pARIvartana

The Role of Data and Analytics in Social Transformation

TATA CONSULTANCY SERVICES

Analytics and Insights | Tata Consultancy Services

pArIvartana

The Role of Data and Analytics in Social Transformation

C O N T E N T S

👪 PUBLIC HEALTH & WELL-BEING

📰 POLICY & GOVERNANCE

📊 **GREEN ECONOMY**

☁️ **TECHNOLOGY & INNOVATION**

Dinanath Kholkar
Senior Vice President, Global Head of Partner Ecosystems & Alliances
Tata Consultancy Services

Preface

On behalf of the Analytics & Insights (A&I) team, it is my honor to announce this special publication **pArIvartana**, which is not merely a compilation of articles but a shining example that showcases how a diverse set of highly motivated experts including CxOs, industry leaders, economists, academicians, researchers, social entrepreneurs, philanthropists have made a positive, and sustainable impact to the community. My sincerest thanks to each one of them for not only sharing their success stories with us but also for sharing their organization's vision to make this world a better place for us and generations to follow.

Over the last 12 months, we have been celebrating two important milestones namely the 150th year of the Tata Group and the 50th year of Tata Consultancy Services (TCS). To make this memorable, the TCS A&I unit had decided to come up with a commemorative publication on the "Role of Data & Analytics for Social Transformation".

Ever since the inception of the Tata Group, all our organizations have been committed to social responsibility. In the words of the Group's Founder, Jamsetji Tata, the community is not just another stakeholder in business, but is, in fact, the very purpose of its existence.

The concept of social transformation refers to the change of society's systemic characteristics that involve technological, economic, political and cultural restructuring. One of the things I have observed is that the social sector and related initiatives depend heavily on data and analytics for their success.

Organizations that make a profound social impact are often those that have realized the power of data and adopted a data centric approach. Also, with Artificial Intelligence (AI) becoming main-stream across various sectors, the same data will play a significant role in the adoption of AI techniques and accelerating the transformation. I am sure, pArIvartana will inspire many of you to volunteer and contribute on various community initiatives by using their core competencies in data, analytics and AI.

At TCS, we are committed in promoting a culture of ensuring safety and protection of environment, and enhancing the lives of this planet's inhabitants by making the most of emerging technologies.

I would like to express my gratitude towards my colleagues in the TCS Analytics and Insights unit for the effort and extra time they have put in to shape this up. Special thanks to Jagdish Chaudhari and Avinash Pattabhiraman for managing this project from concept to execution.

Enjoy the report and, as always, I encourage you to share your opinions on how we could come together and work for a better tomorrow.

Best Regards,

Dinanath Kholkar
Vice President and Global Head, TCS Analytics and Insights

In Conversation with John Kahan

John Kahan
President Aaron Matthew SIDS Research Guild of Seattle Children's Hospital
and Chief Data Analytics Officer, Microsoft, AI for Good.

Sudden Infant Death Syndrome is one of the leading causes of death of infants. John Kahan talks about his plans to address this challenge

How did you get started with the Sudden Infant Death Syndrome mission?

Sudden Infant Death Syndrome, or SIDS, is the leading cause of death of infants between the ages of one month to one year old in the US and in all developed countries in the world. My wife, Heather, and I lost our fourth child, Aaron Matthew, to SIDS.

Under the leadership of Dr. Nino Ramirez, Director, Center for Integrative Brain Research at Seattle Children's Hospital, we have assembled a team of top SIDS researchers in the world and a world-class team of Microsoft data scientists from Microsoft, where I serve as the Chief Data Analytics Officer for our "AI for good" efforts. Tata Consultancy Services (TCS) has also dived in to help with much-needed IT support and genomics researchers.

Together, we are taking a five-pronged approach to solving SIDS. Our team has research going on across genetics, epidemiology, physiology, pathology, and education/outreach, which means we are uniquely equipped to collaboratively address the most difficult outstanding questions in the field from multiple angles.

The most exciting and newest part for us and the world, is to build the first infant genome database ever in the cloud, focused on SIDS, and ultimately on infant mortality. SIDS is the name people understand, which basically means that we just did an autopsy on a child, and we don't know why the child died. What we are trying to do is to reduce overall infant mortality by building a genome database, based on the data science and research the teams are delivering. We believe there are likely biological markers in the genes of children who ultimately died of SIDS. Our vision is to build the first infant genome database so that parents can do a sequencing test of themselves and potential siblings of the child, determine risk factors, and develop preventative measures for a syndrome that haunts all parents in the world today.

What are some of the typical challenges you face in your SIDS Research work and how do you address them?

The first challenge is to check if the problem is clear, and this is a very difficult challenge indeed. In the United States alone, ~60,000 children have died since Aaron died 15 years ago. It's equivalent to two busloads of children disappearing every week for the last 15 years, without knowing why. So the first challenge is the biggest challenge of all, and it has gone on for thousands of years.

The second challenge is of data, and it's one we have all the time. We were very fortunate, as the Center for Disease Control in the U.S. collects data on every birth and death. That's 20+ million births, and 20,000 deaths from SIDS related deaths in the United States. So they had a database, which is great, and it helps us a lot. However, it's not enough, as it doesn't have the biological component of each child. Moreover, a great deal of the data that was collected from autopsies has been collected in an inconsistent fashion, leading to a lot of what we call "noise." This requires plenty of work to clean up the data. In addition, states within the U.S. and elsewhere collect and protect data differently. They also apply different rules, making it difficult to combine the data to understand the challenges. This is why building a genomic database without human bias to further help us understand the challenges with less human judgment, while machine learning will give us better understanding of what is really happening with these children.

Lastly, the biggest issue of all is that such an initiative requires a large amount of resources. I am very blessed that Tata and Microsoft have partnered together with Seattle Children's to help address this challenge. But it will take a sustained level of effort from the public, the private sector, and from the Government, and it will take policies to help us unlock data and clues to solve this problem once and for all.

Who are the stakeholders responsible for bringing the data to the table for the research?

This is a combination of policymakers and the private sector working together. There are many laws in the U.S., and around the world, that prevent the collection of data or lead to inconsistent collection. We need to change those policies. We are working with the representatives in the US government to pass legislation to unlock this ability in the United States, which will greatly open up data for research. That would be one major step forward. We also need to work with medical institutions around the world to encourage parents to agree to an autopsy, or donate their child's tissue samples for genomic sequencing, in case, God forbid, they lose their child. It's very hard to approach grieving parents to help save the next child. It needs to happen earlier in the pregnancy. We also need to work with governments across the world to gather data for the cause of death of children who have died of similar SIDS related causes in a much more consistent way. Data should be shared in an anonymous way to ensure individual privacy. That is the other challenge that we face. We need to build ironclad anonymization into our data so that people and governments around the world feel safe to donate the

information about their children who have passed away. So, we do have some technological challenges, but I am convinced more than ever when you take some of the world's leading companies like Microsoft and TCS together and the world class team of researchers and data scientists who have donated 1000's of hours of time, we can solve this once and for all.

It is important to ensure the privacy and security of data. How would you do it?

I cannot agree more. We must protect the privacy and security of all data here. We can, and in fact this is something Microsoft and other tech companies have gotten pretty good at doing. What's called "Differential Privacy" technology allows researchers to collect data for research purposes in a privacy-compliant manner without identifying individuals. It allows us to do the right research to solve these problems once and for all. We need to expand this capability on a mass scale and make it an open platform to allow everyone to contribute in a safe, compliant way.

The good news is that Microsoft Genomics has built a capability in the cloud that gives us the ability to to protect the data we have today and as we further build out the first genomics database in the cloud focused on SIDS. Using Differential Privacy and techniques such as Confidential computing will allow us to go further and collect more data on a country, state, and location level to get more clues to solve the problem, while protecting individual privacy.

How do you see this project progressing over the next three years?

The great news is, as technology continues to grow at an exponential rate, costs drop, opening up possibilities and capabilities that we could not have imagined before. Just a year ago, the cost of sequencing a child's genome was about $4,000. Today, at our scale it's about $1,700. That's an incredible reduction in cost. As we continue to see more advances in data science and medical research, it will bring down the cost of sequencing even more.

Second, machine learning and AI capabilities are growing astronomically. As a result, our ability to understand large scale data sets is much greater today. Today, with the power of Microsoft Azure in the cloud, AI and machine learning has given us much more capability than we have ever had before and it is growing every day. A short example: when we approached Microsoft Genomics and asked them for help on our genome sequencing work, they asked us for the size of our dataset. We said that we wanted to sequence 256 children, which would have taken us many hours of processing time in servers in Seattle Children's. Now this could be done in the cloud in only an hour or so. As technology improves over the next few years, we will gain the ability to operate faster, and at scale. With partnerships between private and public sectors worldwide we will break down antiquated barriers further. And with technology protecting individual privacy, we will be able to stop terrible diseases in their tracks – not just SIDS – to dramatically reduce infant mortality.

Thus far, we have talked about using technology to solve the SIDS problem from a medical perspective. How about the other dimension - using technology to educate parents?

Our focus has been on research, to understand the problem. Partly because we can't educate parents and take steps to prevent it if we don't understand the problem. But we need to educate parents on the risk factors for SIDS such as the child's sleeping position. Putting a child on their back in the crib and removing soft materials around the child have been shown to reduce the risk. We also need to educate them on other risk factors such as, smoking. One hundred percent of all funding and donations to the Aaron Matthew SIDS Research Guild of Seattle Children Hospital goes to research. My wife and I have fully underwritten the guild and many volunteers have come to help from Marriott, to Adobe, to Accenture, to Microsoft, to Y&R, and Tata Consulting to name a few. We have also been blessed that Cribs for Kids, which is a National charity in this space in the United States, has formed The Aaron Matthew Research Foundation to help fund our research efforts. There focus is on safe sleeping as this is the one of the major known preventative measures a parent can take to reduce

the risks of a SIDS death. Here we are using AI to enable parents' understanding of SIDS and to take action before there is an issue. We have some early prototypes in the cloud and on the edge that use cameras and sensors to better understand safe sleeping positions and alert parents if their child is at risk. Our volunteers are investing their personal efforts on building an app for educating parents on SIDS in conjunction with Robert Wood Johnson Medical School/SIDS Project of NJ. So my personal focus is around Research and that's what we are doing with Seattle Children's, Microsoft, TCS and the top medical researchers and world class data scientists in the world today. But there are others in this matrix who are working on the education space and wherever we can, we will help with our efforts as well.

Can you describe your personal experience in pursuing this transformative journey?

My son, Aaron Matthew Kahan, who died of SIDS, has been the inspiration behind this transformational project. About two years ago, I started on a "normal" crusade to try to give back. I climbed Mt. Kilimanjaro in an effort to donate my time and raise money for SIDS research. I had no intention, or at least no knowledge or thought process of joining the journey of medical researchers, along with our data scientists, to help solve the SIDS problem. A series of incredible events led to the birth of the SIDS mission that we are working on today.

I was very lucky that my lead data scientist at the time had seen a picture of my son, Aaron on my desk. I told him that Aaron had actually died around 12 and a half years before and explained what it was like. He was shocked and was scared for the child that was about to be born in his family, and he learned a lot. He said that he didn't just want to donate money, but wanted to donate his skills, if there was anything that he could do to help. This small meeting unlocked researchers and data scientists worldwide to begin this journey together. That research pointed us to potential genetic markers and led us to begin building the first genetic database ever in the cloud focused on SIDS for researchers to collaborate worldwide. I'd like to say that my son is what led me here. In fact, the thing that gives me chills behind my spine is when we determined, based on our research, that this is the direction to head, Seattle Children's and our data scientists came to me and asked if I would help them build the genomic database and whether I would donate Aaron's tissue samples and sequence to it. I told them I would be glad to donate it, but Aaron had died 13 and a half years earlier and that I didn't have tissue samples or any kind of DNA. They told me that, based on medical and future research, they had been smart enough to put away in an anonymous fashion tissue samples of Aaron. I am very proud that Aaron Matthew Kahan, my son, is the first child in that database and he will gave us the first clues that I am sure will help stop SIDS for the next generation of children and parents in the world. And so that's how I began my journey and I am sure we will solve this, once and for all.

I am forever thankful and appreciative of the donations that people provided. But most important, I am thankful for the time people have invested in this research. I am blessed that leading data scientists from all over the world have donated thousands of hours to help with this journey. They set out to partner with top SIDS medical researchers. I am very thankful to all the data scientists, medical professionals, Microsoft, and Tata Consultancy Services for creating this unprecedented partnership to reduce infant mortality. TCS has given one of our largest donations and engaged every day in everything from fundraising, to research to data science, to help solve this. I can't thank you enough. You have joined us on a mission because we are kindred spirits. This research will not just solve the SIDS problem, but it will ultimately improve the mortality of our next generation of children. That's why we are on this journey together. A huge thank you to TCS and the Tata Group!

To learn more, and to donate much-needed funds to the cause of ending SIDS, please visit *http://www.givetostopsids.org/*

ABOUT JOHN KAHAN

President Aaron Matthew SIDS Research Guild of Seattle Children's Hospital & Chief Data and Analytics Officer, Microsoft, AI for Good, responsible for infusing data science to address the world's great challenges. e.g., promoting sustainable use of the planet's resources, improving opportunities for people with disabilities, protecting human rights, strengthening humanitarian assistance in disaster response, needs of children, refugees and displaced people, human rights and increasing the capabilities of the world's NGOs. Prior to this role, John served as the General Manager of Customer Data and Analytics at Microsoft.

PUBLIC HEALTH & WELL-BEING

Transforming Healthcare and the Society, the Big Data Way

Girish Krishnamurthy

Product Management, Healthcare Products and Platform (HPP) Business Unit, Tata Consultancy Services

A transformative approach, by its very definition, assumes the advent of something totally new rising from the ashes of the pervious form. Social transformations are inherently difficult to achieve due to people's innate resistance to change and the need to impact large population sets while overcoming slow adoption rates. However, a potent combination of processes, key technologies and skilled resources are impacting our world today, helping us solve long standing, complex, large scale social problems.

Healthcare presents some of the most exciting, and yet, challenging avenues for data-driven transformation. Data can directly and quantifiably help save human lives, improve quality of life, optimize cost of treatment, and predict duration of treatment as well as outcomes. It can also help reduce medical errors and hospitalization costs, resulting in wide-spread socio-economic impact of far greater magnitude than in any other industry.

Using healthcare data analytics to drive transformation

Today, healthcare data is spilling out of HIS and EMR systems, mHealth and pervasive wearables, health sensors, social media, and genomics. However, most of the clinically rich data generated today disappears into a massive 'black hole' that fails to provide any meaningful insights. It is estimated that by 2020, healthcare data will reach 25,000 petabytes, a 50-fold increase from 2012.

Data analytics and data science offer a plethora of opportunities in enabling social transformation in healthcare. Here's how:

- Analysis of data streams from rural and remote areas can offer insights to enable virtual diagnosis, care and treatment - at scale. The result: improved community access to healthcare, optimized time utilization, increased social comfort, and minimized travel for patients.

- Remote monitoring and care of elders and at risk patients in a sensor-enabled environment helps communities live with more dignity and freedom.

- Large area analytics of drug adoption data in a particular geography and for particular diseases can be instrumental in improving drug research and efficacy.

- Preventive advice and care based on insights derived from historical records of patients allows individuals to gain better control of their own health.

- Care enablement analytics is useful in measuring the social impact of decongesting critical care facilities through process optimization and care arbitrage.

> From remote monitoring to patient care, improved drug delivery, and cost-effective access to healthcare – Big Data and analytics can help reimagine every part of the value chain. The result: Creation of a connected digital loop that improves, optimizes, and innovates continuously

- Policy impact monitoring that pairs healthcare experts with data science and platform owners to help drive meaningful insights from large scale implementations.

- Personalized healthcare delivery based on macro and micro data analytics of PHR and EMR data offers immense potential to transform the way hospitals interface with the patient community, resulting in a more advanced care model.

- Demand prediction of healthcare facility needs, both for care and operations, can help providers better manage supply needs and predict demand trends and patterns.

While social entrepreneurs and forward looking organizations are already pursuing many of the opportunities, the key to success in such transformations rests on the active participation of all stakeholders – the state, the private sector as well as the not for profit organizations - working together for the community.

At TCS, reimaging the healthcare sector through modern technologies lies at the heart of our engagements. For instance, one of our group companies, in partnership with the Tata Trust, is deploying the Digital Nerve Centre (DiNC) - an innovative platform that leverages digital technologies to connect leading cancer research centers and specialists within the National Cancer Grid. DiNC aims to reimagine patients' access to cancer care and make it more equitable and affordable.

How TCS' DiNC reinvents the healthcare loop

DiNC comprises digital elements, which connect patients and doctors via a centralized platform. It's manned by medical experts who then manage the patient journey across the treatment cycle. Coupled with several on-ground transformation processes at hospitals using technology initiatives such as Shine, HealthX, Clinical Data repository, Interactive Kiosks, Single Window Exit counters, the platform aims to usher in holistic transformation. It continuously collects rich structured and unstructured data points from varied sources, and intelligently analyzes and presents the insights in a meaningful Clinicograph that captures the digital footprint of a patient's journey.

DiNC platform also comprises an event analytics module that analyzes each event in the patient journey, using the corresponding patient health record, and maps it to the central source. The analysis points out anomalies in the patient healthcare events over time, leading to transformative outcomes.

Connecting the healthcare value chain for futuristic operations

Deploying a 'curating-middle-layer' approach, such as the DiNC, can provide superior outcomes in several key healthcare scenarios: from primary to tertiary care, remote to in-campus, and user to specialist focus. Since hospitals operate on an always-on schedule, an unending stream of data is available for analysis at a scale unimaginable even a few years ago. Insights based on this vast array of data can enable personalized health plans, faster and easier access to care, development of better drugs, and cost-effective healthcare operations. It is clear that the need of the hour is to overhaul the interaction dynamics of the society with that of a hospital system using emerging technologies. Doing so can create a more compassionate, insightful, and future-proof healthcare system.

ABOUT THE AUTHOR

Girish Krishnamurthy, as part of the Tata Group is the Head of Strategic Initiatives and Healthcare Innovations. He is a transformational leader in the business of consulting, technology and services. Having an entrepreneurial spirit, he is known to forge incredible culture of learning, innovation, and collaboration for people and organizations to grow.

Girish is the recipient of the Change Agent Award by CRN, in recognition of his outstanding work in transforming the business ecosystem through redefined approaches. Holding a Bachelors in Electrical Engineering and a Masters in Mathematics from University of Madras, Girish is an MBA from the University of Texas.

The Role of Data and Analytics in Boosting Business Outcomes and Quality of Life

Dr. Dhiren Patel

Professor of Computer Engineering and Director of Veermata Jijabai Technological Institute

Data analytics and AI are driving transformation in business as well as societies by reimagining everything - from communications and computing to medicine, manufacturing, agriculture, transportation, and more.

After completing my Masters in CSE from Indian Institute of Technology (IIT), Kanpur, I joined National Institute of Technology, Surat as a faculty where I pioneered an undergraduate program in Computer Engineering and started teaching core CSE courses. Inspired by a short-term training program (STTP) on Artificial Intelligence (AI) and neural networks that I attended at IIT Bombay in 1993, I took up the challenge of offering an elective course in AI to the early batches of computer engineering students at the college. I still remember teaching this elective course using a book titled Artificial Intelligence by Elaine Rich. The book focused on creating interest in Automation Systems with examples of Eliza (Natural language processing), Deep Blue (Chess Playing Computer), A* and AO* algorithms (AND-OR search graphs and reduction algorithms). As part of the course, I also covered potential expert system use cases for addressing primary health care in rural India, industrial robotics, artificial neural networks, heuristics, and decision support systems. Frankly, I could not generate enough interest amongst the students. After a couple of years, we had to drop the course due to limited demand in AI, and instead, cater to a growing interest in internet technologies and applications.

Challenges at the time were centered on understanding and utilizing the potential of networks, the World Wide Web, high performance computing and DBMS/ERP, and verification of real time systems. In 1997, I began teaching an Information Security course at REC Surat, and as we progressed, we noticed that data science and data analytics were witnessing a steady upswing in popularity. We saw a major shift in methods based on learning data representations, as opposed to task-specific algorithms. The importance of data science and analytics in detecting and classifying malware signatures and ensuring cyber security became increasingly evident.

Security best practices are routinely ignored by users. Despite laying down password management best practices, most users typically use the same password for all logins due to the inconvenience of remembering multiple passwords. Ignoring the basic best practices greatly increases the security vulnerability of corporate networks everywhere. So how can organizations balance security with user convenience?

