leverage the central team's incubation expertise—so that our industry units deliver non-linear value to their customers, thus ensuring that they stay competitive and ahead of the curve.

With inputs from Amit Chemburkar and Nitin Hanjankar

Symbiosis: Drones and AI-Driven Machine Vision Solutions

A few years ago, when drones were still in the labs, TCS Incubation understood that drones offered a unique technology driven opportunity to create a new suite of solutions addressing the needs of various field operations teams across industry segments. The program primarily targeted challenges pertaining to human fatigue, personnel safety, and manual inspection of widely distributed assets in the field. The B2B market for drone solutions was not well-established; regulatory policies restricted wide usage of drones. The opportunity, and challenge, was to create and shape a new market and pick the most promising ideas to ensure future customer needs, addressing a wider set of stakeholders beyond the Enterprise CIO – primarily the COOs and Heads of Digital Innovation.

The program was formed by identifying an Entrepreneur in Residence (EiR) with

Figure1: *AI Driven Drone and Computer Vision Solutions*

a team new to drones and associated technologies. Incubation had to provide a rapid learning cycle complemented by a physical space to conduct experiments in controlled environments/labs; access to partners; and identification of anchor customers for real-world field operating environments to perform the tests. "Co-location with TCS' Machine Vision Research team was a big plus," says Mahesh Rangarajan, the EiR for the Drones Incubation Program. "The machine vision research team enabled us to appreciate and acquire the necessary technical skills, which also helped us to better qualify the ideas and feasibility of the business use case. The presales teams brought us a market perspective, bringing in new use cases. It also helped us articulate our differentiators to customers clearly."

A near-shore lab was set up at TCS Cincinnati, USA in April 2017 where our drones and AI machine vision expertise, and solutions were showcased live to visiting customers. It helped us assess the market opportunity firsthand, based on the continuous customer footfalls and the associated feedback.

An indoor environment-based warehouse inventory counting solution was selected as the first MVP with a leading logistics provider as an anchor customer. This was a quick win, as it was not impeded by regulatory compliance. The solution was further extended to retail, CPG, manufacturing, and other industries.

Many other requests and solutions followed: Inspection of large assets such as forests, bridges, rail tracks, power and water utility infrastructure, sewer pipelines, and mine conveyor systems. Flying drones in each environment had its set of operational challenges - image acquisition protocols were different, the image volumes were very high, and data (image and video) management was a clear big data problem. And, the problem to be solved in each case needed slightly different analyses. The learning for the team was huge and a common underlying high-scale core data processing and engineering platform was built to accelerate the individual solution development cycle.

"Incubation helped us with timely course correction and also to pivot – primarily driven by the culture of delivery and validation in continuous 30-60-90 day cycles. The drones team thereby acquired a large machine vision footprint," says Mahesh. The TCS Drones and Machine Vision Incubation Program has built IP and solutions for several industries against high-impact use cases covering transportation, utilities, energy and resources, manufacturing and government enterprises. TCS Incubation's repeatable process has extended the initial idea to 16 drone solutions. A Core Platform is in place and the team is now working jointly with specific business units on multi-million dollar business plans.

With inputs from—Mahesh Rangarajan

Moving Ahead with the TCS Cognitive Transformation Platform

When the Incubation team came up with the idea of creating a cognitive platform, it did an extensive survey. The Incubation marketing and presales teams spoke to customers to understand their pain points. The teams also spoke to analysts to understand the market for cognitive tools. Meanwhile, it became evident to us that multiple initiatives were already in flight, both internally in TCS and within the customers' enterprises, to harness the cognitive and AI technologies. However, customers were facing challenges in selecting the right tools and staying current in the fast-evolving AI space.

"The team decided to build a platform with an abstraction layer that would enable customers to build cognitive solutions with flexibility to select and easily integrate cognitive capabilities for the much-needed inter-operability to migrate to new tools in the future as needed. This approach would enable us to bring in different cognitive capabilities, developed both by our research teams and our partners, seamlessly to help simplify and accelerate the development of cognitive applications", says Pranav Shah, who heads the Cognitive Transformation Platform (CTP) initiative.

Being part of the Incubation team, the CTP project had access to some shared resources, such as Ninja developers and UI designers. It could also use the Ops Framework, which was being incubated.

Today, CTP has evolved into a full-fledged solution that allows enterprises to adopt AI without investing in highly-skilled AI resources or long development cycles. It provides:

- Zero coding business bots that allow business users to undertake AI-powered business tasks on their own without having to wait or rely on AI developers
- Building blocks for cognitive information retrieval with an ability to ingest knowledge, infer, correlate, and dynamically learn from past experiences, data, and human input
- The ability to ingest and process multi-format knowledge and information like humans do from images, audio, documents, text, and structured data
- Several prebuilt adapters for integrating seamlessly with enterprise/third-party systems
- APIs to abstract technology and integration complexity and eliminate the need for large teams of niche-skilled AI developers

CTP can be used to simplify travel claims. Presently, TCS associates applying for a travel claim submit a set of documents such as hotel bills, communication bills, and cab bills, along with their claim form. In future, CTP can be expected to extract information from documents in various formats, and "file" the claim request. Further, it can make the claims process efficient by identifying outliers to prevent potential misreporting or fraud. For example, if a cab bill from Point A to B typically amounts to Rs 500, a claim made for Rs 1,500 can be flagged as an anomaly.

TCS Incubation helped CTP evolve quickly into a robust platform, which can now power TCS and its customers to democratize the development and transformation of their cognitive business applications rapidly, without extensively investing in an AI workforce.

With inputs from—Pranav Shah

NextGen Ops Framework gets Feature Rich

About two years ago, a service desk team, supporting 30+ customers, reached out to the presales team within Incubation with its problem. There was a mandate to improve the operational efficiency through optimization of the service desk function. One of the areas identified was customer email handling. The service desk received around 3000 emails a month, 1000 of which were actionable. Reading each mail, sorting it into different buckets along with their attachments for the next steps was a mundane and time-intensive task. The Ops Framework team used a natural language engine and automated the process: the engine mined the mails, classified the content, detached the attachments with the content and prepared it for the next steps, reducing manual effort by 25-30 percent.

"The Ops Framework grows richer with every engagement," says Jayashree Arunkumar, Technical Architect of the Ops Framework. After the natural language engine, came ChatOps, AI, and various new technology flavors.

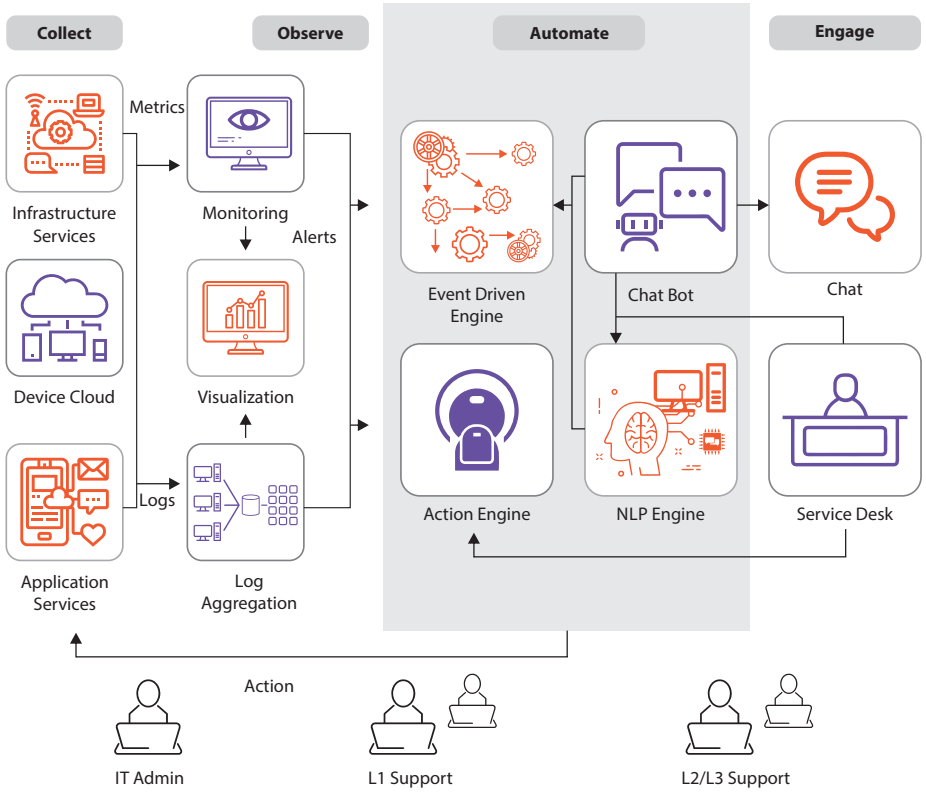


Figure2: *TCS Operations Framework Architecture*

TCS Operations Framework is a TCS IP-based framework, which uses intelligent and integrated automation across IoT-IT-Biz-Dev-Sec-Test-Ops to improve service management processes and seamlessly integrate with existing enterprise technology stacks. This is done by continuously deriving insights about the current operating state using information from various sources, such as metrics and logs, combining them with contextual information and developing self-learning AI models that drive automated actions.

"Rapid prototyping within the incubation team helped fast-track point solutions that validated feasibility of new capabilities," says Ravindran Subbiah, Program Owner of the Ops Framework. "The future holds tremendous promise: concepts of DevOps, Agile, and Lean will enable machines and algorithms to handle operations with humans involved only for exceptions. This will be a paradigm shift in organizational models."

The NextGen Ops will leverage digital twins for IT systems and devices, to ensure that service strategy and design are well-informed of behaviors that can be planned into the service upfront. It will help services provide observability and learn from dependent services, with which they coexist. NextGen Ops will simplify capacity management processes through auto-provisioning and auto-scaling of resources. It will also monitor runtime usage, intelligently identify possible issues and preempt them through self-heal solutions.

TCS Incubation has created the foundation for the NextGen Ops business with automation, intelligence and observability built into the Ops Framework. Using the TCS Ops Framework IP, TCS has been able to tap and convert new business opportunities that have emerged, such as Drones Ops, AI Ops, Quality Ops, Business Ops and Security Ops..

With inputs from—Jayashree Arunkumar and Ravindran Subbiah

Anita Nanadikar

Anita Nanadikar is the VP and Global Head of Incubation at TCS. As the Head of TCS Research & Innovation's Incubation function, Anita is responsible for identifying and building emerging businesses using new technologies and business models. TCS Incubation also manages an Incubation lab where customers, TCS innovators, experts, and partners come together to rapidly convert innovative ideas into disruptive solutions.

The current focus for TCS Incubation includes IoT-IT-Biz-Dev-Sec-Test-Ops, Drones, Cognitive Transformation, AI and ML, Blockchain, and Immersive Experiences.

Anita started her career with TCS and has over 35 years of industry experience.

Digital Impact Square

A pre-incubator for social innovation

Anil Sharma

IN BRIEF

TCS enables scaling of innovation within the enterprise and in society. In addition to incubating high-impact new businesses, setting up innovation labs for select clients, and establishing governance processes for realizing innovation, TCS has been spurring social entrepreneurship.

This essay is about the Digital Impact Square (DISQ) initiative, begun in 2016 and now supported by the TCS Foundation. DISQ operates from the Indian city of Nashik, Maharashtra. It aspires to create change-makers, powerful agents who in turn will engender the progress essential for the betterment of society.

DISQ's mission is to encourage social innovation using digital technologies to address the needs of citizens. By encouraging a culture of creativity and entrepreneurship, DISQ aspires to create change-makers, powerful agents who in turn will engender the progress essential for the betterment of society.

DISQ was founded in Nashik within a robust ecosystem that was created through the active participation of academia, the city's local administration, and the community at large. Close collaboration amongst the participants through a series of innovation camps and build-a-thons led to the birth of DISQ. With the local government playing a key role, DISQ produced remarkable solutions to as many as 15 crucial problems faced by Nashik, in its first year.

DISQ's modus operandi is open in nature. Mentors from TCS bring together university students and young entrepreneurs from all over the country to form multidisciplinary teams so as to bounce ideas off each other. Structured as a platform, DISQ offers innovators the opportunity to launch their innovations in a mass market, with successful teams retaining their intellectual property (IP). With a focus on social innovation, DISQ develops solutions that are tuned to the needs of the user.