The approach taken by a major US bank is worth mentioning here: it places the emphasis on user convenience. In order to ensure security without inconveniencing users, the bank employs a program that seeds, monitors and learns behavioral footprints of individual customers, and detects anomalies or deviations from set patterns for near real time fraud detection and prevention. The origins of this smart program lie in machine learning (ML) and historical data analytics, and the initiative has worked well with minimal false positives.

It is clear from this example that a positive approach to enforcing a security system is one that rewards those who follow policies, instead of punishing those who do not. The framework and policies must provide guidelines for the consistent and secure management of passwords for employees that are comprehensive, practical, and effective for the organization.

Gradually, the importance of pattern matching and classification in intrusion and fraud detection, forensics, and trust and reputation systems came to the forefront. The decades old capability of software to simulate a large array of neurons in an artificial 'neural network', has led to as many disappointments as breakthroughs. However, because of the availability of Big Data and advancements in analytics tools and computing, scientists can now model many more layers of virtual neurons than ever before. Machines can now recognize objects and translate speech in real time. Today, the software learns - in a very real sense to recognize patterns in digital representations of sounds, images, and other data. Indeed, machine intelligence has started to transform everything from communications and computing to medicine, manufacturing, and transportation.

From self-driving connected cars to connected aircrafts, ML and Deep Learning (DL) have reignited the possibility of addressing some of the biggest challenges in AI - beyond image recognition, search, and natural-language understanding. With a greater understanding of data science, AI and ML; the availability of huge amounts of data (available through crowd sourcing, social networks and IoT); and the support of high performance computing and matured cyber security practices; we can arrive at a variety of use cases to solve many fundamental problems plaguing the society and the country.

Today's national leaders have greater awareness and confidence in these technologies and are challenging knowledge leaders and services companies to move beyond e-Governance and venture into areas such as precision agriculture and food security, healthcare, mobility and public transport, chatbots-enabled smart help desks, environment control, disaster aversion and management, and more.

Potential uses of deep learning in machine vision are technologies that use imaging for applications such as industrial inspection and robot guidance. We see personal sensors being leveraged by deep neural networks to predict medical problems. Sensors throughout a city are feeding deep-learning systems that can predict where traffic jams might occur. Data analytics and AI are indeed transforming not only businesses but also the society, inspiring smart cities using simplified and usable self-learning systems and self-care services such as sewerage disposal, waste management, water recycling, air quality management, noise reduction, traffic and signal management, and early warning systems.

ABOUT THE AUTHOR

Dr. Dhiren Patel is the Director and Professor of Computer Engineering at Veermata Jijabai Technological Institute (VJTI), Mumbai. His research interests span across areas such as cloud security, AI, blockchain, IoT, large scale identity management systems, and developing technology for society. Prior to joining VJTI, Dr. Patel worked as a Professor of Computer Engineering Department at NIT Surat, Gujarat. Dr. Patel has academic and research associations with several prestigious global educational institutions and companies: IITs, NITs, C-DAC, MeitY, MHRD and DST of India, City University London, University of Denver, Oslo University, HAW Hamburg, Meiji University, and TCS. He is an alumnus of Indian Institute of Technology, Kanpur.

Statistical Validation of Census Data: Why it is the Need of the Hour

Hemant Joshi,
Head of the Data Office, Tata Consultancy Services

Census data is a critical piece of information that determines how micro and macro level government decisions are made. Leveraging statistical techniques to cross-reference various census data sets is essential to improve the quality and accuracy of data, and ultimately the life of all citizens.

Across the globe, population data is key to a number of decisions. From formulating government policy to making budgetary allocations, estimating market size, and choosing the location for an ATM, a micro and macro level understanding of the population is crucial.

However, ensuring accuracy of census data is easier said than done. Globally, census is carried out via a survey in which citizens are asked to respond to a set of questions. Apart from the personal information about the individuals, the government usually collects information pertaining to religion, caste, education and income, which in turn is used to fine-tune policies. For any country, census is a large undertaking that combines meticulous planning, recruitment and training of a large workforce with coordinated execution to cover every corner of the country on the same day. The members of the population also need to be trained to answer the questions appropriately to avoid under counting or double counting - especially people unavailable at their current location due to travel for work, leisure, or other emergencies. During the 2011 India census, data was collected from 608,786 villages, 7742 towns, and 5767 tehsils across 640 districts, by nearly 21 Lakh enumerators.

There are numerous challenges to such a massive exercise, the key ones being:

- People on the move on the day of counting.

- Difficulty in reaching each individual in every corner of the country.

- Incorrect information provided due to distrust in the government associated with social stigma around disability, transgender population, woman's marital status, etc. or manipulation for social gains.

- Inefficiencies or mistakes by enumerators.

Data anomalies in the 2011 census

A number of initiatives were taken during the 2011 census to improve the quality of information capture. Some significant steps taken included: templates with bar codes for improved data capture, Intelligent Character Recognition (ICR) software for automated data tabulation and de-duplication, as well as extensive training of enumerators. Despite these efforts, a number of issues were detected at both micro and macro levels. In Nagaland, for example, the growth in the population between 2001 and 2011 was tabulated as 0.5% as opposed to the all India growth rate of 18%. This does not match the known fertility rates or population migration statistics of the region.

Another challenge is delineating the population by age. This data when plotted shows a peculiar pattern of peaks at five-year intervals for every state. The same pattern was noticed in the 2001 census as well. This does not hold water as it would mean a sudden, inexplicable increase in childbirth every five years *(see Figures 1-4).*
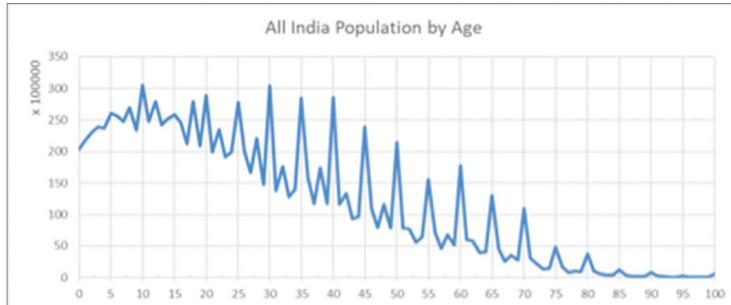
Figure 1: All India Population by Age in 2011
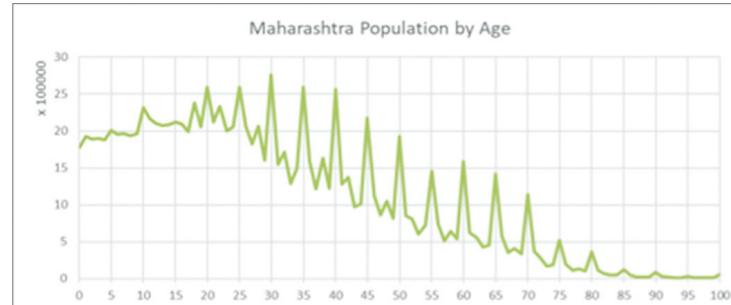
Figure 2: Maharashtra Population by Age in 2011
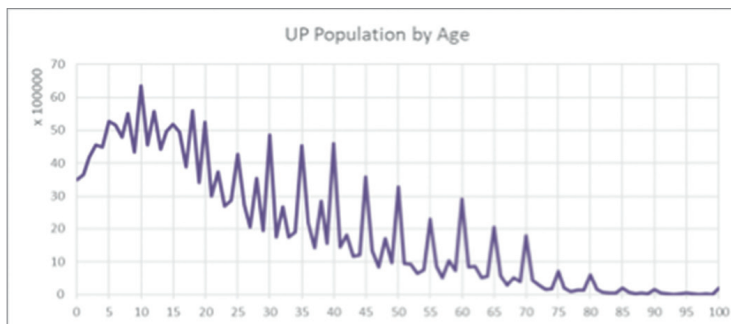
Figure 3: Uttar Pradesh Population by Age in 2011
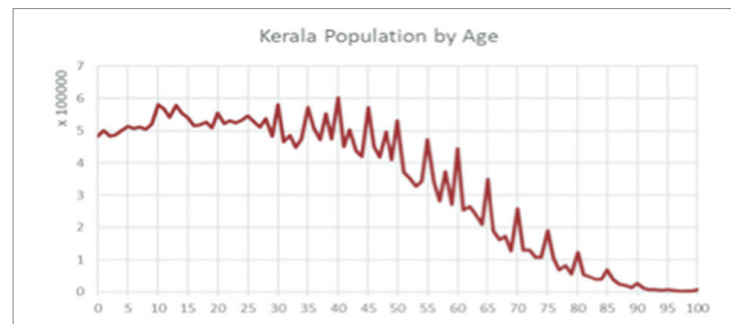
Figure 4: Kerala Population by Age in 2011

Moreover, in many districts, the child population data does not match the student population data collected from Unified District Information System for Education (UDISE), with the student population reported being significantly higher (over 30%) for many districts *(see Tables 1 and 2).* It is clear that at the state as well as district level, there are numerous instances of the population being lower than the enrollment. This is a significant anomaly that needs to be corrected given the importance of the data in the overall budget allocation and other policy decisions.

Ensuring better data for better planning and results

It is clear that there is a need to include statistical techniques to cross-reference various data sets that the government collects with respect to the population to ensure that they all match up correctly. Though the efforts to capture accurate data during census survey are crucial, they need to be augmented through validation of summary information using data collected through other initiatives such as birth and death registry, student population collected via UDISE, employment data, and population migration data within and across states. Government of India's Aadhaar initiative offers significant population coverage and can provide good statistics that can be used for validation. Improving the quality and accuracy of the census information using auxiliary data sources will ultimately help ensure that the policies of the governments are formed based on accurate data inputs and the benefits reach the intended audiences.

% Enrollment	Total School Enrl Std 1	Population Age 6	% Enrollment
Andhra	1625351	1429958	114%
Arunachal	77016	34963	220%
Assam	1009141	701025	144%
Bihar	3349171	3309454	101%
Chandigarh	17298	18423	94%
Delhi	345434	308608	112%
Haryana	495054	498816	99%
Jharkhand	1121457	879796	127%
Maharashtra	2034050	1961161	104%
Punjab	550796	464594	119%
Rajsthan	1996752	1647153	121%
Tamilnadu	1194702	1040265	115%
West Bengal	2498292	1570299	159%

State	District Name	School Enrl	Population	% Enrollment
J & K	Rajouri	21563	16133	134%
Punjab	Kapurthala	17507	13392	131%
Utncl	Rudraprayag	6889	5122	134%
Rajsthan	Jaisalmer	32545	21117	154%
U P	Bijnor	135134	93450	145%
U P	Lucknow	144039	92775	155%
A P	West Kameng	2636	1776	148%
Mizoram	Aizawl	11158	7299	153%
Assam	Kokrajhar	58170	32870	177%
W B	Hugli	102237	76997	133%
J K	Ranchi	108435	80481	135%
Odisha	Mayurbhanj	75170	57188	131%
M P	Indore	83192	63934	130%
A P	Warangal	82082	55698	147%
T N	Thoothukkudi	31464	20466	154%

Table 1: State wise enrollment % for Std. 1 in 2011

Table 2: District wise enrollment % for Std. 1 in 2011

ABOUT THE AUTHOR

Hemant Joshi is the Head of the Data Office for TCS. He is part of the Analytics & Insights (A&I) unit that leverages Data Management and Big Data capabilities to extract insights from the operational data captured within TCS. With over 21 years of experience in the IT industry, Hemant has been an architect and program manager for BI engagements. He has also worked in the CTO organization developing data and analytics-based products.

Changing the Game: Using a Data-Driven Approach to Influence the Sports Culture in India

Arnav Kholkar
Student MMU, Intern KreedOn

Fostering a culture of sports by bringing together sports teams, clubs, governing bodies, corporates, and coaches/ players can help alleviate some of the pressing challenges facing India's youth. Employing a data-driven approach to instill love for sports can help beat lack of fitness, motivation, mental health and inclination to crime.

Sports are an integral part of people's lives and their wellbeing. They connect, empower, inspire and motivate individuals, enable work-life balance, and help shape national character. The ability of sports to drive economic growth, and social and cultural cohesion, has brought communities closer together for years. Those participating in regular sporting activities show greater cognitive functioning and academic performance, as well as the ability to make the right choices in life. On the other hand, those that don't, suffer from poor motivation productivity and fitness. When coupled with the socio-economic scenario of a developing country, lack of interest in sports can push people into negative lifestyles of antisocial behaviour and crime.

A shift in perspective – towards building a holistic understanding of the importance of sport - can help transform individuals as well as the society. A survey report by PwC shows that Indian kids have a long way to go in terms of building body strength, flexibility and so on *(see Figure 1)*. 66% of those surveyed strongly agreed that there should be more health sports programs targeting the 18-64 year age group in the country.

Ushering in social transformation with sports

In order to drive social transformation, it is important to tackle issues such as reducing stress levels, increasing productivity, and reversing the atrophy caused by unhealthy lifestyles. This requires:

Targeting the easy to influence segment – the youth. (27.8% of Indians are between the ages of 0-14).

- Teaching people skills that will motivate them to climb the socio-economic ladder of society. (In India, 87.67 million working adults have a wealth lesser than INR 5 lacs)

- Reducing negative triggers like drugs, violence, lack of hygiene. (Every day 10 drug related suicides take place among the youth (16-28 years. The country has also seen a rise in crime rates, by an aggregate of 7% since 2006).

- Here's where sports can be used as a tool to empower the youth to improve the general well-being of Indian society. According to Statista's (2018) survey of 350,000 Indians over the age of 18, 12% played cricket and 7% badminton. But people who do not partake in sport polled a shocking third placed - at 6.5%.

- Addressing a sticky wicket: Challenges in fostering a sports culture

1 in 3 kids don't have a healthy Body Mass Index (BMI)

2 in 5 kids lack adequate lower body strength

1 in 3 kids don't possess adequate upper body strength

1 in 4 kids don't have desired flexibility

1 in 3 kids don't have desired sprint capacity
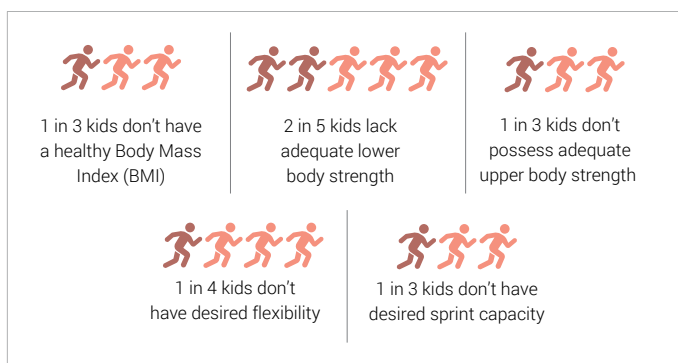
- Several challenges stand in the way of fostering a sports culture in the country that is already influenced by the lack of adequate support from authorities, poor performance in global competitions, and lack of trained professionals and sports curriculum:

Identifying areas with higher concentration of youth and collecting socio-economic, psychographic and geodemographic data from such areas for segmentation. The task is made more difficult by the fact that in four of the top five most populous states – Uttar Pradesh, Bihar, West Bengal and Madhya Pradesh - 70% of the population resides in rural areas.

- Gathering sustainability reports from the local governing bodies as sustainability is a key driver of motivation.

- Identifying funding sources – government grants, investors, public donations, well-wishers, harnessing CSR, etc.

- Recruiting individuals with the right skill set and motivation to drive a countrywide wave of change.

- Analysing how sports transform the general lifestyle and how it applies to the masses.

- Transforming the prevailing mind set where the pursuit of professional sports is looked down upon.

The game plan: Deploying a winning solution

Organised sport is a great way to inculcate the value of hard work, and skills like leadership, teamwork, and problem solving among children in the age group of 6-16. Other important qualities like competitiveness, motivation, and drive can also be imparted through the medium of sport. Sports also create a sense of belonging with one's peers and binds communities. So how can we use sports to socially transform this segment? Here's how.

- Deploy a program that aims to address societal issues and empower underprivileged children between ages 7-16 through sports such as football, badminton and cricket.

- Set up partnerships with:
    - Sports teams and clubs like Pune Youth Club, Deccan Gymkhana, FC Pune City, Vengasarkar academy, and TATA Sports Club.
    - Governing bodies such as Pune District Football Association (PDFA), All India Football Federation (AIFF).

- Fund the program through sponsorships, advertisements, grants, donations and so on.

Hitting a home run with the benefits

Using a data-centric approach can help make the program a viable option for sports authorities and bodies by keeping a lid on the costs while ensuring program success. Done right, the program can deliver several benefits including:

Providing a platform for both teaching and learning.

- Creating a level playing field between people of different economic backgrounds and bringing larger sections of society together by tapping into their team spirit.

- Keeping youth under the age of 21 away from crime, drug related activities, gang wars – common among poor neighbourhoods.

- Educating children and parents about the benefits of a healthy lifestyle.

- Inculcating the importance of maintaining self-control, a positive moral compass, social awareness, and the ability to distance oneself from anti-social behaviour.

Winning hands down: Using sports to spearhead transformation

Sports teach the essence of togetherness, comradery, and community, playing a critical role in social transformation. Teaching children to adhere to activities like sport from a young age helps craft their thought process in ways that prove beneficial to the society in the long run.

ABOUT THE AUTHOR

Arnav Kholkar is a Sales Intern at KreedOn and a Business Developer and Media Advisor at The Butt Ballot. He is currently studying Sports Management at Manchester Metropolitan University and is on an exchange year at Aalborg University in Denmark. Arnav's experience includes marketing positions with FC Pune City, KreedOn, and Manchester City FC; as well as volunteering opportunities with MMU Sport, Manchester City FC, FC Pune City.

References:

- Census of India. (2011), Population Enumeration Data (Final Population), Updated: 15th January, 2019, Accessed on 28th January, 2019, *http://www.censusindia.gov.in/2011census/population_enumeration.html*

- Statista. (2017), India: Age distribution from 2007-2017, Accessed on 30th January, 2019, *https://www.statista.com/statistics/271315/age-distribution-in-india/*

- Statista. (2018a), India: Population by wealth bracket in 2018. Accessed on 30th January, 2019, *https://www.statista.com/statistics/482579/india-population-by-average-wealth/*

- Statista. (2018b), India: What sports do you regularly participate in? Accessed on 27th January 2019, *https://www.statista.com/statistics/562694/india-regular-participation-in-sports-by-type/*

Open Data in the Social Sector: Empowering Communities in Unprecedented Ways

Donald A. Lobo
CEO, The Chintu Gudiya Foundation

Open source software holds tremendous potential for revolutionizing the social development sector, where organizations not only need to work on a tight budget but must also have robust systems to monitor the impact of their work on the community. Open source tool sets that can be easily adopted and used in a short time frame, often have stronger and more diverse vendor support options, and are better suited for integration with other tools, platforms and services. Given that nonprofits often cannot afford to have a large internal technical team, access to such tools is a boon. To understand this better, let's take the example of two nonprofit organizations that are leveraging open source platforms and reaping the benefits.

Shelter Associates is a nonprofit organization working to improve the lives of urban poor through low-cost housing and sanitation solutions in Maharashtra. Shelter uses a geographic information systems (GIS) open source toolset called QGIS to map urban slums and low-income households - across the Sangli-Miraj-Kupwad area near Pune. The GIS software maps the slum as a 2-D image and indicates essential services including water, drainage, sewer, and electric lines. Shelter also uses Kobo toolbox, an open source platform that collects data using mobile phones, to collect crucial citizen data, including demographics, access to water and sanitation, sewage supply and electricity coverage. This data is then merged with the GIS map, allowing Shelter to pinpoint gaps in public service delivery, as well as monitor the age of existing facilities. The data allows Shelter staff and funding organizations to decide where to allocate resources, and identify which houses need sanitation facilities and what their priorities are. Different funding organizations have different goals. For instance, some want to primarily provide toilets to houses with adolescent girls. Having a flexible data collection and analysis framework gives Shelter the flexibility to tailor its work based on individual project needs.

The data is also uploaded onto Google Maps, enhancing public visibility and understanding of the housing and infrastructure issues in the area. Shelter's initiative shows that open source software and data can be used to create an impact that extends far beyond the immediate project. The data visualization also allows Shelter to share information with other non-profits seeking to do similar work in the vicinity. For instance, Shelter provided the collected data to an NGO that is focused on electrification of houses. Based on the available information, the NGO was able to focus on areas where the problem was the most severe.

Shelter has also worked closely with the Ministry of Housing & Urban Poverty Alleviation (HUPA) to provide insights into the challenges faced by the most vulnerable communities in gaining equal access to and ensuring the maintenance of essential communal services. This has helped the government optimally direct resources to create superior impact.

Investing in open source data platforms can generate incremental impact for the social sector as organizations can benefit from a shared data ecosystem to fast-track their operations and widen reach – all in a cost-effective manner.

Another open source data leader in the development space is Aangan, a nonprofit working to ensure child protection in vulnerable communities across India. In 2017, Aangan collected household data on child safety in the community, using hand-held mobile devices. Aangan uses a platform based on Open Data Kit (ODK), an open source project that allows offline data collection via mobile phones. In 2018, they expanded their efforts to the community level and used a pen-and-paper data collection model instead. The latter proved to be more effective in engaging the community as well as the local government officials with whom the information was eventually shared. The model included a community-based map, plotting various resources in the community along with safe and problem areas marked out. Aangan's larger vision is to make data on community perception of child safety (in over 200 vulnerable communities across India) available for public use and analysis, combined with corresponding demographical data on education and health outcomes in these communities.

Aangan's plans include gathering socio-economic information for India at a district level and preferably at a block or village level. When the data becomes available, it will be merged with the data that Aangan and Shelter have gathered at a slum level, to predict which villages or districts they need to focus on next. A large set of machine learning algorithms and libraries from the open source world will be used to make such predictions. The plan is to make the dataset available as an open data set so that other organizations can leverage and build on it.

The compounded impact of open data in the public sector cannot be ignored.

One reason public sector organizations do not want the software that they have spent significant time and effort in creating to be freely available for use by other organizations, is because they compete for the same resources. While this is a fair concern, the benefits that can accrue to the sector from using open source tools and data far outweigh the concerns associated with it.

The software created for one nonprofit's use can be adapted by multiple organizations cutting across issues and geographies, in turn, benefiting a larger number of vulnerable and marginalized communities. This allows us to build a strong, vibrant and stable ecosystem of open source organizations, NGO partners, and beneficiaries. The public sector must invest resources in open source platforms to create more empowered communities, widen the reach of social services, and make the world a better place – one step at a time.

ABOUT THE AUTHOR

Donald A. Lobo is the CEO and Founder of the Chintu Gudiya Foundation that was established15 years ago. His approach to philanthropy centers around understanding pressing and neglected issues on the ground and betting on organizations with the most effective solutions. Previously, Lobo has served as the co-founder and lead developer of CiviCRM, an organization that gives the public sector a free and open alternative to their software needs. He was also part of Yahoo's founding team - 'Technical Yahoo'. Lobo holds a post-graduate degree in computer science.

Saving More Lives with Organ Donation

Tejpal Singh Batra
Co-founder and President, IGiftLife

India has seen an increase in the rate of organ donation in recent years. Yet, the list of patients waiting for successful organ transplant procedures continues to grow. This article attempts to present an overview of the current status and explores corrective measures to address the issue.

Current Status, trends, comparison

As per NOTTO, the number of organ donations in India has steadily grown over the years in the recent past. However, it is to be noted that we do not have an organized central co-ordination system in the country to monitor organ donations. Hence, the figures may not accurately reflect the actual numbers on the ground.

The overall situation on organ donation in India can be summarized as follows. There is a huge gap in the demand for and supply of organs for transplant due to various factors that we will discuss further ahead in this article. This gap however, is contributing to the growth of organ black-market operations and organ trafficking, particularly kidneys.

India has an organ donation rate of 0.8 per million, which is very low when compared to other countries such as Spain with a significantly higher organ donation rate of 40 per million. Implementing the right measures can certainly help us work towards increasing the donor rate in the country, especially cadaver donors.

Listed below are some of the important legislations and government agencies involved in monitoring and regulating the organ donation process in the country:

Indian laws:

- Important laws and legislations pertaining to organ donation include:

    - Transplantation of Human Organs and Tissues Rules, 2014

    - THOA Amendment 2011

    - Transplantation of Human Organs (Amendment) Rules, 2008

    - THO Rules, 1995 (Original Rules)

    - THOA 1994

- The laws stipulate that the physician on-record treating the patient must ascertain the patient as being brain-dead as per prescribed process, and also sensitize the family about organ donation

This article examines the statistics, existing status and trends influencing organ donation in India. It further provides a detailed approach to increase the number of donors and organ donations in the country.
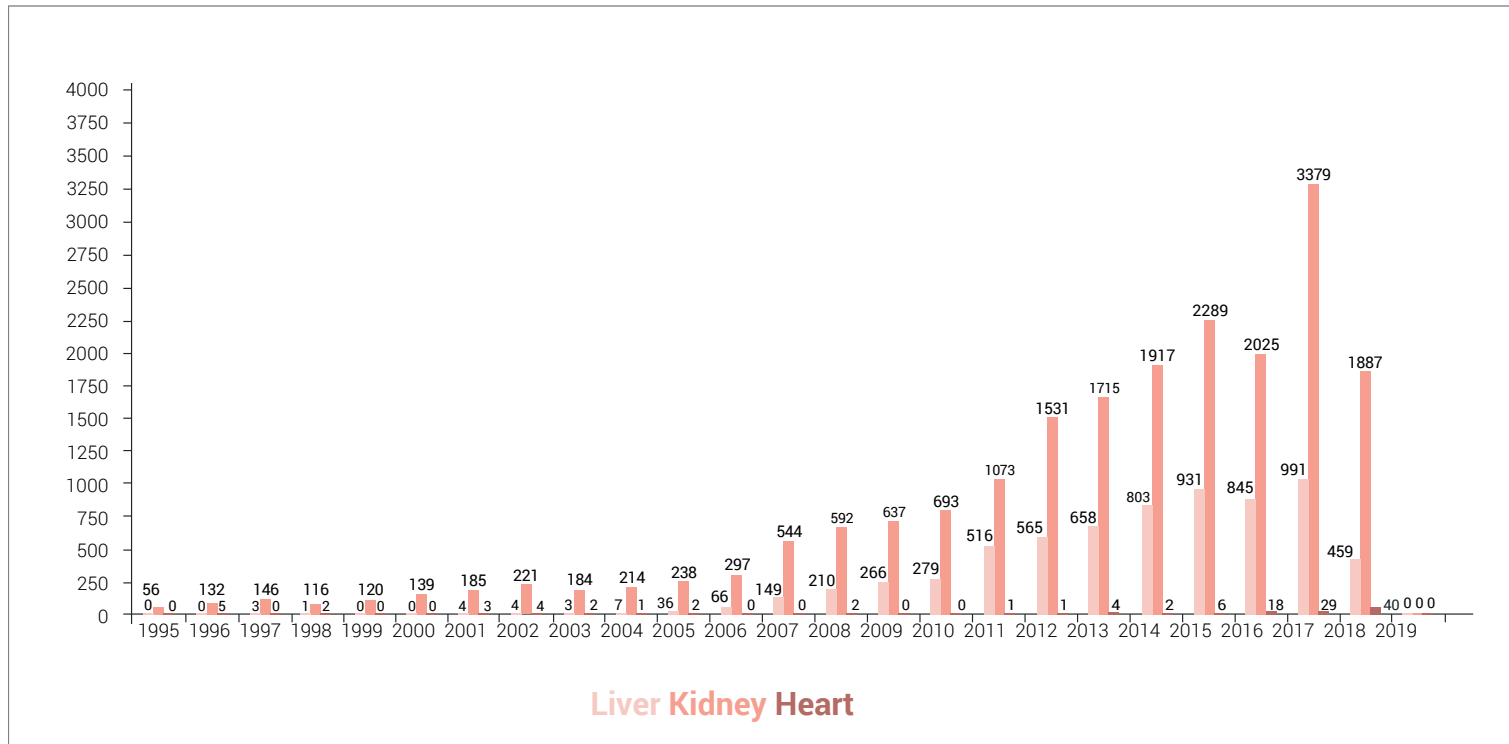
Statistics on the number of liver, kidney, heart transplants (Source – NOwTTO - https://www.notto.gov.in/organreport.htm)

State governments to allow 'non-transplant hospitals' equipped with an ICU and operation theatre to retrieve organs

Government bodies:

- NOTTO – Apex body in India on organ donation
- ROTTO and SOTTO – Function at regional and state levels
- ZTCC – Work at the zonal level within states

What can we do to save more lives?

Increasing organ donation and transplant numbers will require significant coordinated effort by the medical, regulatory and governing agencies. Listed below are some measures that we can adopt to increase the rate of organ donation in the country:

1. Spread awareness across all stakeholders:

One of the key aspects to focus on is to continuously spread awareness on the state of organ transplants across all stakeholders involved. For organ donations to increase, there must be significant social change. People need to be educated and informed about the benefits of organ donation in order to overcome myths, superstitions and customary beliefs that are impeding positive change.
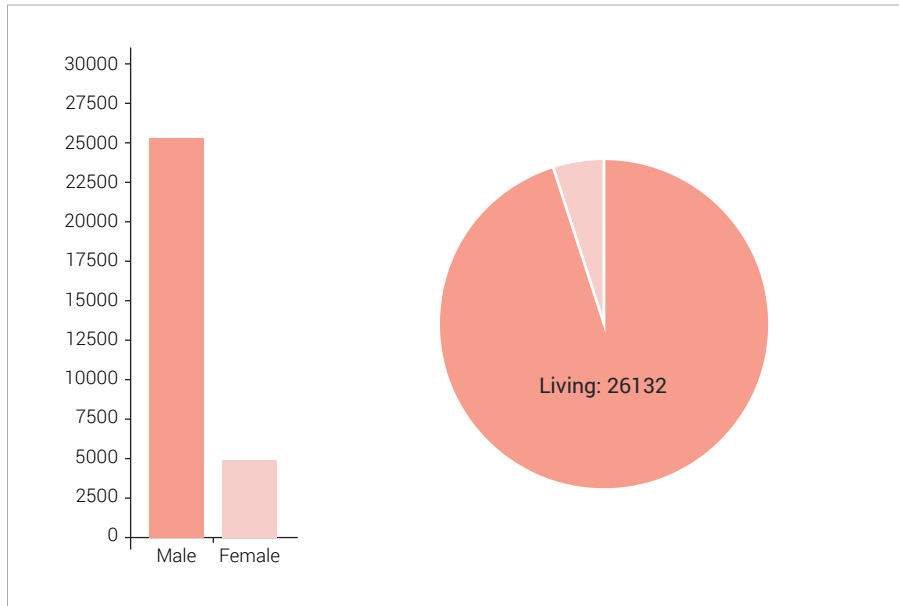
Ratio of male-female candidates
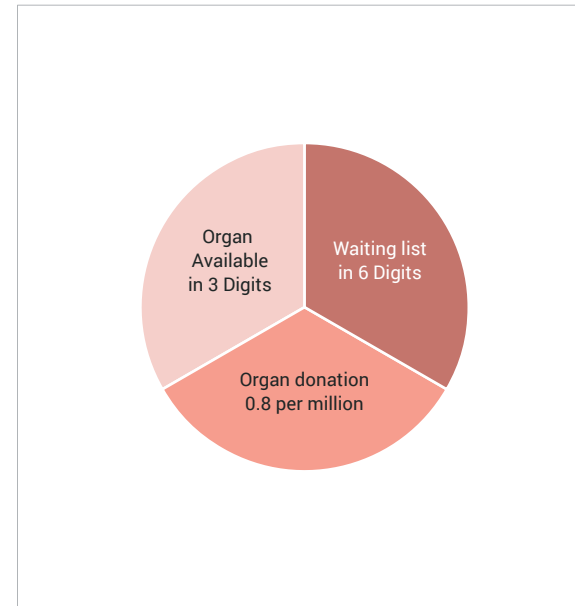
Ratio of living VS cadaver transplants for all organ transplants.

Stakeholder here would include:

- General public

- Doctors

- Transplant coordinators

- Donor families

- Recipient families

2. Streamline processes and infrastructure:

It is extremely important and crucial to put in place appropriate processes and supporting infrastructure to ensure an improvement in the rate of organ donation. These include:

- Availability of seamless transportation services for speedy delivery of donor organ

- Standardized online data base that can instantaneously match donor organs with potential receivers

- Implementation of the latest technology to further hasten and simplify the process

- Increase the number of certified hospitals that can facilitate organ retrieval and transplantation

- Increase in certified doctors to assist with complex procedures

- Streamline and simplify various legalities and regulations pertaining to organ donation

Figure: Living donor rate and Deceased donor rate per million population for New Zealand, France, Austraila, United Kingdom, Spain, United States, Canada.
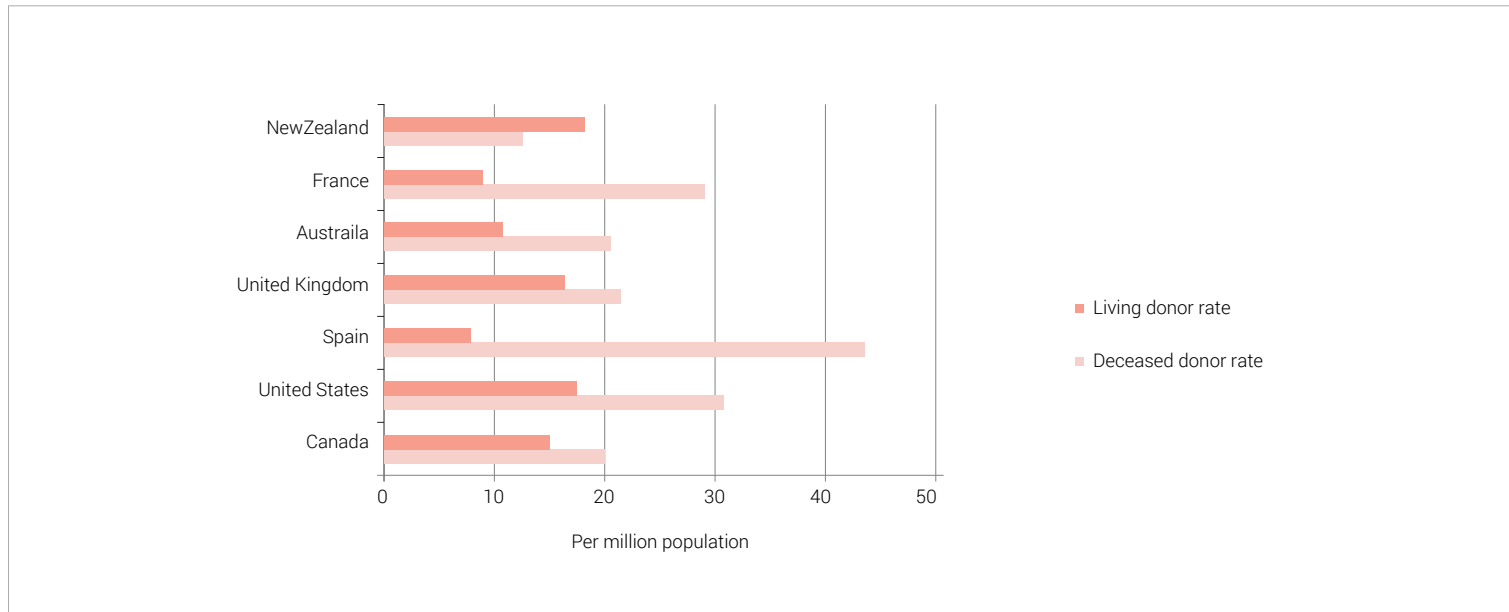
(Figure prepared by the Library of Parliament using data obtained from Global Observatory on Donation and Transplantation, Organ Donation and Transplantation Activities, 2015, September 2017.)

3. Create centralized registries:

Centralized registries that include data on both donors and recipients are a big advantage in identifying immediate and the most suitable matches for organ transplants. A key aspect to bear in mind is that zonal, regional and state agencies will continue to play a major role in a vast country like India, as time is of essence in successful organ transplants. For instance, the heart and lungs must be transplanted within 4-6 hours. Regional co-ordination might be important in this regard to help speed up the transplant process.

4. Presumed consent:

Another key aspect to consider is that of presumed consent. Under presumed consent, a person is deemed to be registered as an organ donor by default unless he opts out. However, given the social and cultural considerations in India, this factor is not considered a major facilitator in increasing organ donation, unless accompanied by significant investment in streamlining processes, infrastructure as well as changing people's mindsets.

5. Incentivizing donor families:

The question of incentivizing donors and their families is a complex issue. Any form of tangible incentive presented towards organ donation is likely to be viewed as unethical and counterproductive. However, one option to consider is to reward the donor's family with the promise of a prioritized position on the waitlist should they require an organ transplant in future.

6. Using data analytics:

While there is a lot of data available about organ donation, data quality and standardization needs careful assessment. Investments must be made in data analytics and related software and applications to gain relevant and actionable insights.

Conclusion

India is seeing an increase in the rate of organ donation, but we have a long way to go to ensure no life is lost due to the unavailability of a suitable donor organ. All parties including medical bodies, government agencies and regulatory authorities and the public at large need to work together to create an eco-system conducive for increased organ donation. This is a slow process and while change cannot happen overnight, consistent efforts will ensure that the organ donation statistics in India grow to match the statistics in Spain and other developed nations.

ABOUT THE AUTHOR

Tejpal Singh Batra is co-founder and president of IGiftLife - an NGO which focuses on saving more lives by facilitating organ donation. Tejpal has over 28 years of global experience in diverse leadership roles. Tejpal is also a visiting faculty member at various business schools across the country and runs his own transformation and analytics consulting firm. Since its inception in 2015, IGiftLife has made tremendous progress in the field of organ donation across 9 cities by conducting 190+ sessions and engaging 25000+ people across the country. The organization has also conducted two major events reaching out to 160,000 people on the topic of organ donation.

The Analytics of Doing Good

Jashodhara Chakraborti
Managing Consultant, Kolkata Network Foundation

Leveraging data analytics to fulfill corporate social responsibility can help explore the art of possible by pushing boundaries and enabling impact measurement with the right metrics. The key to getting this right is to intertwine analytics at the CSR program design stage itself and embed it throughout the lifecycle of the initiatives to enable course correction and optimization.

Social responsibility has always been a difficult thing to agree upon in its finer details, whether as an individual, a social sector professional or a corporate. In this age of metadata and hyperemotional responses, it has taken on perilous proportions.

After all, what is socially responsible when it comes to choosing between two pressing alternatives such as: environmental impact versus lower buying costs, modern manufacturing processes versus protection of artisanal heritage, job creation through global skill standards versus charitable compromise, and sourcing locally versus the next country, state or district?

Which of these alternatives are more important to you than the others? Chances are, they are different for you and your neighbor. The answers depend to a large extent on personal preferences and are emotionally loaded, often preventing a coherent end- to- end strategy.

Wherever there is a tangle of information without a strong analytical framework to provide direction, there is the conundrum of making an informed choice. Compromises are often made based on what the majority agree on or the leader's preference. Needless to say, a thoughtful and responsive analytical strategy is an invaluable tool for strategic decision making and measurement in such contexts.

Businesses know this; which is why there i are numerous studies on impact measurement of CSR programs online, yet there is often dissatisfaction when these are presented or applied. We seem to be looking at off- the-rack options when customization might be the need of the hour.

Intertwined analytics and program planning

Given the complexities involved, effective analytics in this space needs to be woven in from the program design stage itself. It might be well worth the time and effort for organizations to have at least one member of their CSR team with experience in data analytics. , Such a member can then train the team as a whole on the basics of CSR, sustainability goals, and effective application of analytics, especially the possibilities that open up once it is integrated.

Know your recipient's ask

To analyze impact effectively, we need to identify the recipient(s) of our program and decide what outcomes the program needs to drive and measure. Knowing the recipient's ask is the first task to building analytics in the CSR space. Is it a single, high impact intervention program? For instance, for an intervention such as building toilets for girls, the impact measurement could be how many girls are continuing to study in school vis-à-vis a year earlier, whether toilet usage among girls has increased to recommended health guidelines, and the cost per retained girl student. Is the program a sustained intervention where impact is expected to taper over time? For the same example, toilet maintenance schedule versus sustained usage numbers, and number of girls dropping out of schools could be suitable metrics to judge sustainability, even though the numbers are expected to stabilize over time after the initial sharp changes in data.

High impact interventions might be seductove but sustained welfare comes from less dramatic, longer term interventions.

The second requirement to understanding the recipient's ask is to be willing to listen without prejudice or without condescension – a common problem when it comes to dealing with the underprivileged. Trained qualitative analysts can add significant value at this stage. It is tempting for us to say that the solution is obvious but crucial nuances can emerge, which if addressed early on in the program design stage, can make all the difference between success and failure.

Modulate to serve better

The umbrella term of "responsible choice" means different things to different people, as discussed in the introduction. It follows therefore, that an umbrella score would be inadequate in measuring how individuals assess a funding choice. Neither would it be able to explain in a granular manner why an organization chooses one program over another.