DISQ innovations are categorized into seven application areas:

- Health and hygiene
- Education and skills
- Financial and personal security

DISQ offers innovators the opportunity to launch their innovations in a mass market, with successful teams retaining their intellectual property (IP)

- Energy, water, and environment
- Food and agriculture
- Housing and transportation
- Citizen empowerment and transparency

How DISQ Works

DISQ's incubation process has four phases – Spot, Probe, Grow, and Scale – each with its own stringent elimination criteria. In the Spot phase, DISQ selects teams with proven capability in developing innovative engineering-based solutions that have the potential to raise the quality of life of a significant mass market. Candidates tender detailed applications at the DISQ website (www.digitalimpactsquare.com), pitching their ideas to a panel of domain and technology experts and stakeholders from academia, industry, government, and civil society. The panel sifts through the applications to gauge the potential of ideas, team member profiles, project experience, and professionalism. With regard to students, the passion to solve complex societal problems and an entrepreneurial bent of mind are considered in the evaluation process. Apart from final year projects, technical and core skills are also assessed.

The teams selected comprise 19- to 27-year-old final-year students of Master's or Bachelor's degrees, early-stage start-ups, and budding entrepreneurs. Students hail from engineering, management, mathematics, statistics, liberal arts, commerce, medicine, design, and architecture streams, are multidisciplinary, and are expected to work collaboratively.

Next comes the Probe phase. During this phase, each idea is developed into a working prototype at DISQ in 6 months, with teams working full-time for 6-12 months, designing, developing, and testing the prototype at the innovation center. Teams are trained to take a human centric design approach to enable designing of solutions for successful adoption. The teams have access to field-tested research curated by academia and businesses as well as to mentors and experts in the relevant fields who help test their solution

locally. From here on, all teams in each phase are guided by TCS' innovation coaches, design leads, ecosystem mentors, experts, and enablers.

In the Grow phase, the prototype is production-tested. Once it demonstrates business potential, the Scale phase is reached, wherein TCS connects the teams to strategic partners and investors, to help their products attain marketability and real-world viability.

Grown and Scaled

So far, over 300 innovators from 70 cities have worked in DISQ labs. Of them, more than 45 have formed start-ups, with six teams moving to Scale and four to Grow, covering a user/customer base of over 8,00,000 between them.

Vesatogo: One of DISQ's showcase solutions is Vesatogo, an app-based system for marginal farmers to transport their harvests at the highest profits to markets of their choice. Vesatogo connects all upstream echelons in the agriculture supply chain – farmer, transporter, wholesaler, and the processing unit. Procurement agencies – farmer producer organizations, traders, and processing units – log their demand on the app. Their requirement is visible to all farmers and triggers producer-procurer bidding. Based on the information, farmers close deals with the next echelon. Additionally, Vesatogo aggregates transport vehicles and tracks all shipping particulars. This allows farmers who use the app to select the best transporter to ship their produce to the wholesaler or processing unit.

Third I: This solution has helped the 9000-strong Nashik police department anticipate and prevent crimes faster. It also improved their responsiveness toward the general public through informed and intelligent decision-making. Third I is a real-time,

TCS connects the teams to strategic partners and investors, to help their products attain marketability and real-world viability

centralized data assessment and visualization solution. It's a Scale-level solution.

Kibo xs: Trestle Lab's Kibo (Japanese for "hope") is a device that reads out printed, handwritten, and digital text for people with visual impairment and learning and intellectual disabilities. Kibo requires the user to merely place the desired text on a pad approximately a foot from a table lamp-like device; and the software instantly reads aloud the text at normal speed. The current version of Kibo is enabled for English and nine Indian languages, including Hindi. Now, Kibo users can dispense with bulky Braille texts which are hard to come by. (Bulky Braille text may also be the reason behind a 50 percent dropout rate among blind people in pursuing college education).

What makes DISQ tick? It could be the inventor's thrill or perhaps the maker's jubilation. It could be the team member's joy, a user's relief, or even the facilitator's satisfaction. We like to think it's all of the above.

Further reading

1. https://brailleworks.com/braille-literacy-statistics/
2. See: https://nashik.com/kumbhamela-2015-16/
3. See: https://nashik.com/kumbhamela-2015-16/

GG Talk – GappaGoshti®

2015 was a watershed year. Over three days that year, during August-September, 30 million people visited Nashik, a city which then had a population of just over 2 million. The reason? The Kumbh Mela, a major Hindu festival that comes once every 12 years to the city, was held then. During Kumbh, pilgrims gather en masse for a holy dip in the River Godavari, around which Nashik has grown.

Crowd management during the Kumbh Mela had historically been a nightmare for the Nashik municipal administration. Every Kumbh Mela before 2015 had seen thousands of pilgrims going missing, with a few dying.

But from 2015 onwards, history would not repeat itself.

Eighteen months before the 2015 Kumbh festival, the Boston, MA-based MIT Media Lab, Nashik's civic administration, and some corporate business houses organized "Kumbhathon,"

a hackathon in which 800 students, entrepreneurs, and start-ups collaborated to develop 12 social media innovations. TCS was the technology mentor for the initiative.

Thanks to GGTalk – GappaGoshti®, a home grown, TCS patented social collaboration platform,roving volunteers on the ground posted real-time photos, videos, and audio messages, including updates, alerting a central control room to emergencies and hazards related to crowds, traffic, missing persons, sanitation and hygiene, and parking. All ground-up communication was hashtagged for efficient categorization and action. Using location heat maps, trend analyses, issue recurrences, and impact assessments, the control room, in turn, collated the live feeds and furnished situation reports to the top-level civic authorities – the State Kumbh Administration, the Nashik Municipal Commissioner, the Divisional Commissioner, the District Collector, and the Police Commissioner – who then mustered and deployed rapid-action task forces to defuse the crises.

The result: zero deaths, no stampedes, zero epidemics, and no missing persons at end of day.

Kumbhathon's overwhelming success proved that technology-led social innovations developed by young talent is a demonstration of how India's heart beats for the commoner..

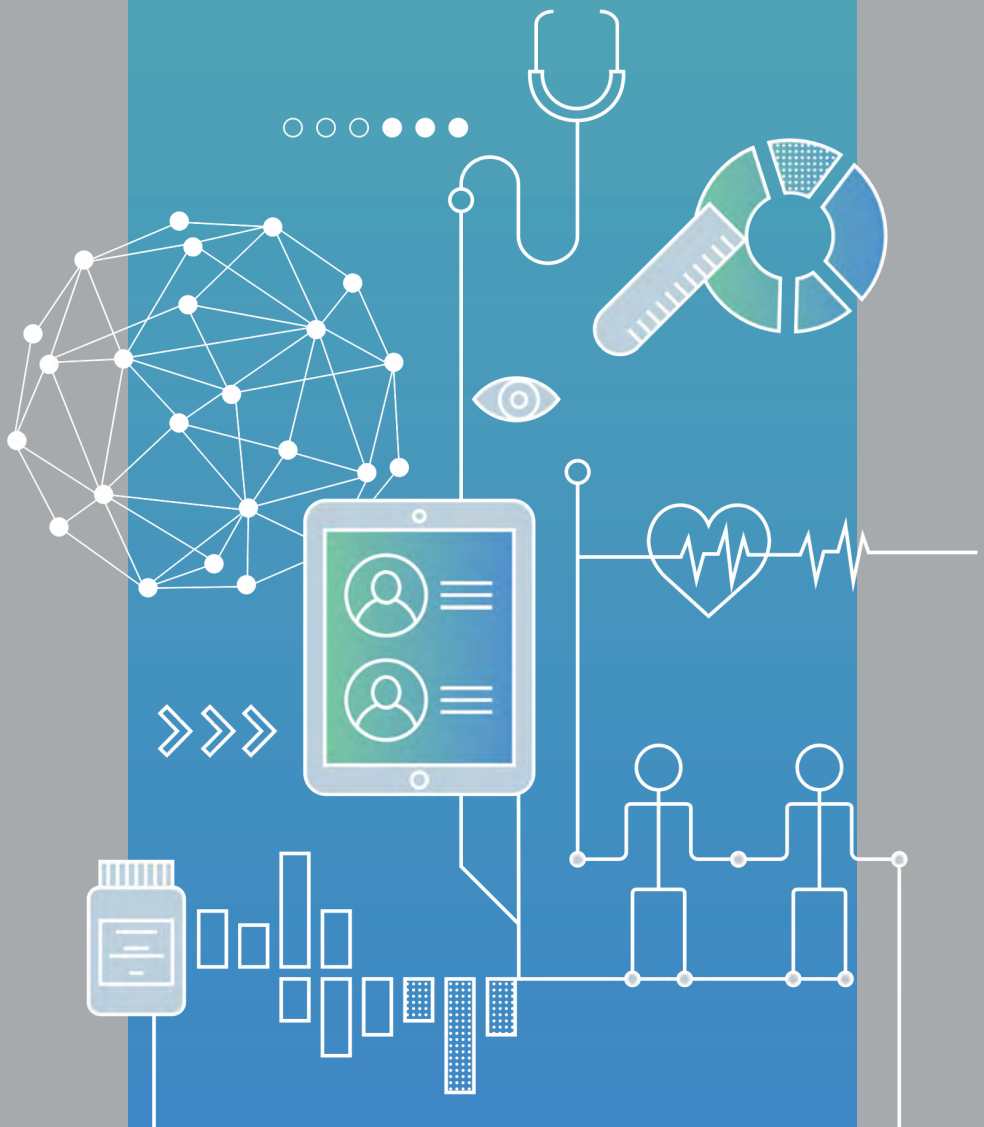
With inputs from—Sanjay Kimbahune, - Pankaj Doke

Anil Sharma

Anil Currently heads TCS Foundation Initiative - Digital Impact Square (DISQ), a social incubator that looks at delivering impact at scale in India by creating viable social enterprises.

He has over two decades of experience in the incubation of various initiatives at TCS. This includes setting up of a Microsoft Solution center in Redmond, big bet initiatives in the areas of Digital Marketing and IoT. He also setup the TCS Intrapreneurship Program (An Incubator for internal startups). He has led teams in technology, architecture, solutions, strategy, ecosystem development, product engineering & business development.

Digitization of Society










Re.Fresh

On the momentous occasion of TCS50 and in alignment with the larger organizational theme of “Experience Energy”, TCS Research & Innovation (R&I) introduced Re.Fresh – a quarterly Research symposium "of the researchers, for the researchers, by the researchers".

Designed to be both vibrant and visionary, Re.Fresh brings together members from our Research ecosystem in events featuring discussions, debates and demonstrations covering multiple areas of interest.

Digital Reimagination of Society

Arpan Pal

A society needs to ensure the physical well-being as well as the intellectual development of its people while providing for sustainable and comfortable living. This calls for research in healthcare systems, sustainable energy and utility systems, efficient agriculture/food production systems, innovative education systems, and citizen-centric services: all encompassed within the fabric of digital technology. In the context of the current pandemic, digitization and remote delivery of these services assume greater relevance and importance.

As Sir Jamsetji Tata said, "In a free enterprise, the community is not just another stakeholder in business, but is in fact the very purpose of its existence". TCS Research and Innovation, in the true tradition of the Tata group's philosophy, has been working extensively to create pioneering solutions to socially relevant problems. TCS Research and Innovation, in the earlier days, had built easy-to-assemble water filters, computer-based adult literacy tools, and mobile-based agro advisory services which made a real impact on society. We continue our mission with explorations in technology-enabled, yet affordable and inclusive solutions in healthcare, agriculture/food supply chain, energy/utility, education and citizen services through the novel use of digital technologies. The focus is not only on building these novel technologies, but also on making them business sustainable on their own, so that they can be scaled up across countries and continents.

In the previous edition of "Reimagining Research", we had covered genomics, connected health, energy and agriculture technologies. In this edition, we introduce novel and innovative applications for digital technologies in education, in food supply chains, followed by well-being at the workplace, and wellness through metagenomics.

In the first article, we introduce TCS iON, which tries to meet the personalization and self-development expectations of the millennial learners. Using a platform-driven approach, iON tries to drive transformations in India and other countries in learning, assessment, education management and operations. The article outlines the unique problems faced and the innovative ways they were solved in scaled deployment scenarios. iON is already proving to be a game-changer in the Indian education scenario in the context of COVID-19.

The next article in this series looks into a completely new area: how to ensure quality of food via food freshness monitors—a smart platform to estimate food quality and reduce wastage. Using sensor-driven IoT architecture followed by a digital twin for food, which employs physics, data and kinetics based models and knowledge bases for analytics, the system offers a disruptive way to improve the food supply chain.

The third article takes a fresh look at a very widely discussed topic: enterprise employee wellness. Instead of looking at the problem from different silos, it takes a new approach that combines multiple disciplines—from behavioral sciences to workplace norms and HR practices, from questionnaire-based evaluation to wearable-based 24x7 sensing, and from agent-based modeling of organizations to suggesting interventions for improvement—the proposed system tries a unique multi-modal approach to solve the very important, yet often overlooked issue of current and future workplaces. With work-from-home becoming the norm in the COVID-19 context, digital technology driven, holistic employee wellness (both physical and mental) becomes very important for all organizations across the globe.

Finally, the last article introduces a new and possibly pathbreaking concept on metagenomics for human wellness. Aptly titled "The Second Genome", it talks about how the genome analysis of microbiomes (also called

metagenome of microbiomes) inside our body holds the key to our future health, especially when screening for risk in asymptomatic patients. Combined with genomic analysis and the digital connected health model outlined in the earlier volume, the metagenomics approach provides a comprehensive way to tackle the problem of non-communicable diseases, which is projected to grow to epidemic proportions in the near future.

Arpan Pal

Arpan Pal is a Chief Scientist with TCS Research and Innovation and heads the Embedded Systems and Robotics Research Area. He has written and presented more than 125 publications and book chapters in reputed Journals and Conferences. He has also authored a book on IoT. He has filed for more than 95 patents and has 38 granted to him. He is a Senior Member of IEEE. He is on the editorial board for reputed journals like ACM Transactions on Embedded Computing Systems, *IEEE Transactions on Emerging Topics in Computing* and IT Professional Magazine from IEEE Computer Society.

Arpan received both his B.Tech and M.Tech degrees from the Indian Institute of Technology, Kharagpur, India, in Electronics and Telecommunications and Ph.D. from Aalborg University, Denmark.

Transforming Education Processes

TCS iON reimagines learning, assessment and related processes

Venguswamy Ramaswamy

IN BRIEF

This essay on TCS iON is distinctly different from the other essays in this collection. Here, the business model innovation is powerful. TCS iON introduced a Phygital model, ecosystem play, service delivery integration and a machine first approach to many processes in the education sector. It helps academic bodies raise efficiency in admissions, recruitment, learning, skilling, and business operations. It offers students new-age skill-enhancing methods and brings them closer to recruiters and corporates.

The intent is to deliver increased speed, widened scale, uncompromising security and superior experience to educators

After a deep understanding of the demands and expectations in the field of education, TCS launched a series of innovations for this sector, and has driven country-level transformations in learning, assessment, and campus processes. The transformations have been powered by TCS iON, an integrated, platform-based approach.

TCS iON is one of the strategic business units of Tata Consultancy Services, that provides business transformation platforms for educational institutions and examination boards. TCS iON enables these academic bodies to raise efficiency in admissions, recruitment, learning, skilling, and business operations. TCS iON has had a track record of reimagining traditional methods and processes in education by leveraging cutting-edge digital technologies. TCS iON's unique IT-as-a-Service model offers innovative, easy-to-use, secure, integrated, and hosted solutions on a build-as-you-grow, pay-as-you-use basis.

The primary intent behind TCS iON has been to exponentially deliver four key values to stakeholders: increased speed, widened scale, uncompromising security, and superior experience. We call this Value 4.0 for the Industry 4.0 era.

A Paradigm Shift in Assessment

When we started out a decade ago, the assessment practices in India were struggling with multiple

Fact File

TCS Research: TCS iON

Outcomes: Reimagination of traditional processes by leveraging digital solutions, thus delivering speed, scale, security, and a superior experience to all our stakeholders

Principal Investigator: Venguswamy Ramaswamy

Techniques used: AI, ML, Deep Analytics, Cloud Computing, New-age Learning Pedagogy, Immersive Learning, Experiential Learning etc.

Industries benefited: Recruitment Process across Different Sectors, Educational Institutes, Exam Boards, Corporates

Patents: 9

challenges. Traditional pen-and-paper exams were the norm, even though they were woefully lacking when it came to tackling the challenges of malpractices, paper-leakages, and the time to execute. We identified these gaps early and leveraged our prowess in the digital sector.

At TCS iON, we believe in using cutting-edge digital technologies to constantly reimagine traditional methods and processes. Through TCS iON, we developed a complete digital ecosystem that pioneered a distributed computer-based test (CBT) platform called TCS iON Digital Assessment. Since then, there has been no looking back. Today, TCS iON has become synonymous with computer-based tests and has recently reached the milestone of assessing 200 million candidates across 20+ countries.

As we got deeper into the education sector, we discovered significant opportunities to reimagine age-old practices.

For us, change is impactful only if it is adopted on a mass scale. By disrupting the standard way that learning, assessments, and campuses

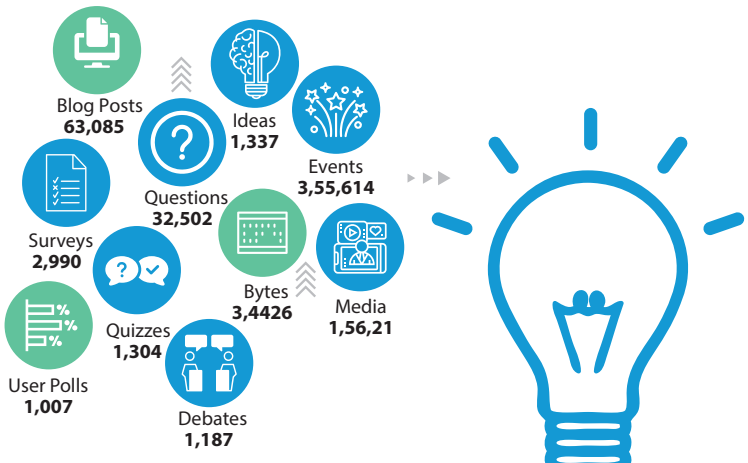


Figure 1: *Digital Learning Communities*

A phygital strategy is used as it is impossible to achieve the multiple objectives of learning through a digital-only approach

function in India and other countries, we have been able to drive country-level transformations. And we have also successfully scaled our business globally, by leveraging our integrated platform-based approach.

Five-Pronged Approach

Our approach to reimagination is deep-rooted in a set of transformative principles. Even though we have pitched our digital capabilities to the forefront of this transformation drive, we have integrated them with an organic human ecosystem that we have diligently built over the years. Let's take a look at our five-pronged approach to drive Value 4.0.

- **A "phygital" strategy:** In our endeavor to reimagine the education sector, we realized, early on, that a single-track digital approach does not work. In fact, it is impossible to achieve multiple aspects of learning through a digital-only approach. To use a simple example, imagine learning to drive a car using only digital learning tools. So, we knew we had to complement our digital solutions with physical infrastructure and hands-on learning methods. Thus, the Phygital strategy was born. Today, TCS iON has penetrated the remotest corners of India with its 6400+ examination centers. A candidate can just walk into an exam center in her city, at her convenience, without having to spend money or time on intercity commuting or accommodation! We have also leveraged the "phygital" model of delivery for the TCS iON Skill Hub, a platform dedicated to reskilling and upskilling learners with vocational skills and standardized content. The Skill Hub not only facilitates digital learning, but also integrates hands-on coaching with local infrastructural facilities.
- **The ecosystem over individual approach:** An example of the success of this approach would be our TCS iON Digital Learning Hub platform. The Hub integrates several learning platforms to provide unique learning interventions and is a market place for the top publishers of the world to come onboard and publish best-in-class learning content. This content can be consumed by millions of learners in an anytime-anywhere, any device model. The Hub further extends this ecosystem by collaborating with educational institutes and leading corporations. While industry and academic experts moderate the many learning communities on the platform to equip students with industry-aligned skills, corporate recruiters get a ready pool of pre-qualified, skilled candidates to fulfill their recruitment needs.

  The community structure of learning helps students interact directly with industry experts and academics. The entire ecosystem of content partners, academics, industry experts, and learners form a symbiotic culture, where each harnesses the potential of the other, thus breaking down silos.
- **Service delivery integration:** The value addition that TCS iON brings increases exponentially due to its ability to control the underlying *hardware*, the overlaid *software* platforms, the *data/content* carrier, and the orchestrated *services*. To

achieve this, the team devised a unique service delivery architecture that integrated hardware, software, content and services into a single digital platform. This helped us provide new, customized offerings for multiple segments in various geographies; offerings that could help manage the end-to-end operations for our customers. To see this solution in action, take for example, the TCS iON Digital Assessment and Marking for language tests. In language tests, there is a need to assess and evaluate candidates on their ability to Read-Write-Listen-Speak (RWLS). To address this need, the offering comprises:

a. TCS iON PAPER™: A handheld educational appliance crafted to cater to the examination requirements of 21st century students. The appliance assesses RWLS skills.

b. TCS iON Digital Marking: A platform to score all formats of responses, whether it is audio, video, digital answer sheets, or scanned copies of answers. The marking scheme details the approach to scoring and facilitates unbiased, fair marking.

c. TCS iON Markers' Hub: This is a service where evaluators or markers are onboarded and trained to render marking services to any exam, anywhere in the world.

By leveraging the individual capabilities of each of these components and bundling them together, TCS iON delivers location-agnostic, fair, and secure marking services to the world, at speed and scale (Figure 2).

- **Machine first approach:** TCS iON has created value in every dimension by adopting the machine first approach, which gives the first right of refusal to machines. By harnessing the power of automation, the application has made a significant difference to the world of assessments and business processes.

a. TCS iON PAPER™, a handheld educational appliance packed with features such as custom OS and controlled access to hardware and software, is an ideal educational tool for students, institutes, as well as exam bodies. TCS iON PAPER™ disrupts the traditional, manual-labor-intensive process that is heavily

A unique service delivery architecture integrates hardware, software, content and services into a single digital platform

Digital Learning Hub

TCS iON Digital Marking

TCS iON PAPER™

Figure 2: *TCS iON Products*

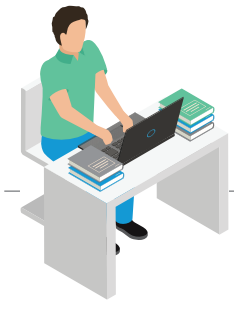
It was important to create a symbiotic ecosystem of content partners, academics, industry experts and learners

dependent on logistical planning. To take a specific example, TCS iON DigiOMR solution powered by TCS iON PAPER™ comes with powerful and intelligent features such as photo and biometric capture, question paper encryption, auto-scrolling, auto-marking, color-coding, and even an integrated timer, among others. It eliminates time-consuming and expensive processes such as printing question papers and Optical Mark Recognition (OMR) sheets, storing and transporting, manually distributing and ensuring security throughout. The solution is so effective that encrypted question papers can be digitally distributed securely, with a single touch. The results can be declared on the exam day! It translates into a superior experience for all stakeholders, while also addressing security concerns.

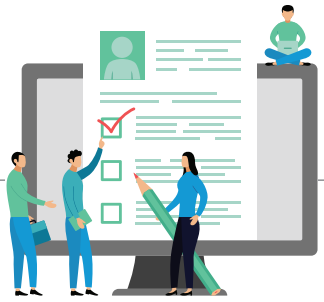
- **Self-learning platforms:** In the Industry 4.0 era, integrating machines with human intelligence by leveraging Artificial Intelligence (AI) and Machine Learning (ML) can drive exponential value. The TCS iON Command Center, which is one of TCS iON's most futuristic innovations, works on this principle and is focused on self-healing and self-learning. It is a state-of-the-art technology that gives administrators a bird's eye view of their operations, in real-time, throughout the process life cycle. The platform has completely reimagined the way high-stake operations are executed. It provides administrators with intelligent alerts for proactive and timely actions. The platform leverages ML to predict the preparedness of operations, based on past incidents. Its IoT and AI capabilities detect patterns and communicate the relevant intelligence to the administrators.

Innovations for a Brighter Tomorrow

With a plethora of innovative features and capabilities, TCS iON has transformed the domain of assessment, learning, and business processes management in India. In the assessment domain, it has successfully ushered in an era of secure and inclusive examinations with TCS iON Digital Assessment



TCS iON Markers' Hub



TCS iON Digital Assessment

TCS iON Digital Campus

and TCS iON Digital Marking. In the field of learning, the application has shifted the focus from traditional, classroom-based learning methodologies to new-age skill-enhancing ones, through learning constructs such as TCS iON Industry Honour Certification and TCS iON Career Skills. We have democratized the fresher-recruitment process by bringing corporations and jobseekers on a single platform to facilitate easy access and best-fit job matches.

Today, TCS iON Qualifier ensures that a fresh graduate from a lesser known college in a tier-3 city has as much access to top corporate recruiters as a graduate of a premier college from a tier-1 city. And that's not all. To ensure that educational institutes can better manage a student's experience of learning, TCS iON has enabled institutions to manage their administrative operations seamlessly with solutions such as TCS iON Digital Campus and TCS iON Digital University.

TCS iON touches the lives of millions of students, families, and professionals. A lot remains to be done, and many miles to go before we sleep, but we continue to work hard to create innovative and disruptive solutions that challenge the status quo.