However, it is possible to modulate the various ethical / socially responsible parameters that emerge while making a decision and provide weighted scores on each parameter based on personal preference or corporate vision. Once all initiatives or social programs under consideration are listed, the weighted scores on each module can emerge for a particular individual or organization, leading up to a collective score for each decision maker. This would immediately make prioritization of effort or fund allocation simple and free of any bias. For ethical customers, such a practical tool would take into account the fact that there is no such thing as a perfectly ethical choice, only an informed choice, given their unique ethical parameters.

Moving from "what is" to "what if" in analytics

A brilliant social responsibility program requires imagination. Too often, we are mired in the existing realities and derive our policies by baselining "what is". This offers incremental benefits at best.

Instead, we need to look at "what if" - where we are willing to look beyond reality to derive our programs from what the ideal situation should be. The key is to do the initial research or qualitative analysis which incorporates such an explorative method.

If analytics can help us identify the "what if" scenario along with its metric scores, it can lead to the following benefits:

- Ability to measure the baseline and incremental impact on the baseline as before

- Ability to accommodate sub-programs which might look unrelated with respect to the baseline analysis or "what is" but might deliver value within the larger ambit of the "what if" analysis, thereby providing value on the big picture to all stakeholders

- Enable the option of exponential benefits instead of negating that possibility right at the beginning.

Effective analytics coupled with the right questions at the design stage can be a powerful catalyst for social change. It can serve by cutting out the fog caused by emotion, bias, and varied perspectives; by providing a multiple indexed measure that is , capable of enabling streamlined fund usage and less wastage.

Most significantly, effective analytics, which explores the art of the possible in CSR, can deliver hope where it is needed most.

ABOUT THE AUTHOR

Jashodhara Chakraborti (Jash) is an entrepreneur, writer, columnist, speaker, and interviewer interested in the vanishing point where technology, storytelling and analytics converge. She is currently Managing Consultant at Kolkata Network Foundation and Consultant, Bharatiya Bhasha Parishad, where she is working on creating a digital literature platform for the Bengali and Hindi languages respectively.

Her passion is to create a sustained, measurable digital presence of quality content in Indian languages and to take stories in Indian languages in real time to every Indian.

In Conversation with Dr. Poornima Dore

At Tata Trusts, data-driven decision management is an integral aspect as data validation is the basis of making accurate and measurable social impact. However, implementing an organization-wide data driven approach is not without its challenges and organizations need effective change management and the right technological intervention to tide through successfully.

Dr. Poornima Dore
Head, Data Driven Governance, Tata Trusts

Kindly describe your current role.

As the head of Data Driven Governance at Tata Trusts my role involves partnering with government bodies to build technology and data backed models for facilitating data-driven decision-making. The focus is to enable key functionaries such as district collectors, municipal commissioners, and members of parliament, to improve service delivery through data based planning, targeted interventions and cutting edge technology – for systematic change in development planning and financing process. Currently, our work spans across 85 districts and 8 cities pan India.

What are you and your organization doing to make a positive impact on the society and environment?

The Data Driven Governance team at Tata Trusts is dedicated to institutionalizing data culture in Indian villages and cities to enhance transparency and accountability. Despite the vision of digital development in India, there is need of accurate information and international standards for smart, resilient and inclusive growth of rural and urban India. At Tata Trusts, we capture data at scale to prepare plans, enhance policymaking and impactful development

initiatives. We have developed a protocol and application called 'DELTA' (Data Evaluation Learning Technology and Analysis) for participatory planning. This also includes capacity building on the use of data and technology.

What are the typical challenges you face while doing this work?

The key challenges that we face include:

• Siloed data: Though government collects a reasonable amount of data, it is a challenge to triangulate and collect all the data, as most of it exists in siloes. In addition, it is difficult to share data amongst different government departments, as it hasn't been mandated.

• Data is not in real-time: Most government data which is granular, exists in the form of census data, which is collected once in 10 years. Or, the rest of the data available is in the form of periodic surveys. As this data is not in real-time it is difficult to administer it.

• Lack of data quality: Most of the data remains restricted within the departments with no effective data monitoring. This makes it challenging to verify the authenticity of data.

What is the relevance of data and analytics in the social work that you are doing?

It is very important for decision makers to focus their efforts and limited time, especially in the realm of public governance, where the problems are complex. Aggregating quality data streams under a unifying hub and analyzing it at the right time enables pragmatic decision-making, efficient strategic planning and proactive policy reviews. Cities such as Pune, Surat and Jamshedpur are now ISO certified, as our team has been able to drive actionable intelligence from high quality city data. This has helped enhance city planning, improve infrastructure investment and streamline public services. Similarly, village level and gram panchayats data has enabled verification, targeted interventions and convergence of government plans and schemes.

How is a government department using data & analytics different from the one which does not leverage the same?

Tata Trusts works very closely with administrative government bodies and decision makers to build culture of data and facilitate accountable governance. For administrative decision makers, the ability to focus and resolve real problems is enhanced by use of data analytics. For instance, data can help the district collector identify which villages have reported issues with access to electricity, sort by type of problem at a click, and galvanize the electricity department to set monthly and quarterly targets to address gaps in specific locations. Similarly, if there are geographic contiguous pockets in remote areas that have not got access to specific schemes, mobilization drives and awareness camps can focus on such locations. As a next step, making select aggregate level data open and accessible can make it possible for the government to collaborate with other stakeholders in the ecosystem; this allows to solve issues swiftly, innovatively and at scale through a participatory approach.

What challenges did you face while implementing a data driven approach with government entities? Was there any kind of resistance to change?

At Tata Trusts, we are proactively involved in looking at technology-backed models that help enhance development outcomes on the ground. On several occasions we have faced a lot of cynicism since inherently there is limited updation of or reliance on datasets. Also, data sharing within government does not happen actively between departments and so

getting multiple departments together to explore synergies is always a challenge. We find that change management has been a crucial element to help policy makers embrace and learn new ways of working with data. Use cases are a good way to build credibility of the value of informed decisions. Our strategy has been to work selectively with enlightened administrators within the system who see the value of such ideas and act as champions of change within the organization. Our role is that of a catalyst. On the ground, we have found that volunteers collecting the data for our surveys gradually become comfortable with using mobile devices after requisite training and evolve as local change agents themselves. Over a period of time, after successful interventions, we have been invited by the Government of Maharashtra, Niti Ayog, and the Smart Cities Mission to partner at scale.

How does a philanthropy differ from a corporate entity while taking a data driven approach?

Tata Trusts takes a philanthropic approach to make impactful and sustainable changes in the lives of communities. We leverage data to primarily make government programs more targeted and assist decision makers to meet aspirations of the people of this country. In addition, we put great emphasis on people participation and make sure that people collecting data are locals. We also earn a lot of credibility within the Government as they know we have no vested interests at play and are purely interested in system strengthening. Through it's work and experience of more than 100 years, Tata Trusts has contributed towards nation building. It is because of the faith that the community has placed in us, that we can continue to do the same even today.

What specific advantages did you derive through a data driven approach?

By leveraging a data-driven approach, we have been able to facilitate action at the local and national level and develop models for unifying government schemes. For a developing country like India, data is still at a nascent stage. However, the top administrators see value in the use of data for better governance. This has helped us take risks to test new program designs and develop them by partnering with proactive government departments and agile technology partners. The result has been tremendous in terms of streamlined and successful planning of development projects. DELTA framework is now being implemented at scale across the state of Maharashtra to create models of data driven development planning in the country. It is also being used as the foundation for the "Transform Aspirational Districts" initiative of the Niti Aayog to map the performance of districts pan India across select indicators. The urban dimension has been enhanced through the city data for India initiative which engages with identified ULBs to potentially break down data silos in individual government departments, by way of institutionalizing a standardized data collection framework (ISO 37120). The "eGov DIGIT Platform" is a comprehensive state of the art open source, inter operable and scalable civic tech solution for citizens and city administration to manage, monitor and interact in a transparent, effective and efficient manner. Our pilot of deputing a city data officer at Pune Municipal Corporation in partnership with TCS, has resulted in the Urban Ministry and Smart Cities Mission nomination 100 city data officers across all 100 smart cities.

What role did technology play in this transformation journey?

Technology has helped us prepare scalable applications, to collect, analyze and host data while using handheld devices. It has improved the speed of processing information and providing insights for District Collectors and Members of Parliament through digitizing village development plans, identifying group areas at a click of a button, and enabling dashboard based services of developing outcomes. In the urban space, the recently launched National Urban Innovation Stack seeks to be an open, single stop platform to bring together all agents in the ecosystem of city governance to identify urban challenges, solve them at speed and scale and strengthen the ecosystem. The platform also hosts capacity building material and we are in the process of developing e-learning modules on data driven governance. Such tools, when developed for the public space, can have a big impact at a large scale

Kindly describe your personal experience while pursuing this transformation journey.

It has been an enriching experience to work with data for service delivery. I strongly feel that through comparable, certified and standardized data, the government will be able to enhance citizen engagement to truly realize the vision of a digital India. This involves a structural change towards a culture of data and will require a concerted effect by all actors at all levels, both within and outside government.

As an economist and management professional, I am fortunate to experience the power of stronger governance systems to solve real development problems. My core team now looks at data and tech, not just as a tool to be harnessed, but as a force for good which if applied wisely and carefully, can yield meaningful results and drive constructive partnerships. Academia and industry have the potential to be valuable contributors as well. The energy and foresight of selected leaders in the political process and administrative services, coupled with the support and backing of senior leadership at Tata Trusts and Mr. Tata himself, has played a big role in giving our team the opportunity to experiment, grow and deliver on an idea whose time has come.

ABOUT DR. POORNIMA DORE

Dr. Poornima Dore is an Economist and a TAS management professional with the Tata Group. She heads the Data Driven Governance vertical at the Tata Trusts, and works on systemic change in the development planning and financing process for villages and cities in India. Her strengths lie in managing large projects, building teams, designing innovative financial structures, and driving meaningful partnerships across companies, government, civil society and academia.

In her earlier roles, she has been part of the start-up team at Tata Capital Ltd, where she headed the Structured Investments portfolio, anchoring some landmark equity and mezzanine debt investments across sectors such as automotive, pharmaceutical, technology and insurance. She has a PhD. in Economics from Indian Institute of Technology, Bombay with an interest in regional growth, employment, finance and urbanisation. She is a gold medalist from XLRI Jamshedpur, and has a bachelor's degree in Economics and is a Principal's Award holder from Lady Shri Ram College. Her other interests include reading and percussion, in addition to teaching and writing on aspects of planning and policy.

POLICY & GOVERNANCE

Going from 'Rural to Robust' with Data Analytics

Pradeep Lokhande
Founder, Rural Relations

Social enterprises need data for direction much like for-profit organizations. Recognizing this, social entrepreneurs as well as activists are seizing the opportunities provided by today's data deluge to deliver social transformation. For experts in the fields of market research, sociology, and humanitarian relief data, analytics helps drive empirically validated problem solving. Project success often depends upon expertise in data management and visualization tools.

How does data analytics help drive wider transformation for a social enterprise? Predictive data analytics helps social enterprises spot future trends, societal preferences, and behavioral changes in the society at the right moment. For example, post-campaign data analytics showed that Bollywood actors promoting agricultural equipment had a persuasive effect on farmers of the previous generation. It also showed that the new generation of farmers are similarly motivated by seeing cricket stars in similar promotions, enabling organizations to fine tune their messaging for greater impact.

Rural Relations, a social enterprise focused on delivering sustainable human development in India, leverages diverse datasets gathered from projects in different states of India through Gyan-Key – an initiative aimed at opening libraries in rural secondary schools. VillagewiKY, Rural Relations' open rural information platform, helps the organization draw insightful correlations from the data sets with a high degree of granularity.

The key Big Data analytics challenge for most social enterprises is tracking and targeting the effect or impact of data-driven strategies.

How can social enterprises make the best use of data?

The first step is to ensure the collection of social data on critical rural issues. Rural Indians are predominantly farmers. The increasing use of mobile phones and internet among this group makes it possible for social entrepreneurs to connect with them easily for data collection. Once the data is collected, deploying analytics enables customized insights for farmers. Farmers can easily access information on new agricultural practices and tweet or message agriculture experts from their mobile phones for additional information.

Help society form a data network

Forming a data network helps social organizations make the most of data collection and analysis. For example, Rural Relations' villagewiKY platform encourages villagers, opinion leaders, and the non-resident villagers to provide data on their villages. If the data network is built with the involvement of the community, it ensures that the community has a stake in extracting benefits from the network. Private companies or social enterprises that need to prioritize social transformation efforts can zero in on areas that they need to focus on based on the analytical insights from the open data platform.

Social enterprises in India are now actively leveraging data and analytics to drive greater impact and build data networks, thanks to increasing smartphone penetration in rural areas and willingness among people to share information. More data spurs better analysis and aids in driving more targeted social transformation.

How the villagewiKY data network works

The villagewiKY data network has set the ambitious target of capturing social data from 85,000 villages in India within 1,000 working days. The goal is to become a major driver, enabling a digitalized republic by 2019. Here's how it works:

On the open rural information platform, individuals can contribute, add or rectify the information on their village based on parameters such as: demography, geography, ecology, sociology, enterprises and institutes operating in the village, and details of workers in each profession.

The information gathered so far is the result of 20 years of engagement with people in villages and opinion leaders.

Villagers and non-resident villagers (NRV) provide the data on the platform that is gathered by Rural Relations' village development officers.

VillagewiKY will present useful information (5% to 80%) of 58,000 villages on its portal *www.ruralrelations.com*.

ABOUT THE AUTHOR

Pradeep Lokhande is the founder of Rural Relations, a social enterprise focused on rural human development. The firm pioneers an entirely new way of reaching out to tomorrow's India with the simple desire to contribute to its development and growth.

Data, Data Everywhere,
Not a Story to Tell

Misha Sharma
Senior Manager, Research Centre for Social Impact and Philanthropy, Ashoka University

Even as the size of India's philanthropic sector grows rapidly, the lack of data, standardized reporting formats, and definition on the typology of social sector organizations hinders transparency and undermines the sector's value. Governments, social impact organizations and philanthropic foundations and individuals must collaborate to create a robust data and reporting ecosystem that can help enhance outcomes in a transparent, scalable, and more accurate manner.

The Indian philanthropic sector has grown significantly over the past few years and is believed to play a major role in providing support to the socio-economic development of the Indian economy. However, there is a lack of rigorous, holistic, and data-driven approaches in estimating its size. Understanding the size of the sector and the total philanthropic capital available is crucial to building its identity and narrative. The lack of complete, accurate and standardized data on Indian philanthropy makes the process of extracting and analyzing data extremely cumbersome.

Several think tanks and consulting organizations have attempted to estimate the total volume of Indian philanthropy in the past. *Table 1* provides a snapshot of some of the prominent research reports that have been published on the topic. In the absence of rigorous and holistic data on the sector, different studies have used different methodologies and definitions to estimate the philanthropic capital in India. For instance, while the India Philanthropy Report 2017 by Bain and Dasra does not present detailed data on individual giving, the report by MacArthur and Intellecap uses estimates based on the level of per capita giving and population size of Indian adults, to measure the size of individual giving. The 'India Giving' report on the other hand, uses a different approach by conducting a primary survey to measure the quantum of individual giving. Reports estimating the volume of corporate philanthropy use public data available through the Ministry of Corporate Affairs website. However, most of the reports base their analysis only on partial data due to the difficulty in accessing the data in its current format.

Mapping the size of Indian philanthropy – A holistic attempt by The Centre for Social Impact and Philanthropy (CSIP)

The Centre for Social Impact and Philanthropy (CSIP) at Ashoka University, as part of its mission to inform sector strategy through credible research and data, recently commissioned a study to map the volume, value, composition, and trends on the available financial resources in the philanthropic sector. The study uses a quantitative approach to analyzing publicly available data on financial flows into the sector as a whole. The research expands on the limited knowledge that exists in the sector, thereby playing a huge role in informing policymakers, philanthropists and social impact organizations about the sources and volume of philanthropy. It enables them to make better-informed and strategic decisions.

The research conducted by CSIP solely uses publicly available data on corporate, foreign, government and individual philanthropy to estimate the total value and volume of the available philanthropic capital. The study examines the patterns in resource allocation across geographic and thematic areas. However, the research is limited by challenges in accessing good quality data on the sector including:

- Lack of standardized reporting formats across regulatory authorities- The social impact sector is currently governed under different regulatory frameworks such as Section 8 of the Indian Companies Act, Bombay Public

SUMMARY OF REPORTS ESTIMATING PHILANTHROPIC CAPITAL IN INDIA

Philanthropy Classification	Indian Philanthropy Report 2017	Strategic Philanthropic Giving and Impact Investing for Development in India	India Giving	Hurun India Philanthropy List 2017	India's CSR Reporting N=100 companies	CSR Analysis of BSE 370 Companies, N=370 Companies
	Bain and Dasra	MacArthur and Intellecap	Charities Aid Foundation	Hurun	KPMG	NGOBOX
Billion (USD)						
Government Spending	23	32	No Data	No Data	No Data	No Data
Bihar	4	No Data	No Data	No Data	No Data	No Data
Chandigarh	2	3.3	No Data	No Data	1.1	0.9
Delhi	5	11.8	No Data	0.4	No Data	No Data
Haryana	No Data	5.1	0.75	No Data	No Data	No Data
Jharkhand	No Data	1.9	No Data	No Data	No Data	No Data
Maharashtra	34	54.1	0.75	0.4	1.1	0.9
Punjab	2016	2016*	2012	2016	FY 2017-17	FY 2017-17

* Data used for estimation however ranges from the years 2011, '12, '13, '14 & '15

Table 1: Summary of Reports Estimating Philanthropic Capital in India

Trusts Act, Indian Trusts Acts, and Societies Registration Act. A social sector organization can choose to register itself in any of these forms depending on its location, nature of work and so on. This leads to different compliance requirements for different types of social sector organizations, resulting in the lack of standardized data across organizations. One way to rectify this issue is for the government to take steps towards standardizing data collected across different platforms to improve the data quality. Better co-ordination between various regulatory authorities can ensure better quality of NGO data, in turn leading to better estimates on the state of philanthropy.

- Lack of transparency - Very little data is available in the public domain on social sector organizations (both NGOs and philanthropic foundations) in India. One leading cause - filing of returns is not mandatory for all types of social sector organizations. Additionally, digital reporting of data such as annual income and expenditure statements of philanthropic organizations is not mandatory in a majority of Indian states. This prevents the sector from ensuring transparency in financial records, making data-driven and strategic decisions hard to come by. Barring the non-profit companies that fall under Section 8 of the Indian Companies Act 2013, there are no accounting standards for social impact organizations registered under other forms (trusts or societies), making data transparency a far-fetched reality.

- Lack of data on private philanthropy by individuals - Data on individual (ordinary and high-net worth individuals) philanthropy is the biggest missing link, as information on individual giving is available only for those individuals who file for income tax exemption under Section 80G. However, individuals donate through other channels as well, not all of which is captured under Section 80G. The lack of research and knowledge on donor behavior prevents us from holistically estimating the size of the sector. Philanthropies need to evolve and implement reporting norms to strengthen the sector's transparency. User-friendly portals and interfaces can help us better manage and organize public datasets on different sources of philanthropy. The datasets can also be standardized and used for cross-comparison by adding a field for a common unique identifier (such as a PAN CARD Number) across databases.

- Lack of standardized definition on the typology of social sector organizations - As per the current regulatory framework, a wide range of organizations is included under the rubric of 'non-profit institutions'. This includes hospitals, schools, colleges, clubs and associations as well as those focused on service delivery, rights, advocacy, etc. Clearly, some of these organizations such as schools, hospitals, and clubs do not necessarily focus on 'social impact'. This dilutes the analytical utility of the current definition of 'nonprofit institutions' for identifying social impact organizations and building useful data on them.

Using data for social good: The way ahead

For stakeholders operating in the social impact and philanthropy sector, building a strong narrative around the sector is extremely important to make strategic, data-driven decisions and further its cause. Data can play a pivotal role in answering key questions about the landscape and identity of the sector. This demands that stakeholders work towards overcoming the limitations and contribute to making the sector more transparent. While social impact organizations and philanthropic foundations must be encouraged to make their data publicly available and comply with the regulatory norms, government efforts to build a robust system to collect and manage good quality data can further enhance the outcomes.

ABOUT THE AUTHOR

Misha Sharma is Senior Manager, Research at the Centre for Social Impact and Philanthropy at Ashoka University. She leads the research agenda at the center and is involved in producing evidence-based knowledge on the Indian philanthropy sector. Previously, she worked as a Project Manager at IFMR LEAD, a research organization where she focused on managing several research projects within the thematic area of financial inclusion. Prior to joining IFMR LEAD, Misha worked with Goldman Sachs as an Operations Analyst. She holds a Master's Degree in Economics from University of Edinburgh and a Bachelor's Degree in Economics from Stella Maris College, Chennai.

Doing More with Data: Using Data for Social Good

Dr. Neha Sharma
Secretary, Society for Data Science

Two decades ago, only one-fourth of global data was in digital form. Today, almost 98% of data is digital. The exponential growth is the result of data pouring in from various sources such as satellite images, sensors, social networking sites, mobile devices, and the internet, in the form of videos, digital images, voice, and text. A large portion of the Big Data is generated as a result of increasing human interaction with smart devices and online services - commonly referred to as data exhaust. It's important to remember that unless the Big Data is put in the right context using analytics, it cannot benefit business, society or humanity.

How nonprofit and for-profit organizations are using data to drive social transformation

The current generation of social entrepreneurs and activists see data analysis as a tool for transforming the society. The inherent diversity of Big Data creates infinite possibilities for using it to address the world's most critical problems like poverty, illness, ecological harm, war, famine, security and so on.

Many nonprofit companies and public-private partnerships are involved in Big Data innovation for the benefit of the society. Datakind, previously Data Without Borders, is a non-profit organization that brings together well-meaning data scientists and members of the civil society to use data science in the service of humanity. Some of the exciting projects the organization is working on include:

Using data to create meaningful growth paths for the homeless

- Creating safer streets through data science
- Advancing financial inclusion in Senegal using predictive modeling
- Using open data to uncover potential corruption
- Giving rural women a voice through mobile surveys
- Forecasting water demand in California where every drop counts
- De-siloing data to help improve the lives of those suffering from mental illness
- Gaining visibility into social issues in the UK
- Delving into child poverty data
- Using open data to prevent home fire deaths and injuries

Another non-profit tech organization Ushahidi, Inc., with a presence in nine countries, offers products that enable local observers to submit reports using smart phones by creating geospatial and temporal archive of events.

Creating data banks that can analyze and curate social data and leveraging citizen science i.e. the involving citizens in the data mining process can help drive targeted social benefits. From reducing crime and corruption to improving the quality of life by forecasting disease, climate, and a range of other geo-social aspects, data, when used right, has the potential to enrich the society in unprecedented ways.

The platform is used by a social activist group 'Humanitarian Tracker' for monitoring violence in the Syrian civil war. The United Nations department of field service and peace keeping also used the platform for surveying Nigerian elections and facilitating eyewitness reporting and analysis of violence in Kenya.

At the same time, for-profit companies are also using Big Data for social transformation. A Hong Kong based company DemystData, mines online data and social media posts to help financial institutions serve people who currently do not have access to financial systems. IBM's 'Many Eyes' website visualizes user generated innovation and insights to attract businesses. Former U.S. President Barack Obama, during his administrative tenure, encouraged the use of the government Big Data set for gaining more insights to drive innovation.

People working in the field of public health and humanitarian relief consider Big Data a significant driver of empirical problem-solving. Big Data is also being used for predicting the happiness index and unemployment and crime rates at national levels. Startups are also increasingly leveraging data for social transformation. For instance, Healthmap uses informal data to predict disease outbreak, Orgnet a social network uses data insights for crime control, and Ayasdi mines text and twitter data for predicting unemployment.

Pushing data analytics to do more

Big Data is still highly unstructured when it comes to addressing grave social concerns like human trafficking. Data is either not shared due to privacy and security issues or is unreliable. Most of the meaningful data is concealed in administrative systems, and to add to that, there is lack of data governance standards. It is therefore important to take necessary steps to create the right ecosystem that can help stakeholders leverage data for solving tough social problems.

One way to do this is for global communities to start creating data banks to helps solve critical issues like poverty, hunger, and human trafficking. The data banks must be supported by a team who can curate and analyze the data. It is also possible to generate deeper insights by tapping into citizen science - where data scientists involve citizens while creating new ideas and products. While data analytics has been delivering significant benefits to both business and science, it's now time to curate data with the explicit purpose of using it to transform the society at large.

ABOUT THE AUTHOR

Dr. Neha Sharma currently serves as the Secretary of Society for Data Science, India. Her role involves integrating teaching with the current needs of the industry to improve the state of employability in India. Previously, she was the Director, Zeal Institute of Business Administration, Computer Application & Research, Pune, India, the Dy. Director, Padmashree Dr. D.Y.Patil Institute of Master of Computer Applications, Akurdi, Pune, and a Professor at IICMR, Pune. She has published several papers in reputed indexed journals, both national as well as international and has organized several technology conferences and seminars. Dr. Sharma holds a Ph.D. degree from Indian Institute of Technology (ISM), Dhanbad.

Betting on Behavioral Change for Social Transformation: How Data Drives it All

Ruchi Mathur
CEO, Pune City Connect

Let's start with a personal account of how data enables behavioral change. Growing up, bread was a breakfast staple - a practice that started when my parents lived in England during the early years of their marriage. The habit persisted with me for all these years, until I read the book "Food: WTF Should I Eat". In the book, the author states that eating two slices of whole wheat bread raises your sugar level more than eating two tablespoons of sugar! Incredibly enough, this information caused me to stop eating bread entirely. A simple data input caused fundamental and immediate change in my behavior.

The power of data in causing behavioral change is indisputable and there are several documented case stories both at the individual and community levels. Behavioral change is, in turn, the basis for driving social transformation.

Data provides strategic insights for the Lighthouse project

Pune Municipal Corporation (PMC) in conjunction with Pune City Connect runs a sustainable livelihood program called the Lighthouse. It is a program that attempts to bring about deep-rooted shifts in the psyche of youth from low-income communities. The core idea is that without inner transformation in terms of self-confidence, self-discovery, and a sense of possibility that we call 'agency', outer transformation such as being able to get and hold down a job or thrive in the workplace, is not possible. The inner transformation journey starts with a foundation course for the youth, followed by counseling and vocational skilling programs. From enrollment to placement, the Lighthouse program was expected to be a one-year cycle initially.

However, once we completed two years of the Lighthouse project, we looked at the data for the period. To our amazement, we found that of the 750 youth who enrolled in year one, the number of youth who were placed in jobs at the end of the second year, had more than doubled when compared to placements at the end of the first year. This data warranted a complete relook at the program in terms of the time horizon for planning and tracking impact of the program. We have therefore, shifted the lens to a five-year horizon, not just for tracking impact but also for providing programmatic inputs. In hindsight, it seems obvious that sustainable livelihood for youth growing up in extreme adversity has to be designed as a long-term project but it required analysis of data to identify the way forward.

Analytics is especially useful when it offers an insight that is counterintuitive. To our surprise, we found that of all the youth placed from the Lighthouse, two-thirds were self-placed and had identified their own job-leads. For such youth, the Lighthouse had helped unlock their sense of initiative to the extent that they decided to be resourceful about getting themselves a job. This metric upstages even the country's top MBA schools, where most students depend upon the institute for placements. The socio-economic need of the youth from low-income communities drives them to seek jobs the moment they sense the possibility. The extent of self-placement is a startling data insight because of which the program is being re-engineered to focus equally on self-identified leads as well as those generated by Lighthouse.

> Behavioral change is one of the biggest drivers of social transformation, the extent and impact of which can be measured leveraged data analytics tools and technologies. From enabling digital literacy to the impact of vocational training and school-based educational programs – data analytics is emerging as a strategic enabler of social transformation in the new digitally-driven world.

Data mapping enables scale for a Digital Literacy program

The Digital Literacy program run jointly by PMC, Pune City Connect, and the NASSCOM Foundation has the ambitious goal of enabling 100% digital literacy in Pune. In other words, one person in every household across low-income communities in Pune must become digitally literate i.e. have the ability to leverage the internet to improve their quality of life. When this program was launched, the challenge was primarily to procure buy-in from organizations and other entities in order to raise funds for running the digital literacy centers. Today, with 20 centers and 35,000 trained citizens, the challenge is to make strategic decisions on where to locate the next set of centers in order to ensure comprehensive coverage of all slum communities.
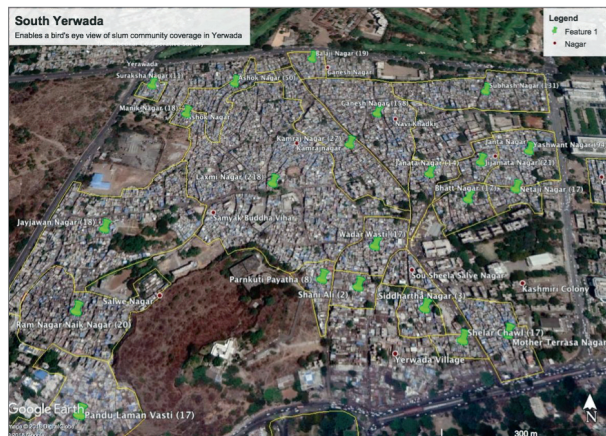

Figure 1: South Yerawada Slums

This is where Google Earth Pro comes in. Data pertaining to the number of citizens trained in each community is visually mapped on Google Earth Pro to identify the next set of geographic areas to focus on. *Figure 1* depicts the slum communities in South Yerwada (Yerwada is an administrative ward of Pune and has the highest number of slum community households, amongst the 15 administrative wards).

Areas enclosed in yellow-lines are slum communities. Within each community, it is possible to see the outlines of individual houses, and thus, approximately gauge the total number of households. The number next to the name of the slum community represents the number of households trained within that community. From the map and the available data, it is possible to pinpoint:

Communities where coverage is far less than 100% - In the South Yerwada area, for instance, the program is in the initial stages. Most communities have many households yet to be trained which will require more centers to be opened. In case a center cannot be opened, digital literacy buses can fulfill the need for smaller communities.

Location for the next set of digital literacy centers - Viewing the coverage information at the city-level helps highlight the wards or geographical locations for the next set of digital literacy centers. Without the visual data tool, making strategic location choices becomes extremely difficult and arbitrary.

Analytics drives impact assessment of educational programs in schools

The social sector depends on CSR funds and individual donations for running programs. Donors are eager to understand the impact of their funds, making it critical to collect and analyze data for assessing the impact. Pune City Connect has partnered with PMC and Leadership for Equity in the education space. The program is one where it is particularly challenging to assess impact as the program is designed to produce long term, systemic change.

Amongst other things, the program works with 1500 teachers in over 200 PMC schools to enhance their capacity to teach mathematics and Marathi. It took a few years for this kind of systemic change to translate into student learning outcomes, which in turn made it difficult to procure donations for the program. It is only now that we have concrete student data showing that 70% of students have moved up one 'level' in Marathi and scored 60% in mathematics, during the last academic year. This analysis was carried out using baseline and midline data. This is a significant shift. With such tangible results, there is hope of securing additional long-term donor partnerships.

Harnessing data for action and impact

Social transformation implies large-scale change and the lack of analytical tools makes it not only tedious but also, at times, infeasible to track and course correct. Any long-term social transformation initiative must undertake data analytics to first understand the efficacy of the program design and second, build a supportive ecosystem. Mission impact and social return on investment must be clear to sustain efforts over the long term.

There are many other possibilities for using data analytics to drive social transformation. These include facilitating citizen participation using social media to drive change, and leveraging predictive, risk, trend and big data analytics for decision making. While social sector organizations are typically short of funds and resources to leverage data analytics to its full potential, the good news is that a few resource organizations are assisting the social sector with data analytics - at no cost. This is an extremely valuable service that is hoped to rapidly gain both traction and scale.

ABOUT THE AUTHOR

Ruchi Mathur is the CEO of Pune City Connect *(punecityconnect.org),* a nonprofit working in partnership with Pune Municipal Corporation in the areas of sustainable livelihood, digital literacy, and education. She has over 23 years of experience spanning both corporate and nonprofit sectors. She has worked in varied functions such as marketing, HR, fundraising, strategy and CSR across several industries, including advertising, retail, IT, and the non-profit sector.

A post graduate from IIM Ahmedabad, Ruchi is also a Business Today awardee. In 2017, she was listed by the publication as one of the 'Most Powerful Women in Business' in the Social Impact category.

References:

- Census of India, Office of the Registrar General & Census Commissioner, India, Accessed February 2019, *http://censusindia.gov.in/2011census*

- Overview of Census 2011, Office of the Registrar General & Census Commissioner, India, Accessed February 2019, *http://censusindia.gov.in/2011-prov-results/data_files/mp/02Introduction.pdf*

- 'Census of India – Lessons learnt and way ahead', Presentation to the UN by Dr C Chandramouli, Registrar General and Census Commissioner India, Accessed February 2019, *http://www.un.org/en/development/desa/population/events/pdf/expert/23/Presentations/EGM-S2-Chandramouli%20presentation.pdf*

- 'A step-by-step Guide to the 2011 Census: Its Success, Failures and Questions', Published November 20, 2015, Accessed February 2019, *https://blog.socialcops.com/intelligence/data-stories/step-step-guide-2011-census-successes-failures-questions/*

- 'How reliable are India's official statistics' , East Asia Forum, Published April 2012, Accessed February 2019, *http://www.eastasiaforum.org/2012/04/06/nagaland-s-demographic-somersault-how-reliable-are-india-s-official-statistics/*

- 'NSSO Surveys along India's periphery: Data Quality and its Implications', Published November 2017, Accessed February 2019, *http://azimpremjiuniversity.edu.in/SitePages/pdf/APU-173194-Working-Paper-Series-9-15.11.2017.pdf*

- Unified District Information System for Education, UDISE, National University of Educational Planning and Administration, New Delhi. Accessed February 2019, *http://udise.in*

Making the Most of Today's Data Science Renaissance

Rob Thomas
General Manager, IBM Analytics

The new–age data renaissance makes now an ideal time to tap data to enable unprecedented business and social outcomes. Organizations can do it by leveraging open source technologies, adopting data governance standards, and making use of machine learning, data visualization, and cloud technologies to build competitive differentiation.

A variety of factors, coming together at the same time, can spark rebirth. The world of analytics is facing such a confluence of factors: economic disruption, rapid re-skilling, and unprecedented access to data. A combination of these factors is sparking the birth of data science, with the expert-led model becoming a relic of the past. History is a great teacher - it demonstrates that this data renaissance is not all that different from the original Renaissance.

During the Renaissance period, artists broke away from the norm and inspired a new era of expressionism. Ultimately, it was all about a new way of thinking, which sparked a period of extended innovation in the arts. In the late 1300's, art became economically-driven, with the wealthy using their riches to hire local artisans. This led to competition which, in turn, stoked creativity.

The modern day data renaissance is also about a new way of thinking, and draws many parallels from The Renaissance of many years ago:

It is economically driven as reducing the cost of compute, storage, and data have enabled the funding of new 'data artisan' enlightenment programs.

The new data artisans can be anyone, not just the wealthy few or those trained in a certain discipline, putting an end to the expert-led model of data science.

The application and confluence of data and science are being combined into a modern day vision of the future: continuous intelligence through the application of machine learning and deep learning.

The most important parallel is the positive impact that insights derived from data science are having on expression – whether in business, government or society – and the advancement of innovation. Five ways to take advantage of the data renaissance

Neither the renaissance in Italy nor the one in data science today would exist without a certain set of pre-existing conditions. In both cases, the market conditions enabled creativity and served as a launching point for future innovation. Today's data renaissance is determining the winners and losers in each industry, and those that adapt will survive, driven by application of a new technology.

Today, organizations aiming to harness data science must first build a prescriptive roadmap, which leads to success and a leadership position. Most companies, however, leap to model building and algorithm selection. While for some companies, this is the right place to start, for others, it may be too lofty a step. Why do organizations jump the gun? This is because most feel lost when it comes to leveraging data science. There are no guidelines, just a set of challenges to contend with.

Here are the top five elements that can enable businesses to take advantage of the current data renaissance.

- Leverage open source: Open source technologies that are free and open enable comprehensive analytics strategies.

- Adopt governance: Unified data governance technologies are necessary for insight and compliance. They organize, catalog, mask, protect, archive, and make any asset instantly findable. A data library not only provides insights but also ensures compliance with key regulations (such as GDPR).

- Manage data for the cloud: Hybrid data management solutions enable organizations to succeed in a multi-cloud world. The future of data management is one that will operate across cloud types, with seamless integration.

- Visualize your data: Data visualization tools are all about data discovery. They are designed to help people see their data dynamically, across pie charts and bar graphs, to unearth insights in real-time.

- Run machine learning and data science: Machine learning capabilities are used to create applications that learn on the go; the more the data that is fed into them, the smarter the applications become, enabling them to accurately predict outcomes. This capability is also the source of many data science 'a- ha' moments that lead companies to automate decision making and operations.

Data science in top gear: Two real world examples

Data science is coming into form, with machine learning use cases leading the way. Take, for example, a pharmaceutical company that was accustomed to new drug launches taking 12–14 years, at an average cost of USD 2.6 billion. Data science and machine learning helped the company reduce the cost by 70%. The machine learning algorithm was trained to learn from two distinct sets of drug data: one on the toxicity of various chemicals and the other on known side effects from approved medicines. From both datasets, the algorithm was able to predict the toxicity of the medicine with reasonable accuracy.

Or consider the case of USAA, the financial services firm that focuses on US military families. USAA adopted data science and machine learning not too long ago to improve customer service among other things. Today, the company leverages insights from machine learning to help predict why clients contact the firm, via phone or online. By using analytical insights from data collected in the past and combining it with dynamic data, they are able to discern why the person is contacting them, and quickly offer them more accurate responses.

The time to leverage data science is now

Let's go back to our parallel between The Renaissance and today's data renaissance. While Da Vinci was the first to correctly draw the Vitruvian man, history shows that he did not come up with this idea on his own. The credit for that goes to Giacomo Andrea da Ferrara, Da Vinci's colleague. Da vinci merely iterated Ferrrara's version to achieve success and history only remembers Da Vinci's version. The difference between the two men was a bias for action. Da Vinci took action, while Ferrara stayed in experimentation mode. The key takeaway here: any renaissance is a call to action, not a call for reflection. The time is ripe for organizations to leverage the data science renaissance for competitive success.

ABOUT THE AUTHOR

Rob Thomas is the General Manager at IBM Analytics. He is also the author of 'The End of Tech Companies', and 'Big Data Revolution'.

Rob Thomas has spent over 17 years in technology, with a focus on high tech and software. He currently leads Analytics software at IBM.

Prior to his career in technology, Rob had short assignments on Wall Street, which cultivated his interest in financial markets. Today, Rob is most interested in the intersection of technology and economics, as the two disciplines create a new set of market leaders and competitive disruption.

Rob is on the board of Domus in Stamford, Connecticut. Domus is a human services nonprofit which opened in 1972 and serves more than 1,100 of the most struggling youth in Stamford. Rob holds an MBA from the University of Florida and a BA in economics from Vanderbilt University.

References:

- 'Active machine learning-driven experimentation to determine compound effects on protein patterns', Research article, eLife, Published February 2016, Accessed February 2019, *https://elifesciences.org/articles/10047*

- AI Hits the Mainstream, Business Report, MIT Technology Review, Published March 2016, Accessed February 2019, *https://www.technologyreview.com/s/600986/ai-hits-the-mainstream/*

Propelling Saudi Arabia into the Next Millennium: The Role of Data Analytics in Realizing Vision 2030

Dr. Amal Fatani
Vice President & Head of the All Women Business Process and IT Services Center, Tata Consultancy Services

The continued progress and growth of Saudi Arabia is not a recent phenomenon. It is the result of long-standing modernization efforts, wherein each generation builds on the progress of its antecedents. For decades, Saudi Arabia has been working to improve the lives of its citizens, using the wealth generated from the Kingdom's natural resources to fund the social and economic development of the nation. Every measure of human development – life expectancy, infant mortality, literacy, per capita income, etc. – has improved dramatically in the span of a single generation.

Against this backdrop, in April 2016, Crown Prince Mohammed Bin Salman unveiled Vision 2030, an ambitious program that aims to drive further development of the Kingdom. The program seeks to build on Saudi Arabia's strengths as an investment powerhouse situated in the heart of the Arab and Islamic worlds, with strong geographic connections to Europe, Asia, and Africa.

Speaking about Vision 2030, the Crown Prince noted, "Our vision is a strong, thriving, and stable Saudi Arabia that provides opportunity for all. Our Vision is a tolerant country with Islam as its constitution and moderation as its method." Vision 2030 outlines 24 specific goals for the Kingdom to achieve in economic, political, and societal development. Vision 2030 further articulates 18 commitments to achieve these goals – with specific initiatives in renewable energy, manufacturing, education, e-governance, entertainment, and culture.

While Vision 2030 is a futuristic goal, leveraging the data available within the country can support the grandiose vision of the program and its execution across various areas.