Venguswamy Ramaswamy

Venguswamy Ramaswamy alias "Swamy," is the Global Head of TCS iON platform solutions. Swamy was instrumental in setting up the unit, and in driving the domain solutions across multiple sectors including Education, Manufacturing and Recruitment. Swamy serves as an Additional Director on the Board of MP Online Limited, a joint venture company between the Government of Madhya Pradesh and Tata Consultancy Services. Earlier, Swamy has held important positions in TCS such as Director of TCS' Global Consulting Practice (GCP), Head of the "GE" account, the Process Consulting Group, and Corporate Resource Management to name a few.

A Six Sigma Master Black Belt, Swamy was named amongst the top 25 consultants of the year 2007 by *Consulting Magazine*. Swamy holds a Masters degree in Computer Applications from Madras University.

Food Freshness Monitor

A smart platform to estimate food quality and reduce wastage

Jayita Dutta,
Shankar Kausley, et al.

IN BRIEF

Food wastage is at the level of a global crisis. Producers, logistics companies, governments, food health enforcement bodies, sustainability advocates, and concerned citizens are trying multiple options to reduce wastage.

Real-time monitoring of food quality has emerged as a viable solution that can benefit all stages of the food supply chain, starting from farmers to end-consumers like us. It helps reduce significant economic losses which occur due to food spoilage and wastage while retaining quality and nutritional value. A novel integrated "smart digital platform" is being developed to estimate and predict food quality. This will enable all stakeholders of the food supply chain to make decisions dynamically regarding altering supply chain logistics and storage conditions, for repurposing and minimizing food spoilage and wastage.

A platform that can gather data from multiple types of sensors, with data processing capabilities, can provide much value in reducing wastage

Roughly one-third of the food produced in the world for human consumption every year—approximately 1.3 billion tonnes—is lost or wasted[1]. A large amount of food produced gets wasted in different parts of the supply chain such as farms, storehouses, transportation, and processing units (Figure 1). Along with this, a significant amount of food wastage is also experienced at the consumer end.

Of the total global food wastage 40–45% is fruits, vegetables, and root crops[2]. Food wastage and spoilage within the food supply chain can be attributed to uncertainties in demand and supply, delays, and changes in environmental conditions at different stages of the food supply chain. This is equivalent to an annual economic loss of USD 1.2 trillion[3]. It also causes a significant environmental footprint amounting to about 8% of greenhouse gas emissions worldwide[4].

As a result, food and drink manufacturers, supply chain managers, retailers, governments,

1 [© FAO] [2017] [FOOD LOSS AND WASTE AND THE LINKAGE TO GLOBAL ECOSYSTEMS] [Page number 1 (for publications)] [http://www.fao.org/save-food/resources/keyfindings/en/] [Date accessed: 21/10/2019]

2 http://www.fao.org/save-food/resources/keyfindings/en/

3 https://www.bcg.com/publications/2018/tackling-1.6-billion-ton-food-loss-and-waste-crisis.aspx

4 http://www.fao.org/3/a-bb144e.pdfmarked

Fact File

TCS Research: Physical Sciences

Outcomes: Novel Platform for Estimation of Food Freshness and Shelf Life, Development of Sustainable Food Supply Chains, Decision-Making System for Food Retail, Consumer Guidance System on Quality and Nutritional Value of Food

Principal Investigator: Dr. Beena Rai

Techniques Used: Chemical Lab Experiments, Physics-Based Models, Artificial Intelligence, Data Analytics, Sensing and IoT, Cloud Computing

Industries Benefited: Food Supply Chain, Food Processing, and Healthcare/Wellness

health authorities, consumer forums, sustainability-focused NGOs, and responsible citizens are all keen on reducing food wastage.

Farm to Plate View

Concerned groups are looking at multiple ways of reducing waste in storage and along the supply chain. Some leading carriers are already equipping their containers with sensors and connecting them to a platform. A platform that can gather data from multiple types of sensors, with data processing capabilities, can provide much value in reducing wastage. In

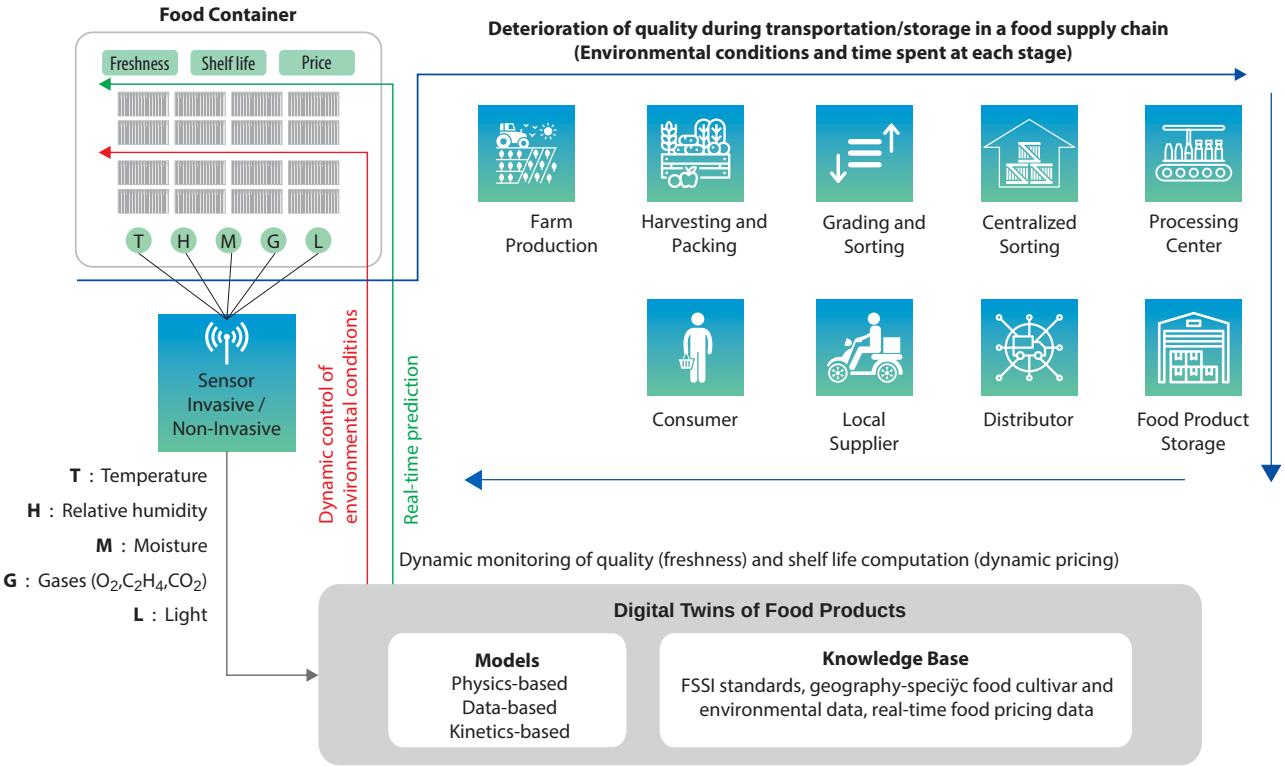


Figure 1: *Framework for Real-Time Food Quality Monitoring in Supply Chains*

order to optimize supply chain logistics it is necessary to equip all stakeholders—from farmers to end-consumers—with dynamic decision-making capabilities and real-time monitoring of various parameters. Such a system enables quick decisions on early offloading, rerouting, ramping-up of storage, and even food repurposing, to extend shelf life. The continuous monitoring and prediction of food quality at all stages of the food supply chain can create direct savings for stakeholders and result in many indirect benefits move these words after impact stable food prices and a reduced environmental impact – for the larger community.

TCS' Smart Food Freshness Monitoring Platform

TCS' platform (Figure 2) for predicting food quality is built using a collection of IoT-enabled sensors to sense Galvanic Skin Resistance (GSR), Near-Infrared (NIR) imaging, ultrasound, pH, air quality, gas composition (CO_2, NH_3, C_2H_4, O_2, etc.), weight, Volatile Organic Compound (VOC) content, camera feeds, and chemically analyzed food parameters. This is coupled with data analytics, image processing, and cloud computing. This platform allows provision for variation of environmental

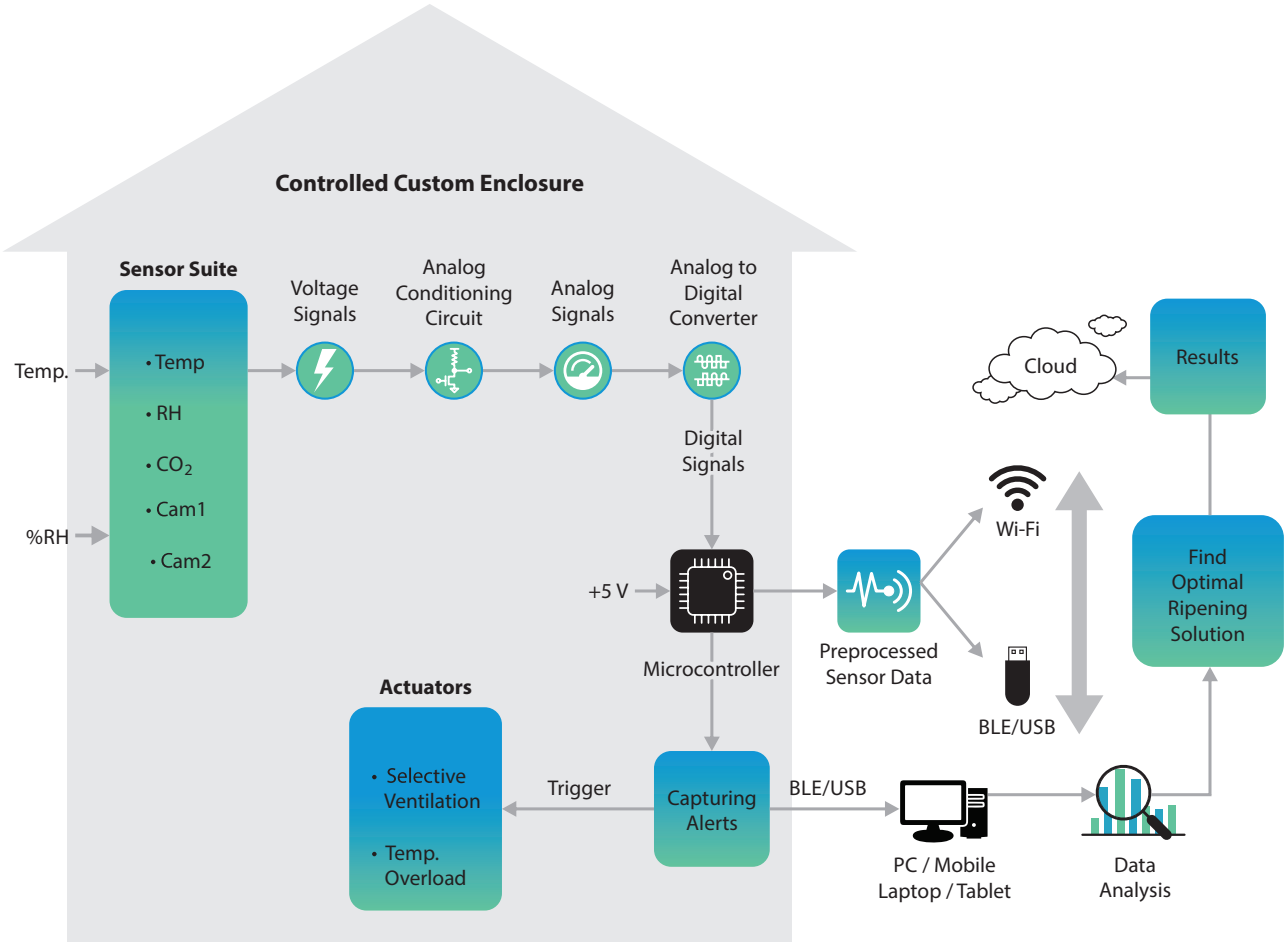


Figure 2: *Food Quality Monitoring Platform*

Predicting Ripe Time

The TCS Smart Food Monitoring Platform was used to predict ripening duration of climacteric fruits (banana, mango, papaya, etc.) and predicting the shelf life of potatoes for different customers.

Climacteric fruits need a lot of flexibility in a supply chain. Overripe fruits become soft and have to be handled with much more care; the act of ripening releases ethylene which may affect the ripening of other fruits close by.