Life Expectancy

The Saudi Arabian Government recognizes that a flourishing nation depends on the health of its people. The health care infrastructure is one of the best in the world and the government is continuing to invest in it. Due to major improvements in healthcare, the life expectancy of the Saudi people has dramatically increased from 52.7 years in 1970 to 75.05 years in 2015. The Saudi healthcare system has swiftly increased the number of hospitals and clinics by 9.5% in the last five years. Annually, the Ministry of Health organizes the construction of 150 new medical centers to serve the Saudi people, in line with its 10-year plan of offering nation-wide integrated healthcare. TCS is involved in integrating the data across primary health care centers and hospitals spanning the country to make it a seamless experience for its citizens

> Saudi Arabia is betting heavily on data analytics to improve life expectancy, infant mortality rate, literacy, and per capita income of its citizens in a bid to realize its vision 2030 – that of a thriving and stable country with ample opportunities for all.

Infant Mortality

The improved health care system supplemented with targeted data-based campaigns to spread awareness of disease, promote healthy habits for child-bearing women, and supplement the care of infants has resulted in a plummeting infant mortality rate. In 1960, there were 292 deaths per 1,000 live births in Saudi Arabia. By 2015, that number fell to 13 deaths per 1,000 live births. The result: a vibrant society with a healthy talent pool.

Adult Literacy

The Saudi Arabian Government is heavily invested in providing education to its citizens. It provides scholarships to high school graduates to pursue education around the world. In 1970, only 8% of the adult population of Saudi Arabia was literate. By 2014, over 94.4% of Saudi citizens were considered literate by United Nations standards.

Gross Domestic Product and Per Capita Income

While oil prices have been on a roller coaster ride over the last three years, Saudi Arabia has crafted economic strategies that have led to constant, unyielding growth. The IMF praised the Kingdom for playing "a key stabilizing role in the global oil market" by leading oil-producing nations in supplying oil during worldwide delays. The remarkable jump from a $42 billion GDP in 1970 to the astounding $753.8 billion in 2014 is evidence of the Saudi success story.

UN Human Development Index

To understand the extent to which the Saudi Government dedicates financial and material resources in the service of Saudi public, one only needs to examine the nation's ranking within the category of Human Development. In 1980, Saudi Arabia maintained a Human Development Index value of 0.575, in comparison to its 2014 value of 0.837. Since 1970, Saudi Arabia has been deemed a "Top 10 Mover" out of 135 countries that were evaluated.

Taking a data-led approach to accelerating growth and development

Saudi Arabia has taken a data-based approach to determining focus areas that can produce an impact on its citizens, and in turn, the social construct of the society. One of the initiatives in line with Vision 2030 was the introduction of knowledge-based industries to help the country diversify from an oil-based economy, with a focus on employing educated youth that graduate from the best universities – both inside and outside the Kingdom.

Aramco, Tata Consultancy Services and General Electric collaborated to open the All-Women Business Process & IT Services Center in Riyadh, Saudi Arabia. The center provides a multitude of services such as Human Resources Operations, Supply Chain Management, Finance & Accounting, IT, IOT, Master Data Management, as well as analytics to customers from various local sectors - including the government, Oil & Gas, manufacturing, and telecommunication. Women employees at the center in Riyadh also support General Electric and its affiliates in more than 55 Countries around the world.

The next chapter in Saudi's Arabia's explosive growth story

The next phase in TCS's contribution to the Kingdom is utilizing data Analytics and insights gained therefrom to support the development of the job market, especially in rural areas, utilizing our experience in the country over the last five years.

In a country that is currently reinventing itself all over again, the role of data analytics is crucial to understanding how to future-proof all sectors of the economy for the next millennium. The multifaceted development plan to realize a brilliant future hinges on stakeholders acquiring a better understanding who they are, where they have been, and where they want to be. The country and its people are looking forward to that journey with bated-breath.

ABOUT THE AUTHOR

Dr. Amal Fatani is Vice President, All Women Business Process & IT Services Center, Tata Consultancy Services (TCS) , Riyadh, KSA. Set up by Saudi Aramco as a joint venture between TCS and General Electric, Dr.Fatani has helped the center grow from a strength of 400 to nearly 1000 females employees as of 2017. Prior to joining TCS, she served as Consultant and General Supervisor of Female Affairs at the Ministry of Higher Education, Riyadh. Dr. Fatani set up, managed and ensured growth of female centers at the Ministry and King Saud University (KSU). At KSU, she was appointed Chair Person of the Female Pharmacology and Toxicology Department, and went on to become the first Vice Dean of Female College of Pharmacy. She then served as the first Dean of the Female Center for nine scientific and medical colleges, and set up female sections in all sectors of the Ministry of Higher Education. She regularly follows up on women affairs in scholarship programs and Saudi universities.

Currently, Dr.Fatani is a Board Member of Saudi Human Rights Commission and has represented Saudi Arabia in many national and international events.

She is also a founding board member of Saudi Research Science Institute at King Abdula University for Science and Technology, member of the International Affairs Advisory Committee at Al Faisal University, consultant to King Abdulaziz and his companion's foundation for giftedness and creativity named "Mawhiba".

She has worked closely with the Ministry of Economy & Planning, Ministry of Labor & Social Development, and Commission for Job Creation and Employment, to support implementation of the ambitious Saudi Vision 2030.

Dr. Amal Fatani holds a Ph.D. in Pharmacology and Toxicology from the Department of Pharmaceutical Sciences at The University of Strathclyde, UK. She obtained her Master`s and Bachelor`s degrees in Pharmacology and Toxicology from the College of Pharmacy at King Saudi University, Saudi Arabia.

Tapping into Social Registries for Data-led Social Development

Roopa Purushottaman
Chief Economist and Head of Policy Advocacy,
Tata Sons Limited

Joshua Abraham
Special Projects Lead, Senior Manager,
Macroeconomics and Policy Advocacy, Tata Sons Limited

Naman Narain
Deputy Manager, Macroeconomics and Policy
Advocacy, Tata Sons Limited

Digital social registries enable democratization of data to enable local and national-level social interventions in a highly effective and cost-efficient manner. Social registries work on data platforms that comprise a unique identification number with associated demographic and communications data – an approach that minimizes manual errors, intervention, and saves valuable time and cost.

Villages in Chandrapur, a district on the Maharashtra-Telangana border, pose significant challenges for social program delivery. Not only are they poorly connected by transport and telecom infrastructure, the local populace also communicates in diverse languages - including Marathi, Telugu, and other dialects. In 2015, Tata Trusts partnered with data intelligence company SocialCops, to create a comprehensive data-based micro plan for these 290 villages across 80 development indicators.

In just three short months, data scientists were able to develop a composite database that allowed government officials to identify households with no grid connectivity, ration cards, LPG connections, toilets, or access to processed tap water. How did the data scientists achieve this incredible feat? They created customized surveys for different demographic conditions on an Android app and linked it to social registries. They could also map households that were eligible for skilling or subsidized insurance programs. This data-driven model now enables Chandrapur's remote villages to participate in programs that were previously inaccessible to them, as well as connect to larger district and state strategic development planning efforts.

The role of social registries in social transformation

From Guinea to Chile, Turkey, Djibouti, Pakistan and India, 'social registry' based approaches are helping people to connect to public services such as free health insurance, social security, financial inclusion and food or fuel subsidies, by prioritizing the poorest. The data platforms catering to the registries comprise a unique identification number with associated demographic and communications data. The simple database saves billions of dollars in wasted spending that could stem from manual processes and errors.

Pakistan's social registry now includes 85% of the population, serves 70 different programs and has generated a savings of $250 million. In Argentina, linking 34 social programs to unique identification numbers revealed eligibility errors, saving USD 150 million over eight years. In India, the migration of benefit programs to Aadhar is expected to deliver significant savings and considerably higher coverage, while effectively addressing social needs as well as privacy and safety concerns.

Social registries exemplify four key pathways by which data and analytics can drive social change:

The use of a digital identity improves both the efficiency and effectiveness of social programs. It minimizes wasteful spending by ensuring that only the eligible receive social assistance. Healthcare, which forms a substantial portion of public spending in the developed world, is a prime example. Digitization of medical records using patient data from genomic sequencing, sensors, and social media, enables insurers and governments to get deeper insights into which treatments work. This is critical given the growing demand for 'value-based' reimbursement.

Patient data and digital tools will eventually give rise to new services that would not be possible without data—services that might involve taking no drugs at all. For instance, United Healthcare, the leading American insurer, offers prevention programs that connect pre-diabetics with special coaches at gyms. Paired with wearable devices that gamify exercise and persuade people to walk a certain distance every day, it provides a far cheaper solution for insurers and governments than years of visits to doctors and hospitals, and drug usage.

In the case of social registry, connecting with every household in a geography that falls below the poverty line would not be possible without the data platform. Many of the impoverished and marginalized will likely receive public aid for the first time under these systems. Flow-based lending systems that enable extension of credit services to the 'unbanked' is another example. By instantaneously generating a credit score based on Micro, Small and Medium Enterprises' (MSME) digital transaction history, they provide credit facilities to small entrepreneurs even in the absence of a formal credit history. Predictive analytics also plays a key role in pre-empting both natural and man-made disasters – from predicting food shortages to identifying high-risk diabetes cases via population health monitoring.

Social problems are often complex and nuanced, mandating solutions that require a high degree of contextual customization. However, what needs to be customized may not always be clear to change-agents. Data and analytics allow for low-cost trials and rapid experimentation, enabling quick identification of the 'right' approach. The universality of data makes it easy to update stakeholders, especially policy makers and funding organizations, on administrative and financial priorities. An ongoing partnership between Copenhagen Consensus Institute and Tata Trusts does just this. It establishes a clear rupee-benefit-per-public-rupee-spent measure for a list of well-documented interventions to critical problems in the states of Rajasthan and Andhra Pradesh. From farm productivity to public health and education, interventions are evaluated and ranked in the context of the state's needs, providing a simple yet effective guide to the government and NGOs.

Data's ability to accurately and persuasively assess the impact of interventions enables greater alignment amongst various stakeholders – governments, funding and international organizations, local NGOs and civil society - on what works for each social problem in its unique context. This can spur greater collaboration, more fruitful partnerships, and a truly multi-disciplinary approach to social change.

Data democratization will lead to highly effective social interventions

Free and accurate data significantly lowers the costs of 21st century social programs in various ways. These include minimal wastage in benefit targeting, generating instantaneous feedback on the effectiveness of an intervention, creating low cost digital service-delivery models, and enabling cost effective transfer of learnings to other contexts.

At the same time, it's important to understand that development problems rarely exhibit universally commodifiable characteristics for easy scaling. For example, Karnataka Learning Partnerships' approach to assessing the quality of primary education by combining data from community feedback, local panchayat contests, and district-level DISE data cannot be transferred to Bihar, where the challenges of primary education are very different. Identifying a reasonable scale for sanitation, health, low female workforce participation and affordable housing may be even tougher, with funding structures, epidemiological profiles, and attitudes varying at a community as well as the aggregate level.

Despite this, data and analytics are equipping social change agents with invaluable tools for their intractable challenges. The democratization of data will lead to highly effective local and national-level social interventions. While this may be a slow process, it could very well prove to be the legacy of our digital age.

ABOUT THE AUTHORS

Roopa Purushothaman is the Chief Economist and Head, Policy Advocacy, Tata Sons Limited. She is widely recognized for her path breaking research on BRIC countries, and prior to joining Tata Sons, Roopa has held leadership positions at several leading firms - from Goldman Sachs to Everstone Capital Advisors. She is also the founder of Avasara Leadership Institute, a non-profit educational institution focusing on accelerating academic and leadership outcomes for high potential adolescent girls in India. Ms. Purushothaman is a graduate from Yale University and a postgraduate from the London School of Economics.

Joshua Abraham is a part of the Macroeconomics and Policy Advocacy team at Tata Sons Limited. He leads thought leadership and research efforts centered on India's development priorities. Prior to joining Tata Sons, he spent seven years as a management consultant with A.T. Kearney India, delivering multiple large-scale strategy consulting engagements across Europe, South-East Asia, and India. Joshua's vast experience includes long-term strategy development, large-scale transformations and operations improvement primarily in consumer goods, online and offline retail and automotive sectors. He holds a Bachelor's degree in Electrical Engineering from Indian Institute of Technology, Bombay.

Naman Narain is a part of macroeconomics and policy advocacy unit at Tata Sons Limited. Previously, he worked as an Investment Analyst at Kalaari Capital, where he evaluated investment opportunities across healthcare, enterprise, consumer internet, and other sectors. He is a graduate of London School of Economics and St. Xavier's College, Jaipur.

In Conversation with Prashant Girbane & Hitendra Singh

Prashant Girbane
Director General, MCCIA and Director (Hon.),
Pune International Center (PIC)

Hitendra Singh
Head, Social Innovation Lab,
Pune International Centre.

Kindly describe your current role.

PG: As the Director (Hon.) of PIC – a liberal and independent think-tank that focuses on policy research across multiple domains including Defence and National Security, Economics and Growth, and Social Innovation, my role involves working closely with the President, Vice-President, and other trustees of PIC in delivering various activities such as research, programs, events, round table conferences, and development of policy papers, in the domains listed above.

HS: I work on the social innovation aspect at Pune International Centre and help social enterprises (both for-profit and not-for-profit) that use an underlying innovation to solve a social problem. We do this by running a nation-wide social enterprise discovery platform called National Conference on Social Innovation (NCSI),

a mentorship program called Social Enterprise Mentorship Programme (SEMP), a shared service platform eponymously called SSC for Social Enterprises, and by carrying out attendant research on various topics of relevance.

What are you and your organization doing to make a positive impact on the society and environment?

Ans: PIC leverages the strength of its more than 400 strong members and their participation, network, and goodwill to create and execute its research and various programs. The Social Innovation vertical, in particular, is dedicated towards enabling improvement in the most marginal sections of society. We believe the present day technology and innovation allows for newer products, services, and processes to emerge that will enable us to provide the solutions to our marginal people that didn't exist before.

What are the typical data challenges you face while doing this work?

There are three key challenges:

1. Availability of reliable data: More than the dearth of data, the fact that whatever data we do have exists only in silos makes it challenging to access data at the time and point of need.

2. Finding data that is usable: Often, data that we do find exists in formats that are technically incompatible, which makes that data unsuitable for further analysis and interpretation.

3. Contradictory information: It is not uncommon that data and statistics from different social organizations contradict each other. At such times, given our scattered data ecosystem, it becomes challenging to determine who is the custodian of which data, further impacting availability and reliability or authenticity.

What is the relevance of data and analytics in the social work that you are doing?

Data is one of the core elements of doing social good and that's because in order to do anything impactful, one needs to consider three key dimensions:

1. Base lining activity data: This involves answering some basic questions like - how many people have access to education or healthcare, what are the outcomes of any previous intervention, or say what is the state of tribal or women welfare, etc.

2. Data on the social innovation ecosystem: This includes information such as which institutes are there, what kind of research they do, what are the outcomes, how research ideas come to become final products, etc.

3. Data on market reach: This includes information on what are the different network organizations one can partner with, how they can add value, etc.

How is a NGO/Social Enterprise using data & analytics different from the one which does not leverage the same?

An NGO/Social Enterprise that leverages data and analytics is better placed to optimize spend where it is most required, track its impact, and measure progress to benchmark goals and objectives. Such a NGO/SE is also more likely to partner intelligently with other NGOs/SEs or government bodies to widen reach and efficacy of social initiatives.

What challenges did you face while implementing a data driven approach in your organization? Was there any kind of resistance to change?

At PIC, since most of us come from techno-managerial backgrounds, we were very clear that leveraging technology is the best way forward to make incremental social impact. We track all our programs and their outcomes through a rigorous data framework.

However, the challenges arise as the level and frequency of data updation varies across organizations, further leading to availability and reliability issues that we already discussed.

How does a NGO/SE differ from a corporate entity while taking a data driven approach?

NGOs and most SEs do what they do without attaching the profitability aspect to it. Corporate entities on the other hand, are mandated by the CSR law or business related aspirations to tap into rural economies or marginalized population base, leveraging a data-driven approach. The NGOs traditionally have been carrying out extensive field surveys that help them in demonstrating their work and also in reporting it back to the donors. Though, given the lack of reporting standards or consistent practices, in the NGO domain there is wide-disparity in how data is collection, analyzed, and reported.

What specific advantages did you derive through a data driven approach?

Leveraging this approach, we have been able to build a robust data ecosystem with participation from around 400 members. We have been able to make our decision-making more objective than allowing us to sway by our perceptions. Being data-centric has helped us in evolving the overall PIC brand, and specifically within the Social Innovation space, we have been able to identify our niche, find suitable partners, and track our work better.

What role did technology play in this transformation journey?

Technology plays one of the most important roles in enabling social transformation as the basic need of any SE or social innovator is to have a portal that supports data collection, standardization, and acts as a hub to connect social innovators. Technical resources as well as platforms can help bridge this gap that NGOs and governments have not been able to do, due to lack of expertise, time, and resources. Without losing continuity and certainty of the job at hand, technology firms can contribute by engaging a small set of resources to manage the data related tasks such as aggregation, cleaning, sanitization, uploading, and much more. Academic institutions also have a role to play by getting involved in these projects and enabling research of longer-term projects that may get sidelined in the more market driven framework. Specifically, academic institutions can enable a robust talent pool of potential people willing to work on such projects so that the quality and speed of work is never compromised.

Kindly describe your personal experience while pursuing this transformation journey.

PG: On the personal front, I feel deeply honored to be working alongside renowned social innovators like Dr.Mashelkar and Dr.Kelkar to drive large-scale social impact. I particularly feel that by giving the youth a platform to channel innovative, progressive thinking in all aspects of public life, the PIC is making long-term investments in ensuring sustainability and growth of our nation.

HS: I share Prashant's sentiment; it is a dream-come-true to be able to work with, and learn from some of your role-models like Dr. Mashelkar and Dr. Kelkar. I feel grateful to have found this opportunity where some aspects of my training and skills can be put to use in helping drive better quality of life and prosperity in our society.

ABOUT PRASHANT GIRBANE & HITENDRA SINGH

Prashant Girbane is the Honorary Director at Pune International Centre (PIC), a global think tank based in Pune, India. He actively contributes to PIC publications that include both national and state public policy papers. Mr. Girbane also serves as Director General of Maharashtra Chamber of Commerce, Industries & Agriculture (MCCIA), President of IIM Ahmedabad Alumni Association, Pune Chapter, and is associated with management institutes such as Symbiosis and MIT through their 'academic councils'.

He has previously worked with Tata Consultancy Services in the domains of Capital Markets and e-Governance across geographies including the UK, the US, Switzerland, Singapore, and Hong Kong. Prior to joining TCS, Mr. Girbane served as a National Consultant at United Nations in India for its Development Programme (UNDP). Mr. Prashant Girbane holds a Bachelors degree in Chemical Engineering from UDCT (now ICT) and an MBA degree from IIM, Ahmadabad.

Hitendra Singh is a researcher and Social Entrepreneurship enabler at Pune International Centre (PIC), a global think-tank. He teaches hands-on entrepreneurship to school-going children and considers entrepreneurship to be a great life skill. Prior to joining PIC, Hitendra worked with a technology company servicing the commodities and currency desk of a leading London bank. He then went on to become a partner in a London-based Distressed Debt fund that managed over USD 1 billion in assets. In in his last corporate role, Hitendra led as the COO of a USD 40 million company in Dubai. After turning entrepreneur, he started India's first Alternative Investment Fund (AIF) distribution platform for IFAs in Mumbai. He is an alumnus of Indian Institute of Management (IIM), Ahmedabad and Delhi University.

GREEN
ECONOMY

Agriculture 4.0: New Pastures for Big Data in the UK

Paul D. Wood

Data & Analytics Strategy Advisor, A&I, TCS UK

According to the UK Government's Industrial Strategy, tackling low productivity growth is the most important economic challenge facing the country. While the UK's agricultural industry boasts of some of the world's best farmers and growers, the productivity growth in the country's agriculture and horticulture has failed to keep pace with that of major competitors. The Agriculture and Horticulture Development Board (AHDB) estimates that if the rate of growth of UK farming had kept pace with the US since 2000, its contribution to the rural economy would have been £4.3 billion higher by 2013.

The AHDB has consequently identified five focus areas to drive growth in productivity:

- Powering research and innovation
- Creating an evidence base for farming decisions
- Setting up a coordinated knowledge exchange (KE)
- Providing training to improve skills
- Enabling farmer-to-farmer learning
- The case for Big Data in agriculture

As technology pervades every aspect of the farming value chain, the information revolution is making inroads into agriculture. Sensors, GPS data, and internet access to a myriad other sources allow farmers to measure, quantify, and analyze data to enable on-going improvements in an unprecedented manner.