The TCS platform was used to study the natural ripening process of climacteric fruits, both for a farmer and a leading metro retailer. It helped prevent over-ripening, by providing alerts on the variation of environmental conditions. This platform was used to monitor the quantitative variation in the quality of "Cavendish"—a popular banana cultivar w.r.t. ripening index (RI). Combined quantitative sensing of O_2, CO_2, sugar content, and images during ripening stages of the banana were used to calculate the respiration rate (RR) and establish correlation between the RR and the RI - to find an optimal ripening schedule. Further, the platform was used to predict the ideal ripening duration and factors affecting the natural ripening process at different environmental conditions.

For the retailer, real-time information on changing quality and remaining time for ideal ripening was provided so that the store could introduce dynamic price against quality grading. For example, if there were three stacks of bananas with different ripening durations, they could be priced differently to ensure economic benefits and waste reduction.

The platform was also used for predicting the shelf life of potatoes for different supply chain scenarios. Potatoes have different uses such as the production of chips and fries, and as potato starch for domestic consumption. The supply chain for potatoes includes harvesting, curing, storage, processing, and logistics. Our platform predicted the shelf life of different varieties of potatoes at different environmental conditions for different applications in terms of weight loss, sprouting, fungal growth, change in sugar content, and dry matter.

Our platform with soft sensor technology coupled with precision modeling and digital imaging can revolutionize existing food supply chains.

conditions such as temperature, humidity, light intensity, etc., thus, simulating conditions similar to those experienced by food in the field, during storage and in transit. The platform incorporates a modular framework. It consists of a custom-designed enclosure where different supply chain scenarios can be simulated.
The enclosure consists of a multi-modal sensor suite with a modular framework and is inclusive of multiple IoT-enabled sensors.
The sensors and offline data are interfaced with digital media via a

Multivariate, multi-modal, synchronized, time series, online, and offline lab data collected at predefined intervals have been used to train these models to accurately predict food quality

Bluetooth Low Energy (BLE) module or a Wi-Fi module. The data is further sent to the cloud to allow real-time monitoring of the sensed parameters across the globe.

The platform incorporates hybrid models (AI/ML/physics-based) built on non-invasive sensory and offline lab data simulating all possible supply chain scenarios. These act as soft sensors in the real-time prediction of food quality. Multivariate, multi-modal, synchronized, time series, online, and offline lab data collected at predefined intervals have been used to train these models to accurately predict food quality. These hybrid models provide a standardized digital signature of the food under study and enable prediction of its quality, freshness index, and remaining shelf life. The predictions enable development of a feedback system, which allows alteration of the environmental conditions to increase shelf life or repurpose the food (such as ship it to a nearby location or take other appropriate action), and in turn, reduces food wastage.

This platform, once deployed in a real-world scenario, will assist stakeholders to take dynamic decisions relating to modified logistics and changes in environmental conditions. Thus, this platform will help in building sustainable supply chains and reducing impact of global food wastage on hunger, the economy, and the environment.

Jayita Dutta

Jayita Dutta is a Scientist with Physical Science Research at TCS. During her six years in the company, she has worked in various research areas, including perovskite material discovery and synthesis, fabrication of perovskite solar cells and solar panels, food sensing, and data analytics.

She has filed a couple of patents in India and other geographies. Her current work involves monitoring of food quality, which is aimed at benefiting society and the environment. She has completed her Bachelor's in Engineering (BE) in Electronics and Communication Engineering (ECE) from Bengal Engineering and Science University, Shibpur, and an M.Tech in Solid State Technology from the Indian Institute of Space Science and Technology (IIST), a premier institute under ISRO. Her work has been published in IEEE Standards.

Shankar Kausley

Parijat Deshpande

Well-Being at the Workplace

A multidisciplinary approach and scalable digital tools

Mayuri Duggirala,
Sachin Patel, et al.

IN BRIEF

Many organizations have employee well-being programs, but even with these in place, mental health issues that need clinical intervention and/or counseling go undetected. Often, help arrives only after the problem is very severe, or when it has resulted in adverse outcomes. Behavior, Business, and Social Sciences (BBSS) research at TCS supports proactive initiatives for associate well-being. Here, we discuss some elements that can enable realistic modeling of organizational agents and an easy-to-consume interface that can improve well-being among employees.

A 24-year-old man with chronic depression takes to drugs and alcohol, unable to deal with stress at work; a 30-year-old woman ends her life because she is deep in debt due to multiple family responsibilities; a middle-aged telephone operator hangs himself in the office in the middle of his work day because of multiple health issues. It is not uncommon to see such headlines in newspapers. To many of us, these may not appear to be strong enough reasons to hurt ourselves, or end our lives. What becomes apparent though is a need for greater awareness of mental well-being and support for those with mental health issues at the workplace.

Constant and rapid change often requires today's workforce to not only adapt to external unforeseen changes but also have sufficient psychological resources and capabilities to respond to change in an effective manner. Capabilities such as resilience, stress management, and sound decision-making are some of the often-cited employee-related strengths critical to an enterprise's success as well. The TCS R&I has well-being as one of its focus areas as part of Behavior, Business, and Social Sciences (BBSS) research. At BBSS, we use a multidisciplinary lens to understand states of well-being within the enterprise and to guide associates toward improved well-being.

Fact File

TCS Research: Behavior, Business, and Social Sciences

Outcomes: Behavior sensing and analytics, evidence-based interventions, purposeful games, knowledge repository, well-being app

Principal Investigators: Vivek Balaraman, Sandeep Athavale, Sachin Patel, Mayuri Duggirala

Academic Partners: MIT, University of Toronto

Techniques Used: Data Science, Modeling and Simulation, Behavioral Research Methods, Game Design Research

Industries Benefited: Life Sciences, Healthcare, Pharma

Patents: 3

Papers: 8

Need for Multi-Disciplinary Approaches to Well-Being

Many organizations have employee assistance programs. Even with these mechanisms in place, many problems and challenges that need urgent clinical and/or counseling attention go undetected or reach the employee assistance program only after the problem is very severe and has resulted in adverse outcomes such as self-harm or suicide. This is due to several factors, ranging from the stigma associated with mental health problems, privacy concerns, and the lack of mental health awareness, to name a few.

Given the sheer scale of the problem of mental illness incidence worldwide, and the criticality of its management for the individual, enterprise, and society as a whole, there has been a sustained interest in the use of digital technologies to develop more scalable solutions for mental health management. Emergence of areas such as telepsychiatry[1], digital mental health[2], and digital therapeutics[3], reflects the growing application of technologies in the field of mental health. While there have been significant strides made in the development of these technologies for mental health, concerns regarding robustness, scalability, and evidence-driven technologies continue to dominate the field[4].

1 https://www.psychiatry.org/patients-families/what-is-telepsychiatry

2 https://psycnet.apa.org/fulltext/2017-39812-001.html

3 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6393746/

4 https://www.weforum.org/agenda/2019/01/power-digital-tools-transform-mental-health-care-depression-anxiety/

Given the sheer scale of the problem of mental illness incidence worldwide, and the criticality of its management for the individual, enterprise, and society as a whole, there has been a sustained interest in the use of digital technologies to develop more scalable solutions for mental health management

Digital Solutions for Well-Being

Building a scalable digital solution which can help address issues such as early identification of burnout, stress detection, and management and prevention of self-harm at the workplace not only demands expertise in the behavioral sciences, but also in related areas (e.g. workplace norms, HR policies) and in areas such as image recognition, voice analysis, big data management, and analytics, among others. The solution has to have several components such as:

- A knowledge base that supports behavior modeling related specifically to the workplace
- Image and voice recognition algorithms that can detect anomalies in the subject
- A platform that can fuse data from various sources and anchor the solution
- An engaging mobile app that serves as a companion and confidant to the user
- A robust measurement and intervention system for enhancing well-being in the enterprise

BBSS research has focused on each of these components and we would like to briefly describe them here.

A Knowledge Repository

The need for repositories as sources of information in a domain has been a critical requirement across scientific disciplines. There are knowledge repositories related to behavioral sciences, but the focus here must be on:

A structured and usable repository amenable to computational behavioral modeling, synthesis of knowledge, generating evidence, identifying and generating new areas for research.

The repository has to be a structured entity, such as a set of tables in a well-recognized standard entity such as a relational database management system (RDBMS). This is to facilitate both the entry/modification of information into the repository as well as its querying to obtain its contents. Note that querying and consumption of the content should be possible both by a human or a program.

The creation of a minable repository of a vast area such as human behavior requires many steps such as:

1. Creating a corpus of papers in the behavioral sciences domain
2. Identifying relevant papers using a low-cost multiclass classifier
3. Identifying sections in each relevant paper that have empirical results of interest using intra paper classifiers (This second-level classifier is built because techniques such as information extraction in natural language processing are compute-intensive) The process splits as below:

3.1 Information tagging and extraction techniques to identify behavioral variables, the relation terms, the strength and the confidence of the relation, as well as other information considered relevant from text, using a text relation miner;

3.2 A table extractor to process the tables (if any) in the paper to extract quantitative empirical findings

4. Combining the information obtained through the relation miner and table extractor to put together a complete record of an empirical finding using an integrator

5. Verification of the extracted knowledge by domain experts, after which validated knowledge is now moved to the main behavior repository

Several challenges have to be overcome in creating the repository such as the lack of domain experts as well as varied structures and formats of the reports and tables. There is also the issue of the lack of definitive solutions for processing tricky areas of natural language and written text. A separate machinery may be required to create resources for activities such as correcting error terms, mapping synonyms to each other, expanding acronyms and abbreviations, identifying other semantic relations between terms such as hypernymy and hyponymy, adding context to a term in a particular usage so that there is a differentiation between different uses of the same term, among others.

Modeling Realistic Organizational Agents

One of the goals of the development of the Behavioral Relations Repository is to demonstrate the use of hypotheses from behavioral and management sciences in modeling behavior across the organization. Agent-based modeling and simulation (ABMS) is amenable to the study of a wide range of behavioral dynamics in a given context. ABMS has been traditionally used in the study of complex systems, but its applications to organizational psychology have been sparse. ABMS has been applied to the study of macro behavior in the realms of social psychology; however, its use in studying the micro-foundations of behavior has been limited. To bridge this gap, fine-grained agent-based models can be used to develop realistic models of behavior anchored in the domain of organizational behavior.

Fine-grained agent-based models can be used to develop realistic models of behavior anchored in the domain of organizational behavior

A Well-Being App

There are many mobile apps that are meant to help with measuring and improving one's well-being, but many of them are meant for those who are already seeking help or already know they have a problem. Today, it is possible to sense emotions through voice, facial expressions, and activity. A mobile phone today can sense so much about its owner: acts of physical activity, tone of voice, facial expressions, and lifestyle behaviors (sleeping/shopping/spending/interacting), among others, that it becomes a source of information about a person's well-being. Within the limits of GDPR, much of this data can be analyzed and alerts may be provided to the subject/caregiver on anomalies from usage patterns within the well-being app. If the app is further gamified and endears itself to the user, both accurate data collection and intervention may be possible.

Engagement	Sensing	Analytics	Interventions	Intelligence
T says "Good Morning," detects good weather and says - "lets take it one step at a time, start our morning routine and go for a walk."	T senses if the user really got out of bed, completed the scheduled routine and walked using pedometer	T knows that most mornings the user sleeps till late, has an erratic routine and avoids walk	T says I am going to hide in the park, find me. T asks "Would you like to connect with the doctor today?"	T knows that user does not go for walk suggests new interventions or interventions that have worked in the past
T asks "How are you feeling, how do you see your day going?" and shares a suggested plan for the day	T asks user to take a selfie. Senses that user is anxious and sad.	T knows that the user is anxious and sad when they go to work or find themselves alone at home	T suggests that the user meets or call friends and/or family after work	T learns that user is anxious and sad when sitting alone or lying in bed not sleeping but lost in thought
T suggests the user eats some fruit and stays hydrated using a nutrition counter	T senses that the user has been sleeping erratically for most of the day and has ignored all messages so far	T knows user has ignored most calls from family, friends or work this week	T suggests that the user write a daily diary of experiences or speak to the therapist or complete the assignment given in therapy	T learns that the user is cheerful while around a few people the user likes and trusts T nudges joint activity with trusted ones
T plays soft music and asks user to breathe along, relax and go to bed.	T senses that user is unable to sleep and has not eaten much during the day	T knows that user is in good mood during 2nd half of the week	T suggests user to meet friends and exercise one day a week and increase it gradually	T monitors prescribed medication intake as per prescription and sets reminders for timely medication

Figure 1: *T for Depression Sensing and Intervention— A sample case*

Advancements in NLP and related areas have helped evolve novel approaches in addressing challenges in creating knowledge repositories

An Anchoring Platform

Central to a digital solution is a platform for multi-modal behavioral data capture through numerous data capture devices, sensors, and secondary data sources; behavioral knowledge mining from multiple sources; behavioral analytics; behavioral knowledge delivery to various simulators, games, and reasoners. The platform acts as the hub for the entire ecosystem of tools and applications. It must have the Behavior Knowledge Repository as its source of knowledge. The platform must have analytics capabilities such as a personal well-being dashboard, population level well-being analytics, and a pattern discovery module. Other capabilities include a manual intervention module and a behavior study design wizard in addition to the behavior relations repository and an interventions repository. In keeping with current GDPR policies, the platform must be compliant with privacy and other aspects outlined in standard GDPR practices that are in effect today.