The Big Data revolution in agriculture comes from the expansion of small data i.e. the growth in farmers' ability to collect data from soil sampling sensors, farm machinery telematics, GPS, and farm aerial imagery in combination with weather and farm management datasets. One conceptual example of farm data decision making is analysis based on information in the genetics-by-environment-by-management scenario or GxExM relationship.

The utilization of data can guide decision making not only at a farm level but along the entire lifecycle - from field to fork. Data from farms aggregated into a community is more valuable than the data from any one farm individually. However, in the UK, farmers are not well served by aggregated data providers - a situation that contributes to the productivity challenges. Data analysis is further compounded by the challenges in querying and combining the Department for Environment, Food and Rural Affairs (DEFRA) data with other sources such as weather data and individual farm management data.

The challenges do not end there - encouraging farmers to make individual data available for the greater community good is a key obstacle.

A new digital agricultural revolution, Agriculture 4.0, is taking hold as Big Data analytics, robotics, sensors, and drones become a part of the farm-to-fork value chain. In the UK, Agriculture 4.0 promises to revolutionize productivity, boost efficiencies, and scale margins by a big count.

Policy changes: A potential productivity driver

The UK government recently published A Green Future: Our 25 Year Plan to Improve the Environment, 2018 – a 25 year environment plan that makes the sustainable use and management of land a key goal. Further, with the UK exiting the EU, there is an opportunity to re-design the Common Agricultural Policy (CAP) which presently has two pillars: payments for land area (accounting for approximately 75% of CAP payments) and payments for public goods (accounting for approximately 25% of CAP payments). For instance, maintaining a wildflower meadow strip around a field to attract pollinating insects comes under the purview of 'public goods'.

The UK government is currently considering the form that 'UK CAP' should take in the post-Brexit era. The likely outcome, consistent with the 25 year environment plan, is a greater proportion of 'UK CAP' payments for 'public goods'. The claim process for CAP payments will require data and this policy can prove to be a catalyst for creating a consolidated data collection and analytics application in agriculture in the UK. This could be modeled on the lines of the Ag-Analytics "Farm meets Big Data"TM application in the US which performs a related function for US farmers as a crop insurance calculator and what-if analyzer.

On the road to agricultural innovation

The UK government realizes that with Brexit, the need to unlock the hidden potential of food supply chains to boost production efficiency, transparency, and consumer trust is more pressing than ever. At the National Farmers' Union (NFU) annual conference in February 2018, the government announced a £90 million fund for agri-tech. This investment is designed to bring together AI, robotics, and earth observation to improve supply chain resilience in the agri-food sector. Agrimetrics: the Big Data platform for the agrifood sector is another UK based platform that aims to boost collaboration between government, academia, and the industry – with a mission to solve the UK's food challenges. Drones are also making inroads into England's agricultural scene, enabling automated farming, disease detection, and harvesting.

UK agriculture is ripe for disruption

What the country's agricultural industry requires most today is effervescent leadership – that can guide farmers to work innovatively, attract investment from across the supply chain, and provide a blueprint for a sustainable agricultural system. A coherent agricultural policy, backed by evidence-based decision-making and greater adoption of digital technologies can pave the way for a highly productive 'farm to fork' experience.

ABOUT THE AUTHOR

Paul D. Wood works as a Data and Analytics Strategy Advisor in the Analytics and Insights unit at TCS, UK. Previously, he has served as the IT Managing Director in a Global Investment Bank where he led numerous business transformation technology projects. Paul holds a Horticulture (Plant Science) degree from the Eden Project in Cornwall, UK.

References:

- 'In rural England, a fleet of drones monitors and harvests crops', Article, Wired Magazine, Printed January 2018, Accessed February 2019, *https://www.wired.co.uk/article/future-of-farming-and-agriculture-drones*

- Agrimetrics, Company Website, Accessed February 2019, *https://agrimetrics.co.uk/home*

- '£90 million to revolutionise tech in UK agriculture', Article, Open Access Government, Printed February 2018, Accessed February 2019, *https://www.openaccessgovernment.org/business-secretary-announces-90-million-revolutionise-tech-within-uk-agriculture/42961/*

- Ag-Analytics, Company website, Accessed February 2019, *https://ag-analytics.org/*

- 'A Green Future: Our 25 Year Plan to Improve the Environment', Report, Department of Environment, Food and Rural Affairs, United Kingdom, Published 2018, Accessed February 2019, *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/693158/25-year-environment-plan.pdf*

Big Data for Bountiful Grape Production

Dr. Jayram Maruti Khilari
Ex-President of Maharashtra Rajya Draksh Bahaitdar Sangh

Precision viticulture is helping India's grape growers to leverage Big Data technologies for managing variability in their vineyards, acquiring highly fertile soil, and boosting productivity in a cost-efficient manner.

In 2017, 214,440 metric tonnes of grapes were exported from India, of which Maharashtra's share was 95%. As of July 2018, 350,000 acres of land (2.01% of the total area) in Maharashtra has been allotted for grape cultivation. Maharashtra's spectacular contribution towards surplus production and export of grapes is the result of hardworking farmers backed by data-driven and scientific agricultural techniques.

The scenario in the past

Previously, India lagged other dominant grape producers of the world like Chile and United States as the farmers lacked access to the varieties of grapes that are in demand in the global market. Additionally, insufficient labor, ever increasing labor costs, and expensive pesticides and fertilizers further added to high production costs. To overcome these issues, growers have now adopted scientific approaches and are using data analysis to promote automation and increase the export produce

Today, grape growers with the help of scientists and the Indian Government are trying to import new patented and non-patented seedlings. This is helping them magnify the grape export figures multi-fold. There has been a 40-50% YoY growth in grape export since the activity started in India in 1986-87.

How precision viticulture helps grape growers boost efficiencies

Precision viticulture – a precision farming technique applied to optimize vineyard performance can help grapes growers increase potential yields, maximize profitability, and enable them to respond to market opportunities with agility. The tools of precision viticulture are used to collect large amounts of data on the vineyard and the attributes of individual production areas at a high spatial resolution. A number of enabling technologies including GPS, GIS, VTR, soil sensors, yield monitors, and remote and proximal crop canopy, can aid in understanding the variation in both crop and soil. Recognizing these variations and managing different areas of land differently can help growers exert more control over their production system and volume. Precision viticulture can also potentially result in efficient use of water and fertilizers, targeted management of nitrogen pollution, and maintenance of soil fertility.

Leveraging Big Data, can help farmers better manage the variability in their grape production system by tailoring inputs to desired outputs. This helps growers address sustainability issues by optimizing productivity in the productive parts of the landscape while conserving biodiversity and the natural resource base in the less productive ones.

ABOUT THE AUTHOR

Dr. Jayram Khilari is the Ex-President of Maharashtra Rajya Draksh Bahaitdar Sangh (MRDBS), Vice Chairman, Research Advisory Committee at Maharashtra State Grape Growers' Association, and Retired Soil Scientist from the Mahatma Phule Krishi Vidyapeeth, Rahuri, Maharashtra.

Famously known as the Grape Scientist of India, Dr.Jayram has been working for over 45 years in the areas of viticulture R&D in Maharashtra He has received several prestigious awards, including 'Draksha Mitra' by Maharashtra Grape Growers' Association and 'Best Grape Scientist' award by Krishibhushan Vasantrao Arve Research Development Trust. Dr.Jayram holds a Master's degree in Agriculture and a Ph.D. in Viticulture.

Data-Driven Water Management: Lessons from Pani Panchayat

Kalpanatai Salunke
Managing Trustee, Gram Gaurav Pratishthan

Dr. Gurudas Nulkar
Professor of Management, Symbiosis Centre for Management

While the Indian economy has consistently grown over the last few decades, the problem of water scarcity remains unresolved. Various establishments like central and state water resource ministries, local governments, NGOs and international institutions have been engaged in addressing this issue on various fronts, and yet, there has been no respite to farmers and rural communities. Even as villages struggle with water deficits, market driven cropping pattern decisions of farmers lead to inequitable distribution of water and worsen the problem. The sugar industry concentrated in the drought prone area is a glaring example of this mismanagement.

A key reason why this happens is in the way decisions on policy and social investments are made in India. Broadly speaking, policy decisions are based on

- the agenda of the ruling political party

- demographic attributes

- the negotiating power of lobbyist groups

- maintaining status quo (do nothing).

In each of the situations, the final decision may not be the optimal choice, and in each case, the decision is prone to subjectivity and personal biases of decision makers.

On the other hand, if decisions are based on data, there is a higher probability of optimizing and contextualizing them. The Pani Panchayat has believed in this approach to assist villages in managing their water resources, by minimizing the chances of favoritism and discrimination through the use of data. Late Vilasrao Salunke, the founder of Pani Panchayat, laid the concept of water as a community resource, which must be conserved, managed, and shared by the community. This evolved during the drought of 1972. Though the micro watershed model was developed and demonstrated on Naigoan Farm, the extension program was limited to Pani Panchyat community lift irrigation schemes based on minor irrigation projects or percolation tanks. The basic data taken into consideration for the rightful entitlement of water was: the total area of the village, average rainfall, and the size of the population. The requirement of a family for protective irrigation to sustain on agriculture was also considered. The water allocation per capita was set at 1000 meter cube per capita, with a minimum of 5000 meter cube per family.

A case study of Khalad Panchkroshi

Agriculture is the biggest consumer of water and hence cropping decisions significantly impact the water demand of a village. In a free market economy, consumer demand drives production, and producers have the freedom to choose what they produce. Farmers across the country have discarded traditional cropping wisdom which was based on rainfall and climatic conditions and have resorted to growing what markets demand.

Leveraging a data-driven water management approach helped Pani Panchayat, an NGO, to alleviate the long-standing water crisis in Khalad Panchkroshi. Integrating data-driven policy making with resource management is the ideal way to overcome short-term challenges, while future-proofing long-term welfare and sustainability.

Data	Unit of measurement	Data source
Population	Number	Census data, Gram Panchayat
Total area of the watershed	Hectares	-do-
Area under cultivation	Hectares	-do-
Irrigated agriculture	Hectares	-do-
Wasteland	Hectares	-do-
Forest	Hectares	-do-
Sources of water (Well, borewell, farm ponds)	Number	Actual data
Livestock population	Number	Animal husbandry
Installed water pumping capacity	HP	MSEDCL
Energy use from agricultural meter connections	KW	MSEDCL

Sr. no	Particulars	Unit of measurement
1	Average annual rainfall in last 5 years	mm/annum
2	Runoff water (24% of rainfall)	CUM
3	Underground water storage (10% of rainfall)	CUM
4	Water stored in soil (3% of rainfall)	CUM
5	Surface water tanks, nala bunds (3% of rainfall)	CUM
6	Evaporation losses by crop use (60% of rainfall)	CUM
7	Water available for use (3+4+5)	CUM
8	Water available per capita	CUM/capita
9	Water required per capita (ideal)	CUM/capita
10	Shortage of water in the watershed	CUM
11	Shortage of water per capita	CUM/capita

This has resulted in a multi-fold increase in the water requirement of the village, which exceeds the natural supply in the watershed. Moreover, in rain shadow and drought prone areas, the water availability per capita is far below the minimum required for livelihood.

The Khalad Panchakroshi faced two droughts – in 1972 drinking water was available but food was in shortage, while in 2002 food was available but no potable water. Because of climate change (global warming) on one hand and the changed cropping pattern on the other, there was accelerated deterioration of biodiversity, water, and land in the area. Pani Panchayat believes that the solution to this problem lies in decentralized planning, program implementation, and governance based on Panchkroshi as the unit of planning.

Panchkroshi Khalad is a cluster of five villages in the Purandar taluka of Pune district. It is a part of two micro watersheds, the BM-58 and BM-59, of Urvarit Bhima river basin. It comprises five villages – Kumbharvalan, Ekhatpur, Munjavadi, Khanvadi to the North of the river and Khalad, lying to the south of Karha river.

Pani Panchayat considered the irrigation water demand as a critical factor in addressing water sufficiency. This required an objective assessment of the water needs of the communities that inhabit the micro watershed. For this, we accessed various secondary data sources and primary data collected at the Panchkroshi level. From this, we could calculate the quantity of water available and the water consumption per capita. Statistical analysis clearly identified the drivers of water deficiency for the Panchkroshi communities. [3]

It was evident that the conventional cropping pattern of the Panchkroshi must be changed to reduce the water shortfall. To achieve this objective, Pani Panchyat proposed that farmers divide their production into two categories: self-consumption and market produce. Doing so would reduce the financial requirement of buying food for individual households. The household's basic needs would be provided for through

intensive cultivation on a small plot (about an acre per family). During droughts, it would be easier to provide irrigation to this concentrated area, reducing the burden on the farmers. Furthermore, it was proposed that the Panchkroshi villages evolve a common cropping pattern, based on traditional knowledge and experience, to ensure more efficient use of water and sustain the competitive power of markets.

The data-driven approach proved to be more effective in instilling faith in the villagers, rather than having them trust opinions and assumptions made by an organization outside their village. Empowered with this analysis, Pani Panchayat is trying to convince the Khalad Panchkroshi to create a micro watershed level policy on farming to resolve the problem of water deficiency. For this initiative to succeed, the five villages must cooperate - rather than compete, for the overall benefit of all stakeholders.

Data analysis

Pani Panchayat identified the key data on which the water management was based upon. The data was compiled for the five villages which make up the Panchkroshi. *(See table 1).*

The next step is to compute the water balance in the watershed. The parameters for this are shown in *Table 2*.

The hydrological zones of the watershed were characterized according to physiography, geology, hydrology and suitability. The data was then analyzed to establish a methodology for water use, which is consistent with the characteristics of the watershed.

The data yielded two key indicators –

- rainfall versus the water availability for the watershed and
- water use in the watershed.

Benefits of data-driven resource management

- In developing and implementing the Panchkroshi model of water management, Pani Panchayat has successfully leveraged the benefits of a data-driven decision support system. Some of the key lessons learnt included:
- Villagers place more trust on objective assessments of data, especially when the sources of data are trustworthy. We used census data, IMD rainfall data, and actual primary data collected from their own villages.
- Water footprint and ecological footprints of villages within a watershed are useful indicators in creating awareness on resource use. They remove the subjectivity and any chances of favoritism.
- With data backup, the Panchkroshi model empowers villagers to solve their own problem of water deficiency, instead of placing it in the hands of irrigation engineers or local politicians.
- When cropping decisions are placed in the hands of the beneficiaries, it is easier to form and implement a system of restraints which promises a greater common good, while considering the needs of all stakeholders.
- The bottom line: Overcoming short-term challenges without compromising on long-term welfare can be achieved by integrating data-driven policy making with resource management.

ABOUT THE AUTHORS

Kalpanatai Salunke is the Managing Trustee of Gram Gaurav Pratishthan, a charitable trust that is deeply engaged in work related to sustainable rural development through equitable distribution of water. Established by her late husband Shri Vilasrao Salunke, the trust promotes organic farming, healthcare and education to villages. A social reformer in her own right, Kalpanatai is actively involved with villagers and farmers to bring the changes in the water usage by practicing water equity through the Equitable Water User Forum.

Dr. Gurudas Nulkar is a professor of management at Symbiosis Centre for Management and also serves as a trustee of the Ecological Society, where he teaches environmental economics and sustainable development courses. Dr. Nulkar has authored several books on sustainable development and the key focus areas of his research include management of natural resources, and usage of water in agriculture and industrial ecology. He holds a doctorate in industrial sustainability.

Harvesting Data and Analytics to Revolutionize Agriculture

Dr. Srinivasu Pappula, *Ph.D., CISSP*
Global Head, Digital Farming Initiatives, Tata Consultancy Services

More than 65% of the Indian population depends directly on agriculture for their livelihood. Yet, pressing agricultural challenges stemming from climate and market related uncertainties threaten to jeopardize the future of food security in India. Leveraging analytics to transform Indian agriculture can help change this equation.

Making Indian agriculture 'climate and market smart'

TCS Digital Farming Initiatives (DFI) aims to make Indian farmers and agriculture 'climate and market-smart' through the extensive use of data and analytics. Our market-demand and crop-protocol centric Crop Rotation, Optimization and Planning System (CROPS) forms the hub of the revolutionary initiative. CROPS uses data-driven decision making techniques to advise farmers on the best crops that can be grown on their land, thereby optimizing field-level management and resources.

This involves considering various static data sources such as the historical cropping patterns, geographic climate zones, weather patterns, soil composition, water tables, and water availability maps. In addition, dynamic data elements such as individual farmer constraints, market forces, and predicted prices are also factored into the equation. Through the application of deep learning techniques across disparate data sets, the system generates a shortlist of the best crops that can be grown in a certain area. It also generates the cost of production and the expected income for each option. Once a farmer selects the desired option, CROPS creates a Crop Protocol for the chosen option.

The Crop Protocol comprises personalized information for each farmer. Following the protocol instructions helps farmers achieve the expected yield while maintaining the predicted quality. The revolutionary protocol-centric approach helps farmers better anticipate and smooth out the unexpected vagaries of weather and market using historical and predictive data analytics.

Optimizing crop protocol in real time

During the crop cycle management phase, data from real-time monitoring is analyzed to fine-tune the crop protocol. Here's how:

The first step is to determine the optimal sowing period based on the Moisture Adequacy Index (MAI) - the standardized measure used for assessing the adequacy of rainfall and soil moisture required to meet the potential water requirement of crops. Historic climate data spanning over 30 years is used to determine this sowing window.

Thereafter, real time monitoring of the weather conditions and application of various AI-based techniques helps create an early warning system for pests and diseases. This helps farmers enhance their farm income in two ways.

Smart farming and precision agriculture involve the use of Big Data and predictive analytics to analyze humongous amounts of field and crop performance data to optimize the crop protocol in real time. Indian farmers are taking to smart agriculture with the ultimate goal to develop a probability model that can help them weather rough times, improve efficiencies, and generate superior profits.

One, by obtaining advance intimation on the incidence of pests and diseases, farmers are also able to avoid reactive usage of chemical pesticides, saving money as well as the environment. Second, they are able to reduce wastage through proactive disease prevention. The DFI group at TCS has created models for forecasting the incidence of 80 diseases and pests for 30 different crops.

Cognitive Remote Sensing Services in Agriculture (CRSSA), the imagery acquired by satellites, is the cornerstone of CROPS. The type of imagery, more specifically its wavelength, plays a key role in the nature of the agricultural problem being addressed. For instance, crop growth measures (such as leaf area index, stomatal activity, chlorophyll content) and soil health metrics (such as soil moisture, salinity, pH) need to be correlated with various wavelengths and scaled up for using satellite imagery. In crop identification, specific crop growing practices are considered when using satellite imagery. For instance, identification of rice across large watersheds can be based on the ponded conditions under which most rice is grown. Crops such as oil palm can be identified by applying machine learning algorithms using their unique leafing pattern as a signature. A combination of spectrum choices and plant phenology-based markers can help in the development of enhanced versions of CRSSA.

Enabling smart farming through predictive analytics

The true potential of analytics lies in its ability to crystal gaze into the future and accurately predict success. Deep learning algorithms can analyze a decade of field data comprising insights into how crops performed in various climates and inherited certain characteristics, and use the information to develop a probability model for predicting which genes will most likely contribute a beneficial trait to a plant. Machine learning algorithms can also be used for grading and sorting produce in a low cost manner.

Clearly, the time for precision agriculture is ripe and data and analytics will play a key role in ushering a new era in Indian farming.

ABOUT THE AUTHOR

Dr. Srini Pappula works as the Global Head of Digital Farming Initiatives Group at TCS. He is responsible for driving the usage of Information and Communication Technologies (ICT) for improving farmer livelihoods across the world through the creation of rural nuclei of growth called PRIDE™s (Progressive Rural Integrated Digital Enterprises). His ambition is to create technology-powered avatars of the Gandhian Gram Swarajya, thereby fostering village-level self-sufficiency, employment, innovation and growth while making farming a profitable profession.

With over 28 years of experience, Srini possesses varied expertise in the technology sector and has previously worked as a Senior Product Leader and General Manager at Intel Corp, California. He has also co-founded and served as the Chief Executive Officer of successful start-ups in the wireless security space in the Silicon Valley.

Srini holds a Bachelor's degree in Electronics Engineering from the Indian Institute of Technology and advanced degrees in Computer Science and Management from the University of Texas and University of California respectively. He is also a Certified Information Systems Security Professional (CISSP).

How Data-Driven Social Transformation Can Uplift Urban Slums

Pratima Joshi
Co-founder, Shelter Associates

Resham Badri
Architect and PR Manager, Shelter Associates

Ross Plaster
Architect, Shelter Associates

Shelter Associates (SA), a Maharashtra-based civil society organization is leveraging data and analytics to realize the dream of 'One Home-One Toilet' (OHOT) across the state. Leveraging open source data mapping techniques, the organization empowers Urban Local Bodies with powerful insights to improve the quality of sanitation in urban slums.