The Road Ahead

In building the digital solution with the components mentioned above, we have had a taste of success (View Box). The knowledge repository that we are building will keep growing as behavioral science is a vast area encompassing many domains such as psychology, cognitive science, and sociology, among others, and new knowledge is added constantly across all areas of science. Challenges in building the knowledge repository include the reproducibility debate, which focuses attention on the need for replicating evidence to test for its

robustness and generalizability across different settings.

Advancements in NLP and related areas have helped evolve novel approaches in addressing these challenges in creating such repositories, and work on these aspects is underway at BBSS. The serendipitous value we have discovered is that many other areas of research are keen on learning our techniques in building knowledge repositories in other domains as well.

Both the digital platform and the app will gain new features and usage through our continuing field trials. We envision these tools being used in enhancing employee well-being as well as other contexts where human behavior measurement and change is of importance. This could include problems in the space of change management, competency management, learning, and many more.

T

TCS endeavors to be proactive in employee support initiatives. Being a technology company spearheading digital transformations for leading businesses, it uses digital tools for several employee initiatives such as fitness, marathon running, accessibility, and social responsibility. As mental well-being is a top concern for TCS HR, it has approached the BBSS research team for various studies, notably in workplace stress and productivity among support service teams; individual productivity and well-being; and the link between perceived supervisory support and subjective and objective performance measures, among others.

We were asked to build a scalable solution to foster mental well-being. The solution has come from our research in BBSS and is anchored on a platform called Beacon. It draws knowledge from the domain repository we built, it analyzes facial expressions and voice anomalies of the user, it holds the mobile app we call T, and it has scoring mechanisms and intervention protocols. We would like to highlight the consumer end of the solution "T" here.

T is a digital companion based on the concept of a pet dog . The 1990s saw the rise of virtual pets. Since a large percentage of TCSers are millennials, the idea of a virtual pet is an accepted concept. Our Serious Games research team created the character and tested it with the target audience who said virtual pets are fun to engage with, and valued the chance to share emotions without being judged.

T uses the latest techniques in the area of purposeful games. It is data privacy-conscious and built upon numerous sensing technologies to measure well-being and assist people in adopting behavior changes that help them stay in the positive end of the well-being

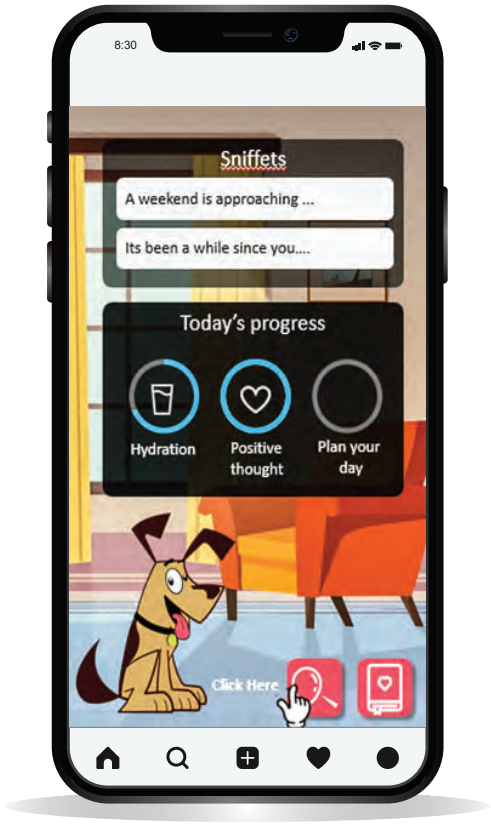


Figure 2: *T Landing Page*

spectrum. We have carried out initial user pilots with T and are currently refining T to incorporate these inputs as well as carry out larger pilots across the enterprise. In terms of physical sensing, Tappy's capabilities include emotion detection with speech, text, and camera images as inputs. Other capabilities include laughter and blow detection, physical activity sensing, and heart rate sensing.

Scenarios for T

T is currently envisioned as a Digital Companion App for well-being at work including measurements and interventions targeted at workplace stress management. Other behavioral change interventions that can be enabled by T include a more balanced lifestyle, improved relationships, enhanced motivation, and other well-being behaviors of interest. In addition, T is also amenable for applications in other scenarios such as managing depression. Figure 1 illustrates this application. As shown in Figure 1, T's sensing, measurement, and intervention capabilities enable an individual living with depression to track their moods, activities, and lifestyle over time.

As a companion for therapy for individuals with depression, T presents a promising scenario that is being explored with our partners in healthcare and insurance contexts as well as with clinicians working with patients suffering from depression. Many patients require a combination of both medication and therapy to manage their symptoms and to stay productive. T's capabilities potentially allow for the therapeutic interventions to be delivered to the user at prescribed times as well as be sensed/tracked for medication and therapy follow-ups with the clinicians. Given that the doctor's appointments are likely to be short and sporadic, T allows the patient, clinician, and the caregiver to manage the treatment regimen longitudinally. Once we have a more robust version of T ready for the workplace well-being context mentioned above, we plan to explore T's applications for depression management (Figure 2).

Mayuri Duggirala

Mayuri Duggirala is a senior scientist with the Behavior, Business, and Social Sciences Research team at TCS R&I. At BBSS, her work focuses on the applications of behavioral science research to various problems of interest to the organization and the larger stakeholder ecosystem. In particular, Mayuri works on examining well-being and its many facets and areas of impact using a technology-driven approach. Her work in BBSS research also focuses on applications of fine grained behavioral modeling and simulation, technology-enabled behavioral measurement and change, serious games and service design for behavior change. Her recent work spans well-being and related areas such as stress, productivity, engagement, employee satisfaction, etc.

Mayuri holds a Ph.D. (Management) from the Indian Institute of Technology, Madras, with an academic background in applied psychology and organizational behavior. She has also worked in the areas of technology-mediated learning, healthcare services quality, and IT-enabled services research, technology adoption, learning, etc.

Sachin Patel

Sachin Patel is a Senior Scientist with the Behavior, Business and Social Sciences Research Team at TCS R&I. His research interest is in behavior sensing and its applications in personal informatics/quantified-self systems. He has led multiple innovation projects from ideation to deployment. In his previous stint as a software professional, he acquired extensive experience in managing and delivering enterprise software projects.

Sandeep Athavale

Vivek Balaraman

The Second Genome

Microbiome-based diagnostic markers

Sharmila Mande

IN BRIEF

We carry more microbe cells than our own cells! We know that these microbes play a crucial role in our health and well-being. Although one can now obtain genome sequence data corresponding to these microbes using state-of-the-art sequencing machines, gaining meaningful insights from such complex and voluminous data is challenging.

TCS' Life Sciences Research team has not only developed efficient algorithms for every step of analysis and management of the data, but has also carried out cutting-edge research to understand the link between microbial community (called "microbiome") and diseases. The team's innovative microbiome-based solutions can be utilized for not only monitoring the health status of an individual, but also for predicting the risk for a number of asymptomatic diseases and disorders at an early stage.

Our algorithms can be used for not only understanding microbial communities present in our body, but for obtaining key insights into the role of such communities which are present in other environments (soil, water, etc.) too

Can you believe that the microbes residing inside us outnumber not only our own cells, but also contribute close to a billion microbial genes as against about 20,000 human genes? Thus, apart from our own genome, we have a second genome (also called metagenome) contributed by the DNA of all the inhabiting microbial communities, collectively called the microbiome. In order to understand whether and how these tiny microbes residing within us influence our health, we need to study their genomes. As most of these microbes cannot be cultured in the laboratory, they cannot be studied using traditional genomics approaches. Thanks to the progress in science and technology and engineering, a new generation of DNA sequencing machines can now sequence the entire DNA of all the microbes that reside in any environment. This field which bypasses the culturing step is called metagenomics.

The field of metagenomics has advanced rapidly in the last decade. Enormous amounts of DNA data has been/is being generated and analyzed. However, there are some challenges in analyzing DNA sequencing data obtained from a microbial community:

- The sequencing data is voluminous and noisy
- Analysis requires complex algorithms, given the fact that we have no prior idea about the

Fact File

TCS Research: Metagenomics and Microbiome

Outcomes: Several patented algorithms for analysis, visualization, and management of metagenomics data, first-of-its-kind "Microbiome-based biomarkers" for risk assessment of diseases/disorders

Principal Investigator: Sharmila S. Mande

Academic Partners: National Institute of Cholera and Enteric Diseases (NICED, Kolkata), Translational Health Science and Technology Institute (THSTI, Faridabad), Dr. Mohan's Diabetes Research Foundation (MDRF, Chennai), Hinduja Hospital (Mumbai), Institute of Genomics and Integrative Biology (IGIB, Delhi), National Centre for Cell Sciences (NCCS, Pune), Institute of Advanced Study in Science and Technology (IASST, Guwahati), Tata Chemical (Pune), Novo-Nordisk Foundation Centre for Basic Metabolic Research (University of Copenhagen, Denmark)

Techniques Used: Next-Generation DNA Sequencing, Genomics, Metagenomics, Functional Metagenomics, Data Mining, Machine Learning

Industries Benefited: Healthcare, Insurance, Pharma, Biotech

Patents: 84 filed, 23 granted

Papers: 92 in peer-reviewed international journals

types and amount of bacteria to look for

- At each step of the analysis, specialized algorithms, visualization tools, and statistical methods are needed to obtain biologically meaningful insights from metagenomic data

TCS' Life Sciences research team has developed efficient algorithms for not only every step of the analysis, but also for management of metagenomic "big data". These algorithms can be used for analyzing metagenomic data as well as for comparing multiple metagenomic datasets sampled across space and/or time. This is specifically important if one wants to compare datasets corresponding to healthy and diseased states. Our algorithms can be used for not only understanding microbial communities present in our body, but also for obtaining key insights into the role of such communities which are present in other environments (soil, water, etc.).

Screening for Risk of Asymptomatic Diseases

When we talk about human health, the first question which comes to our mind is whether we can diagnose diseases at a very early stage so that curative measures can be taken. A large number of diseases and disorders, like cancers, diabetes, cardiovascular diseases do not show any symptoms in their early stages. In a healthy

Unlike existing preterm birth diagnostic solutions, which are only applicable at later stages of pregnancy, our biomarker works in the first trimester with significantly high accuracy (>95%)

individual, there is a fine balance between various microbial groups residing in the body. This balance is lost in diseased individuals. Therefore, "variations" in the bacterial communities residing within us have the ability to serve as "diagnostic biomarkers" that can foretell the presence and/or stage of disease. Taking this as the clue, TCS' Life Sciences Research team has captured these patterns of imbalances, i.e. "variations" in order to develop microbiome-based diagnostic markers.

Preterm Deliveries

To give an example, take the case of preterm births (PTBs). Every year, about 15 million babies are born preterm[1] . Out of them, more than a million die due to medical complications. Premature birth results in more neonatal deaths than any other condition. Every 30 seconds, a new born child dies. A preterm delivery is difficult to predict and at the moment, there are no diagnostic methods that can accurately raise an early alarm. Therefore, TCS' microbiome research team looked beyond physical traits as well as biochemical tests and focused on studying the bacterial communities in pregnant women. We studied a lot of publicly available microbiome data from pregnant women and deciphered patterns that have the ability to act as biomarkers for preterm delivery. Unlike existing PTB diagnostic solutions, which are only applicable at later stages of pregnancy, our biomarker works in the first trimester with significantly high accuracy (>95%). Our biomarker relies on metagenomic sequencing and analysis of a single microbiome sample (either a vaginal swab or a saliva/stool sample) collected from a pregnant woman. The

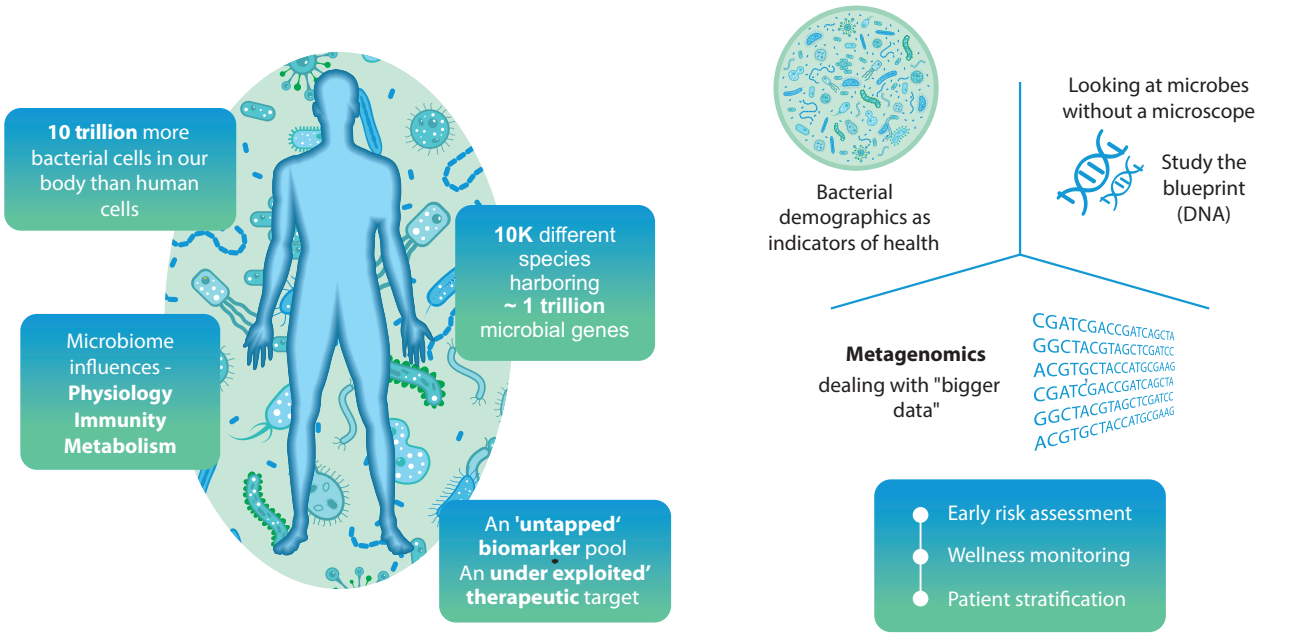


Figure 1: *Microbiome: New Descriptor of Health*

1 https://www.who.int/news-room/fact-sheets/detail/preterm-birth

The Gut Factor

The TCS "Gut Health Score" can be used to monitor an individual's gut health status with ease.

Your stool sample can point to the presence of several diseases, some of them serious and often asymptomatic. No bullshitting, this! Analysis of metagenomes is growing in importance as a diagnostic tool; and a stool sample is an indicator of not just the health of our gut, but of other organs as well.

The gut microbiome is a fascinating one; here are some reasons why:

- A majority of microbes in our body reside in our gastrointestinal tract (gut).
- While moving down the alimentary canal, food is stopped by trillions of microbes in the gut microbiome. (The total weight of all the microbial community in the gut is estimated to be around 2 kg – which incidentally is similar to the weight of our brain!)
- Two-thirds of the gut microbiome is unique to each individual, the uniqueness being dependent on the food we eat, ethnicity, health condition, and other environmental factors.
- Some food we eat cannot be digested by our own enzymes. When the stomach and small intestine are unable to digest such foods, certain gut bacteria help ensuring that we get the nutrients that can be absorbed by our intestine. Certain gut bacteria modulate the secretion of specific hormones (e.g. serotonin) and vitamins (vitamin B and K). The gut microbiome also plays a role in bioavailability of drugs as well as resistance to antibiotics.
- Aberrations in gut microbial communities have been shown to be connected with several diseases and disorders. Interestingly, the gut microbial community also controls functioning of distal organs like the heart, brain, lung, and liver.

TCS' Life Sciences Research team has been actively working on the gut microbiome and has demonstrated expertise in not only developing metagenomics algorithms, but also in translational microbiome research.

method has been validated using publicly available microbiome data pertaining to pregnant women (on over 1,100 clinically collected samples from diverse geographies and ethnicities). Since our microbiome-based diagnostic biomarker can accurately predict the risk of preterm birth as early as in the first trimester of pregnancy, clinicians will have enough time to act and prevent the preterm delivery.

Apart from predicting the risk of preterm delivery, the TCS team has successfully developed microbiome-based biomarkers that can be used for accurate screening of colorectal cancer and

breast cancer using microbiome sequence data obtained from an individual's stool sample. The biggest advantage of using our biomarkers is that these are non-invasive (unlike traditional clinical diagnostics/screening methods) as well as relatively inexpensive. They can also predict the risk of a number of asymptomatic diseases/disorders early. This gives doctors a chance to start timely therapeutic intervention and manage the patient and the disease in a much more effective manner.

Future of Personalized Healthcare

While linking imbalance in gut microbiome composition to diseases isn't new, TCS' Life Sciences Research team has pioneered in (1) microbial function-based assessment of gut health and (2) microbiome-based biomarker discovery for early risk assessment of diseases/disorders. With the technological advances and rapidly declining costs of DNA sequencing, it will be possible in the near future to routinely sequence gut microbial communities just from stool samples of individuals in a cost-effective manner. A simple analysis of this data using TCS' innovative methods will be able to provide clues on the health condition of the individuals. This will help in not only defining the best diet and nutrition plan for them, but will also indicate whether they need to undergo further medical investigations or interventions.

What makes for a healthy gut microbiome? We carried out collaborative work with the National Institute of Cholerae and Enteric Diseases (NICED, Kolkata) and the Translational Health Science and Technology Institute

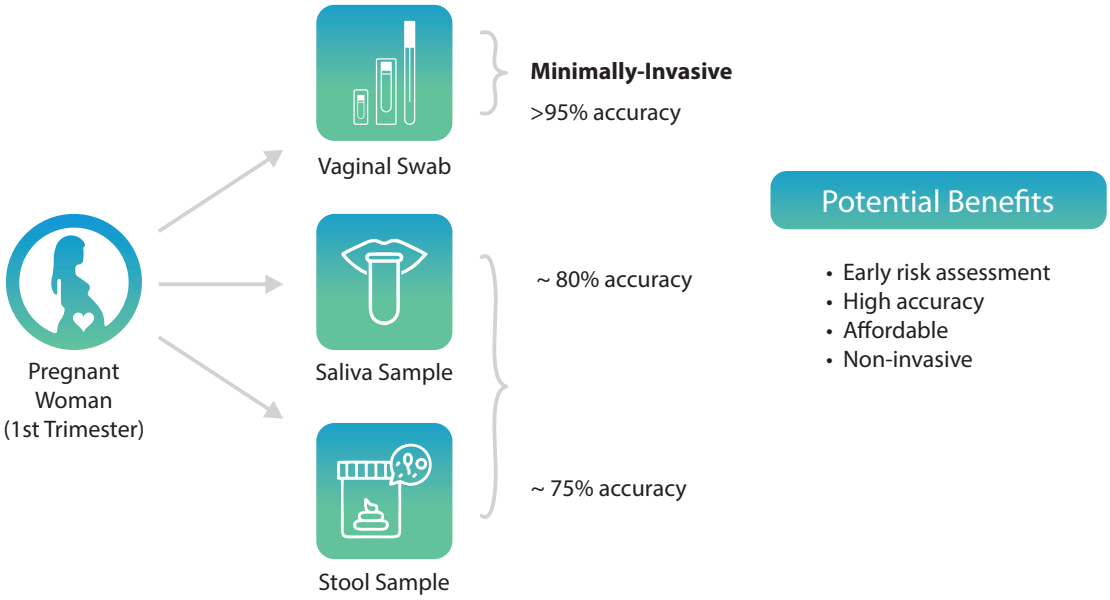


Figure 2: *Preterm Birth Diagnostics*

(THSTI, Faridabad) to understand the link between gut microbiome and malnourishment. Since the gut microbiome plays a crucial role in nutrient pre-processing, assimilation, and energy harvesting from food, one would expect dysbiosis of the gut microbial community in malnourished children. We analyzed microbial DNA extracted from faecal samples (which represent the gut microbiome) obtained from children from rural areas in West Bengal. We selected children with varying nutritional status, i.e. ranging from severe acute malnourished children at one end of the spectrum to apparently healthy children at the other end. An interesting question is this: "Regardless of living under similar conditions of hygiene and sanitation, why are some children malnourished while the others are apparently healthy?" Our study findings suggest that the indicative characteristics of the gut environment rest not only on the presence of specific microbial groups, but also on the inherent inter-microbial interactions present in it. The results showed that with decrease in the nutritional status, there is a possible increase in functional interdependence among the various microbes residing in the gut. As per the analysis, it is possible to modulate the gut microbiome by disrupting certain key players (microbes) so as to achieve conditions that could result in the regression of the malnourishment phenotype. In other words, appropriate nutritional regimes as well as new probiotic and prebiotic supplements can be developed for tackling the malnourishment problem.

Monitoring Gut Health: Toward a Personalized Diet and Therapy

Realizing that a healthy and balanced gut microbiota is one of the keys for ensuring good health, it would be ideal to be able to monitor and understand health of the gut (and its microbiome) periodically. It is to be noted that microbial communities may be very different in two healthy individuals from two different geographies, but they may perform similar biological functions. Thus, in order to assess "gut health", one needs to look at not just the microbial groups residing therein, but also their functional potential. TCS' Life Sciences Research group has explored specific functional signals from the microbiome for developing a method for assessing and quantifying gut health status. TCS' "gut health score" can be used to monitor an individual's gut health status just from collected stool samples.

How robust is the TCS gut health score in differentiating between a dysbiotic and healthy gut? We have validated our score using around 600 publicly available microbiome datasets corresponding to healthy and diseased samples. The diseases included Crohn's disease, inflammatory bowel disease (IBD), diabetes, asthma, colorectal cancer (CRC), non alcoholic fatty liver disease, Parkinson's disease, and cystic fibrosis. Our gut health score could indeed differentiate the disease samples from the healthy ones. The findings also revealed that this score could indicate disease conditions which affect not only the gut (e.g. in the case of IBD, Crohn's, CRC), but also distal organs (like the

The TCS team has successfully developed microbiome-based biomarkers that can be used for accurate screening of colorectal cancer and breast cancer

brain, lung, and liver). This suggests that one can utilize a simple stool-based test to assess gut health of an individual and identify the risk of various diseases and disorders. Going beyond this, identifying functions determinant of the gut health condition of an individual can provide insights into designing personalized therapeutics which may be in the form of prebiotics, probiotics, or nutritional supplements.

Sharmila S Mande

Sharmila S. Mande, a distinguished chief scientist at TCS Research, has 35 years of research experience in the Life Sciences domain. She received her Ph.D. degree in the year 1991 in Physics from the Indian Institute of Science (IISc), Bangalore. She had her postdoctoral training at the University of Groningen, The Netherlands, and the University of Washington, Seattle, USA. After returning to India, she continued her research at the Institute of Microbial Technology (IMTech) and Post Graduate Institute of Medical Education and Research (PGIMER) in Chandigarh, before joining TCS in 2001 to start TCS' Life Sciences R&D activities. She has published several papers in international journals and holds a number of patented algorithms and software solutions that address challenges faced by researchers in analyzing large-scale biological data.

AFTERWORD

Research and Innovation Architecture for Business 4.0

K Ananth Krishnan,
EVP and CTO, TCS

Computing is truly "the technology" for the new millennium. It is helping humanity hunt for subatomic particles and get insights from planets and stars. And of course, it is changing the way business is done, across the world and in every industry. And we technology watchers, who sit behind a futurescope, live in perpetual excitement of discovery—new technology stars that will make the world better; and catastrophic black holes we need to dodge. And from here begins the journey of a technology from a concept to a resource that can change the way we work, play, and think.

Futurescaping

The Business 4.0 era requires agility in organizations, demanding velocity in every phase of research and innovation (R&I). Identifying the technologies is a crucial first step in enabling this agility. Here is a set of techniques we have learnt on our R&I journey:

Using multiple antennae: TCS has an Insights and Foresights function within the Research and Innovation organization that scans the technology space using some primary and plenty of secondary research. Our academic partners give us an up-close view of nascent technologies. Our venture capital partners in the Emerging Technology COIN™ help us understand where they are placing their next big bets. Startup partners show us how much of new technology has actually hit the road in a useful way. Our Research and Innovation Advisory Board consisting of academics and technologists of global standing offer an outside-in view. TCS business leaders and customers share problems that need to be solved. These multiple resonances are like a "deep space network" that guides a space journey—we listen to the strong signals and pay attention to the weak ones too.

Studying digital exponentials and intersections: We have charted out the next technologies that might impact us in an exponential fashion. We look at technologies individually; and also at domain trends. The trick is to see exciting possibilities at intersections: intersections of technologies, intersections of domains, and the intersection of technologies and domains.

Observing mass adoption patterns: For every trend, we look at its propensity for mass adoption. We give proportionally more weightage to areas which are likely to proliferate, in both domains and markets. For instance, as Web 1.0 and Web 2.0 proliferated, a number of domain offerings and services rode on top of it; similarly, we saw the same with the adoption of mobile and wearables.

Applying a societal filter: The impact of societal behaviors has a counterbalancing effect on adoption and proliferation. Demand for data privacy, Net Zero goals, and the disruptions created by the COVID-19 pandemic are some examples of societal trends that course-correct trajectories for technology growth. These pathfinders enable TCS to decide where to pivot next in our technology exploration.

Reinforcing TCS' Innovation Architecture

TCS has built a set of Enterprise Capabilities that enable new technologies/ideas to move from research to the delivery of mature business offerings. We outlined our basic innovation architecture in our book *Research by Design* published a decade ago. We had, by then, created a strong foundation with the following:

A portfolio approach: We based this on Harvard professor Clayton Christensen's (a former independent director on the TCS Board) three-horizon segmentation of research investments: Derivative Innovation (improvement on current products and services), Platform Innovation (a move to adjacent technologies and markets), and Disruptive Innovation (game changers, completely new markets) across time horizons – current, 2+ and 5+ years respectively.

Open innovation: We decided to scale enterprise innovation to include partners in the innovation landscape—emerging tech companies, venture funds, academic research, entrepreneurs in residence, and our anchor customers with the TCS Co-Innovation Network (COIN)™ program (refer essay *Co-Innovation with Academia*).

Idea-to-innovation cycle: We created a well-defined ideas-progession path— from an "Explore" phase when resources are plied to create a research finding/algorithm, to the "Enable" phase where a prototype is engineered and piloted, to the "Exploit" phase where pilots with market traction are incubated to scale for commercialization. This trajectory is now governed by the New Products and Solutions Deployment (NPSD) framework (refer essay *Taking Research to Business*).

Innovation culture: We had instituted awards, contests, and crowdsourcing mechanisms to enable any associate in TCS innovate and be rewarded for creative thinking and work. These capabilities have been further reinforced with our 5V framework by which TCS R&I aims to be **Visionary**, to offer **Variety**, have **Velocity**, create **Value** and imbue the company with **Vibrancy**.

Setting the Vision

Our vision is to be a company that is noted for its technology R&I. TCS has, in fact, been listed in Forbes' top 100 Innovative Companies for many years. We want to target grand challenges, aim for bigger disruptions, and prepare for completely new technology paradigms that are emerging.

Our researchers, scientists, and technologists are publishing at research conferences and scientific journals, filing patents and winning awards across the world. They collaborate with our Academic Co-Innovation partners— leading global universities and technology institutions in a growing number of relationships and labs. R&I leaders are tasked with finding the next big bet to focus on. We have invested in foundational research areas that we believe will offer us a strategic advantage (*Preface*).

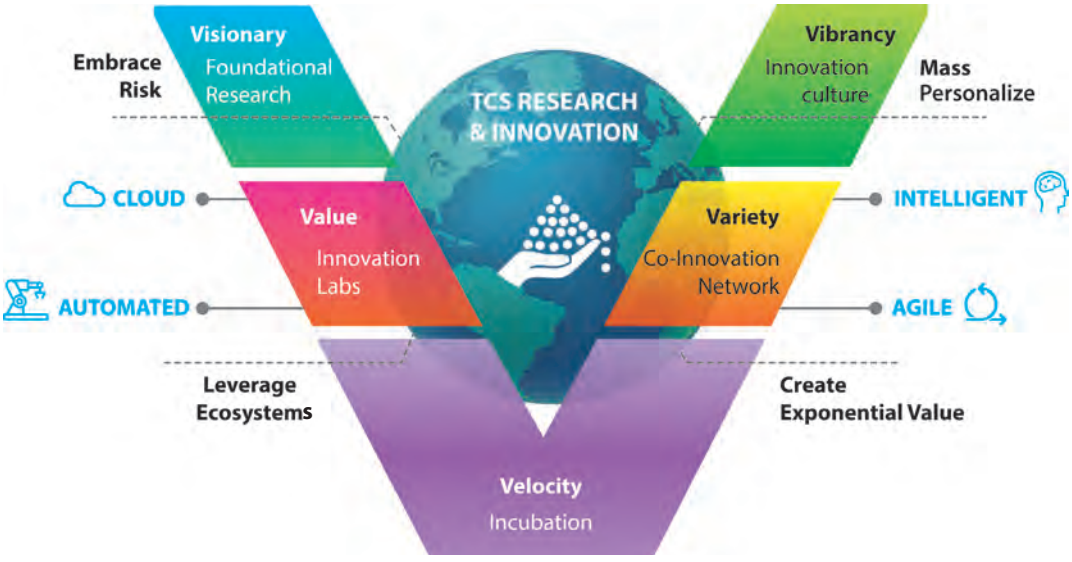


Figure 1: *The Five Vs of TCS Research and Innovation*

Dazzling with Variety

TCS serves twelve major industry sectors, and multiple sub-sectors. We serve our clients in 46 countries. Our requirements demand variety. TCS R&I has many sources that seed explorations.

Apart from the long-term research in foundational areas carried out by TCS Research, TCS draws in new technology ideas from its Co-Innovation Program (COIN) that works with academia (refer essay *Digital Citizen*), partner ecosystems (refer essay *Earth to Cloud*), emerging tech companies, venture funds, and customers themselves. The company also encourages ideas from the grassroots with its Innovista contests and hackathons (refer essay *Building an Innovation Culture*). The entrepreneurship program invites associates who have disruptive ideas and allows them to work like a start-up within the company.

Driving Velocity

Too often a good invention languishes in the lab and loses its early-mover advantage. The trick is to get it out of the lab to meet the business need quickly. We attempt to strike a balance in investment between innovation programs that have a time-bound delivery schedule and those that need long-term exploration. TCS R&I has put in a rigorous process to move inventions from the drawing board to delivery through a series of technological readiness reviews. We encourage teams to put forth a minimum viable product within a defined time period. A dedicated team incubates innovations that have market traction and scales these up for commercial deployment.

Periodic reviews with business units sharpen the focus of the offering and facilitate business sponsorship.

Innovations that are relevant here and now, as extensions of current solutions or with currently mature technologies, that we term Horizon 1 innovation, are performed by technology offices within our business units. This scale-out contextualizes innovation and enables speedy delivery.

The internal entrepreneurship program office vets ideas we receive from our large workforce, bets on the most promising ones, and fast-tracks them. The enterprise-wide hackathons that are conducted also produce MVPs that are evaluated for scale-up.

Generating Value

The most important function of the TCS Corporate R&I unit is to offer a continuous pipeline of new business opportunities. The technology readiness assessments, the incubation program, and the co-innovation program create a stream of new offerings that impacts the company's business, the company's internal efficiencies, or the company's social causes. Most of the essays in this volume are examples of inventions and innovations that have created value for the company. Hundreds of customers have been consumers of TCS Innovation and have testified to its effectiveness.

The R&I Evangelize team is an effective connecting channel that brings business ideas into research for exploration, and takes the best of R&I to the market. It goes beyond a sales and marketing team (refer essay *Evangelizing Research and Innovation*) and enables the creation of value for both research and business.

Creating a Vibrant Culture

Industrial research can never be confined to the ivory tower. It will have failed its mission in part if it does not infuse creativity within the organization. In an organization that is as large as ours, with 400,000+ associates, we feel creativity can come from anywhere. The TCS Innovista 2020 Competition had 6450+ entries. Coming from different units in the organization, this shows a healthy growth in innovative thinking through the company. Each of these entries is filed by a four-to-five-member team, who benefit from the mentoring, visibility, and the encouragement they receive from leadership teams.

Apart from events and contests, we have thoughtfully designed spaces that encourage design thinking, brainstorming, collaboration to infuse energy, and creativity to problem solving. The TCS PACE™ program will launch several such events in different geographies where TCS operates. TCS PACE™ Ports will be touch points for our customers and partners to experience and participate in TCS Innovation.

The Future

Early and continued investments in research, a method to pick the next technologies, and a comprehensive innovation architecture keep TCS Research and Innovation optimistic about the future. We are looking at a fast-changing landscape and a number of exciting options. At TCS 50, we hope to continue our voyage for another 50 years in cyberspace, powering our customers to race ahead of competition, offering the best to their consumers.

K Ananth Krishnan, EVP and CTO, TCS

Ananth directs Research and Innovation in TCS. Under his leadership, TCS has created a significant portfolio of patents, papers, and other intellectual property, delivering business value.

Ananth has served on several Governing Councils of academia, industry advisory boards, and government committees. He has been a regular invitee to the Board of TCS since 1999.

He was elected a Fellow of INAE in 2013. He was named a Distinguished Alumnus of IIT Delhi in 2009. He has been listed in Computerworld's Premier 100 IT Leaders (2007), and Infoworld's Top 25 CTOs (2007).

Ananth holds an M. Tech. degree in Computer Science and an M. Sc. in Physics from the Indian Institute of Technology, Delhi. He has a Bachelor's degree in Physics from Fergusson College, Pune.

Our thanks to ...

They say it takes a village…

This book has drawn on the strengths of many TCSers. Among them are:

Amit Chemburkar
Anupama Kasturi
Arvind Sreekumar
Beena Rai
Bhanu Raj
Devasena Ravisankar
Dr Pradip
Hari Hara Subramanian E.P
Jyothi B
Jyoti Bisht
Kanchi Usha Parthasarathy
Karan Gandhi
Kirubanandham N
Lakshmi Murali
Mohammed Monzoorul Haque
Murali Menon
Parthasarathi V
Preetha Rajiv
Prof Mathai Joseph
Ram Viswanathan
Ramasubramanian Suriyanarayanan
Santoshi Tiwari
Sarjinder Sethi
Saswati Collam
Shampa Sarkar
Shanmuga Priyan Viswanathan
Shivaraman G
Smita Ghaisas
Sougata Murmu
Suchita Rajesh Mantri
Sudipta Mandal
Suresh Kumar Ramasamy
Sutapa Singha
Tamanna Desai
Varsha S
Vinodhkanna Jayaraman

This Book would not have been possible without process support from:

Alamelumangai Mukundan
Anand Endur
Celine Rodrigues
Della Sajan
Gopal Sankaran
Kashmira Jijina
Manash Paul
Manjula K
Manjula MS
Manjunath Munivenkatappa
Saritha K
Seema Kukreja Kadam
Sriganga H
Srikanth Venkatesh
Srinivasan Mani Iyer
Usha Divakar
Vidhya Jawahar

Index

W

Research Careers

TCS Research and Innovation pays special attention to research students, by encouraging lab visits, sabbaticals, poster presentations and internships. TCS researchers have informal meet-ups with students from premier institutes to inspire a career in research. TCS supports Ph.D. researchers in India with a scholarship program.

About Tata Consultancy Services Ltd (TCS)

TCS' mission is to help customers achieve their business objectives by providing innovative, best-in-class consulting, IT solutions and services and to make it a joy for all stakeholders to work with us.

A part of the Tata group, India's largest multinational business group, TCS has over 443,600 of the world's best-trained consultants with offices in 40+countries. The company is listed on the BSE (formerly Bombay Stock Exchange) and the NSE (National Stock Exchange) in India. TCS' proactive stance on climate change and award winning work with communities across the world have earned it a place in leading sustainability indices such as the Dow Jones Sustainability Index (DJSI), MSCI Global Sustainability Index and the FTSE4Good Emerging Index. For more information, visit us at www.tcs.com/research.

Reimagining Research

In a fast paced world, TCS Research and Innovation (R&I) has to speed solutions using emerging technologies to solve customer problems. This book gives a sampling of such work and how it impacts industries, enterprises, and society; the transformations it brings to computing; and how it embeds intelligence in business processes. TCS' research outlook and innovation architecture discussed here, highlight our reimagined approach to research.

TCS has invested in Research and Innovation since 1981. From creating early CASE tools and compilers to current AI based algorithms, TCS R&I has been in the forefront of technology innovation. A steady stream of research output finds its way into commercial offerings that bring business value to TCS.

The essays in this book are by our researchers who are experts in technology and have acquired considerable domain knowledge. TCS researchers are passionate about their work while creating sizable intellectual property for TCS. We hope you, our reader, will be able to share their excitement and spirit of enquiry reflected in these essays.