India's urban population has grown dramatically due to rapid migration from rural to urban areas. One of the major consequences of this population shift is the growth of informal settlements within the country's urban areas, creating several challenges for Urban Local Bodies (ULB) .

The rapid influx has put a severe strain on an already inadequate ULB infrastructure, resulting in environmental deterioration as new informal settlements become established and existing ones become denser. The deterioration is occurring at all levels: the home, the neighborhood, the ward, and the wider urban area as a whole. While remedial projects are being deployed by ULBs in an effort to address the problems, they are proving to be unsustainable, largely due to absence of granular data to inform realistic solutions. In addition, the use of a top-down approach hampers the participation of beneficiaries during the process.

A civil society organization's solution to combating the challenges faced by ULBs

One of the key hindrances to effective delivery of services to the poor is the reluctance of government bodies in institutionalizing data and updating it regularly. Consultants are only appointed on a project basis to track information. Over the years, several consultants have logged information that remains unused. As officers constantly move or transfer, there is no continuity in the use of data. Over a period of time, the data becomes inaccurate, irrelevant, and therefore, unusable.

 Shelter Associates (SA), a civil society organization (CSO) established in 1993, has been working with the residents of informal settlements in urban areas across Maharashtra. The organization's goal is to design and implement impactful projects that achieve tangible and long-lasting change while elevating the quality of life of India's most impoverished people. Creating spatial information to enable informed decision making is an integral component of SA's methodology.

Through various projects, SA has consistently emphasized the importance of data and analytics as the foundation for delivering services in a systematic, phased, and targeted manner. According to the organization's Co-founder & Executive Director, Ms. Pratima Joshi, one of the major stumbling blocks in the early years was lack of comprehensive and detailed information as even the government lacked accurate data reflecting ground realities. SA quickly understood that to make meaningful interventions, they would need to get on the ground and collect primary data themselves.

Figure 1: Data collection on smartphones

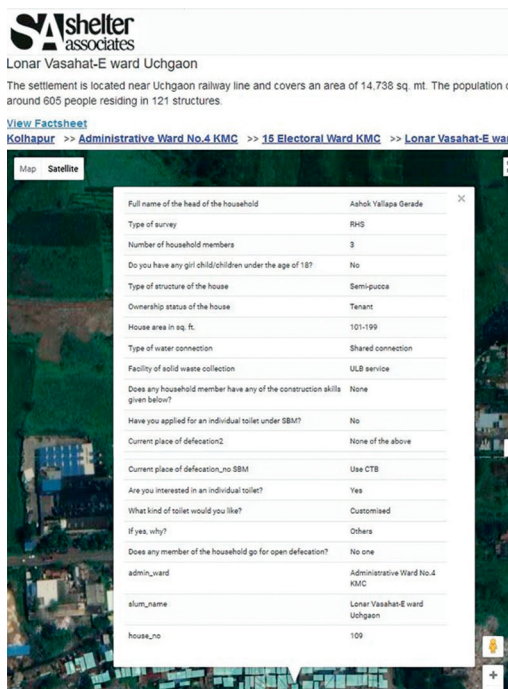Figure 2: Structural mapping using Google Earth

Figure 3: Household level data is linked to respective households spatially on the GIS platform
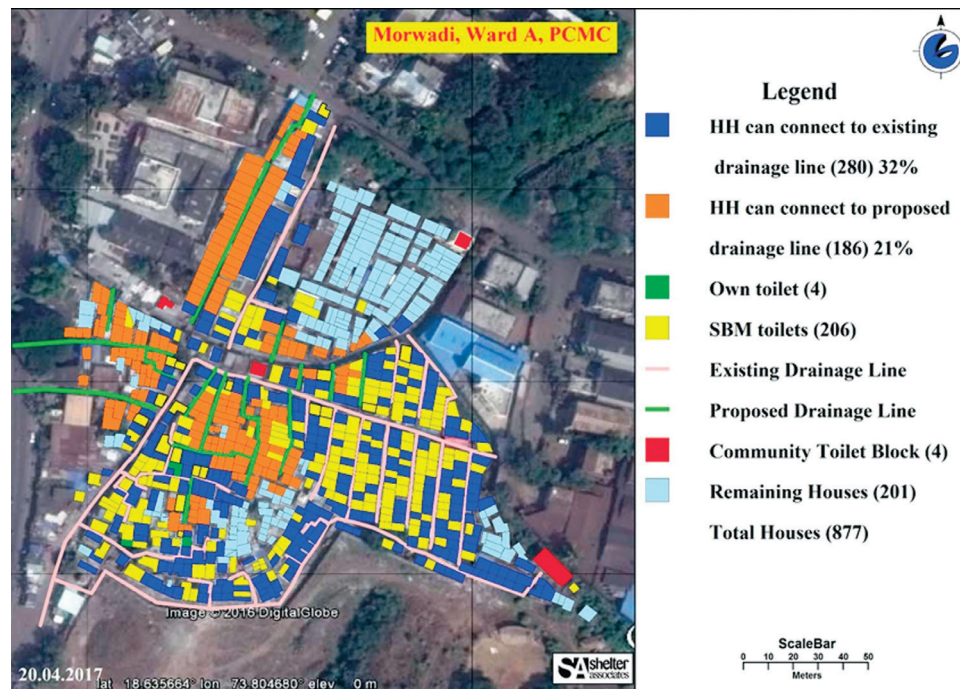

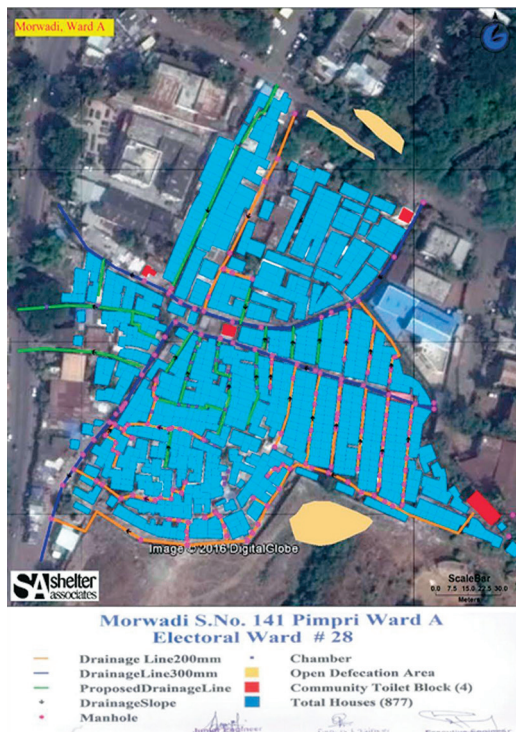Figure 4: Slum level map highlighting phased interventions

Figure 5: Drainage map validation from ULB based on drainage network (existing & proposed)

SA adopted a data-driven approach using surveys to identify high priority gaps and found that access to sanitation by urban slum dwellers was the most neglected service across most cities. To address this challenge, SA has been successfully evolving a scalable, replicable, and sustainable model called 'One Home-One Toilet' (OHOT). It uses a unique data-driven, community-centric, and cost sharing approach, through a partnership between SA, the ULB, and community members, to facilitate household toilets in urban slums. Since 2013, the 'OHOT' model has been implemented successfully across six cities in Maharashtra and facilitated over 14,000 household toilets, impacting over 3,60,000 individuals directly and indirectly. This model is also aligned with the Government of India's program 'Swachh Bharat Mission' (SBM).

How SA leverages data and analytics to drive the OHOT model

The first step towards providing sustainable household sanitation is maintaining a comprehensive spatial database by leveraging technology platforms such as Geographic Information Systems (GIS), Google Earth (GE), and mobile technology. SA pioneered a solution that uses GIS and GE for mapping poverty. A GE map serves as a base map for Rapid Infrastructure Mapping (RIM) where the settlement level infrastructural systems such as drainage networks, water supply, roads, community toilet blocks, waste management systems and so on are marked on the map - along with the existing structures - using unique numbers. A Rapid Household Survey (RHS) is then conducted to gather household level information such as family details, water connections, sanitation facilities, and area of the house, through survey forms *(see Figure 1 and 2)*. SA uses a web-based application called Kobo Toolbox, available on an open source Android platform to collect slum and household level data.

The survey data and mapping information is collated and integrated spatially onto a GIS platform and analyzed to strategically target households lacking in household sanitation *(see Figure 3)*. Community mobilization activities to generate awareness and demand for household sanitation are also tracked and monitored on the platform for efficient service delivery.

The data collected is made available on SA's data portal for open viewership. The data is further validated by each ULB and linked to its website *(see Figure 4 and 5)*. SA has successfully created a city-wide spatial database for cities like Pimpri-Chinchwad and Kolhapur. Data for other cities like Sangli, Pune, Navi Mumbai, and Thane, are also on the data portal.

The key challenge is not only facilitating household toilets, but also strengthening the drainage networks in cities during the implementation of such initiatives. The spatial data shared with ULBs enables informed decisions, and ensures accountability and transparency. The data leverages budgetary provisions for laying, repairing, and augmenting existing drainage networks in a phased manner to generate economies of scale. The data can also be used for effective delivery of other services.

The impact: Successful household sanitation installations at scale

SA has successfully installed over 1700 household toilets within a span of one year in Balaji Nagar - one of the largest slums in Pimpri Chinchwad *(see Figures 6 and 7)*.

With the help of SA's city-wide spatial data, ULBs have been able to scale household sanitation deployments under SBM and strengthen drainage infrastructure within slums *(see Figure 8)*.
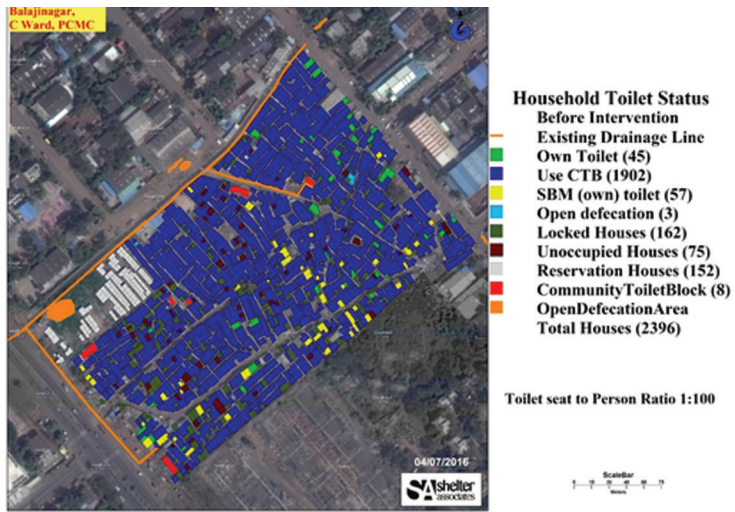
Figure 6: Pre-intervention status of Balaji Nagar
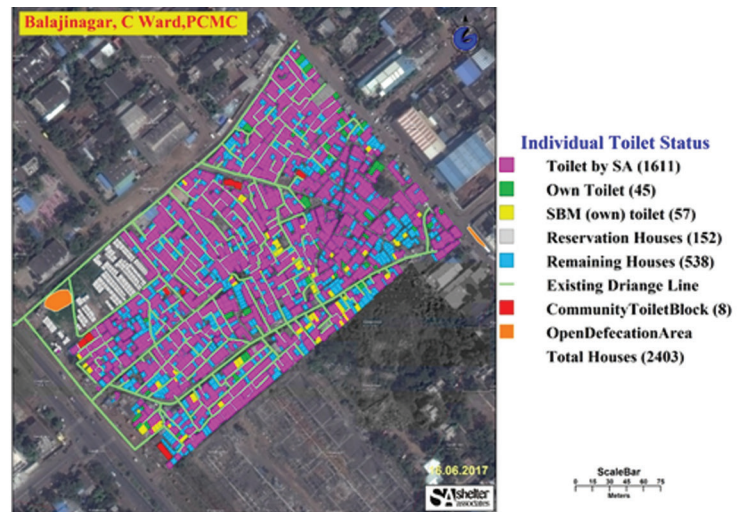
Figure 7: Post intervention status of Balaji Nagar

	Pimpri-Chinchwad	Pune	Kolhapur
% of HH toilets	9% → 36%	27% → 57%	22% → 40%
% of drainage coverage	30% → 64%	90% → 90+%	14% → 42%
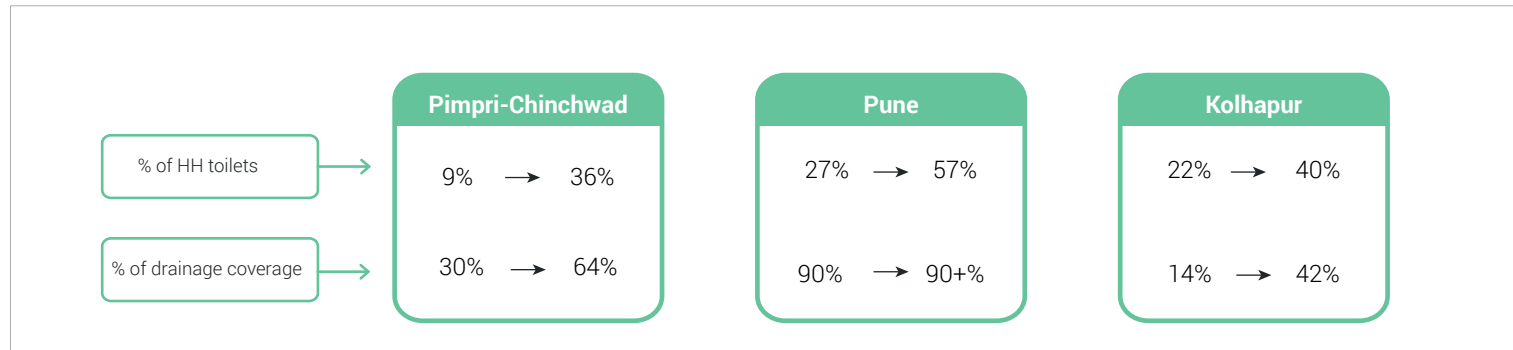
Figure 8: Impact of SA on Swachh Bharat Initiative

SA's successful methodology has received global recognition. The organization has received several awards, including the Google Earth Hero Award (Only Indian NGO to receive this award), HUDCO award for sanitation, as well as the NASSCOM Social Innovation Forum 2017 award under the 'Environment' category. Recently SA also bagged the India Sanitation Coalition-FICCI's special recognition award for innovative use of spatial data to deliver sanitation.

SA is set to launch an easily replicable toolkit that leverages the OHOT delivery model. A third party impact assessment of SA's OHOT model and a policy review paper that compares the SBM model versus the SA model, with recommendations to make the SBM model more robust in implementation, have also been compiled.

While many cities are undergoing rapid transformation, it is imperative for urban planners, policymakers, and managers of urban affairs to reimagine their approach leveraging data and analytics to drive more robust and holistic policies. Such an approach can help strengthen the infrastructure and accelerate transformative service delivery for social transformation.

ABOUT THE AUTHORS

Pratima Joshi, an architect by profession, is one of the co-founders of Shelter Associates (SA), which aims to convert slums into housing societies for the poor by providing access to basic services like water, sanitation, and electricity. An Aga Khan scholar, Ashoka fellow and Google Earth Hero, Pratima has received several accolades for her work at SA over the last 25 years.

Resham Badri heads the Public Relations vertical at Shelter Associates. She aspires to offer skill sets, be able to reach out to communities at the grass root level, and make an impact on their habitats and livelihoods through the spread of knowledge. Resham holds a Bachelor's degree in Architecture.

Ross Plaster is Architecture Consultant for Shelter Associates. He provides architectural, project management, and practice management services to support the development of data collection systems to facilitate efficient and effective community participation to ensure the appropriateness of the designs and the long-lasting impact of the projects. He is also instrumental in designing and implementing the NGO's various slum rehabilitation projects, both housing and essential services. He holds a post graduate certificate in professional practice from University of Bath, UK, and master's degree in Architecture from Glasgow School of Art, UK.

References:

- 'United Nations Expert Group Meeting On Sustainable Cities, Human Mobility And International Migration', Research report, Published September 2017, Accessed February 2019, *http://www.un.org/en/development/desa/population/events/pdf/expert/27/papers/V/paper-Bhagat-final.pdf*

- Squatting Rights: Access To Toilets In Urban India', Research report, Published March 2013, Accessed February 2019, *https://www.dasra.org/resource/improving-urban-sanitation*

- 'Infrastructure Development Strategies for Inclusive Growth: India's Eleventh Plan', Report by G.S. Joshi, ASCE Library, Published March 2010, Accessed February 2019, *https://ascelibrary.org/doi/full/10.1061/%28ASCE%29LM.1943-5630.0000057*

- SA uses the term 'informal settlements' to refer to both notified and non-notified slums; notified slums are eligible for welfare from the UrbanLocal Body (ULB) but non-notified slums effectively do not exist from the perspective of the ULB.

Using Data to Bridge the Educational Divide for Girls

Safeena Husain
Founder and Executive Director, Educate Girls

Educate Girls, an Indian non-profit organization leverages data and analytics to tackle the rampant gender inequality in the country's educational landscape and measure the impact of its initiatives. The organization goes by the overall goal of balancing a human-centric approach with advancements in technology to increase girls' access and uptake of quality education.

Millions of girls across India are denied an education, with social and cultural barriers preventing their enrolment, hampering their progress through school, and perpetuating gender discrimination across thesystem. For the team at Educate Girls, an Indian non-profit that aims to tackle gender inequality in education, the mission is three-fold: identify out-of-school girls, enable their villages to find a solution for their enrolment, and address the learning deficit of the children.

But before we could tackle the complex issues around girls' education and enable social transformation, we needed to find answers to a few key questions. Who are we serving? What is their need? How do we understand what is working and what is not? Finally, how do we know if, and how, we have been successful? The fact is we cannot answer any of these questions without carefully collecting and analyzing data.

Let's take a closer look at each of the key questions and how data helps answer them.

Who do we serve and what is their need?

We conducted an extensive door-to-door survey, covering over three million households across Rajasthan and Madhya Pradesh to identify all out-of-school girls, and created a targeted approach to their enrolment. Just to understand the magnitude of the operation, our monthly data collection for village and school profiling stands at over one million data points each year.

Armed with this data, we inscribe our logo on the wall of every house that's part of our program and assign it a number. By doing so, we are able to pinpoint the location of out-of-school girls to better understand the barriers to their enrolment. As we scale up, we want to ensure that we stay as accountable to every girl who comes into our program as we were to the very first one. This demands an individualized approach without losing sight of the efficiency and codification necessary when working with such a large number of children. It's easy to see why data sits at the heart of our work - not just in terms of what information we collect, but also how we analyze it, and how quickly we use what we have learnt.

How do we keep a tab on what is working and what's not?

We have digitized our pen and paper based monitoring and evaluation processes and deployed acustomized mobile app. Our field staff uses the app to collect and process necessary information for program management and decision making. On the other hand, our performance management system is fed with data pertaining to the location of our staff and their work schedules, as well as information gathered from the field about meetings, surveys, and school assessments. This builds greater transparency and accountability, besides ensuring rapid transfer of data.
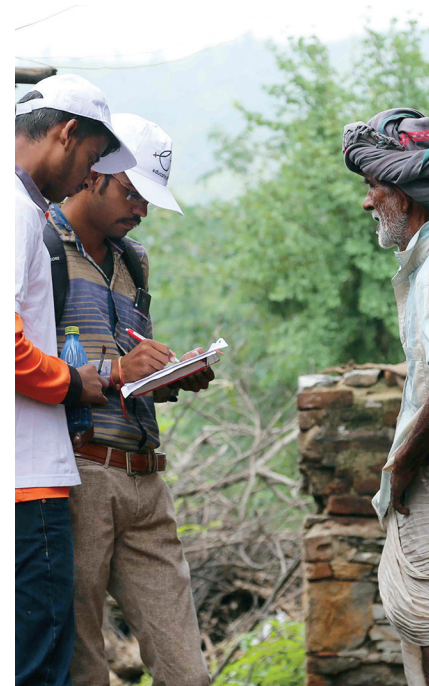
Once the girls are in the classroom, we work with government schools to deliver a child-centric and gender-sensitive curriculum. We collect data around planning, classroom observations, and testing. The data is then used to identify children who are struggling to understand the challenges they are facing, what resource might be needed to help them, how the field staff is performing, and what additional training might be required. We also assess whether home visits and additional lessons are needed for absentees.

How do we know if, and how, we have been successful?

The purpose of data collection is to drive program improvements and results, rather than just accountability and transparency. Simply considering a 'before and after' scenario in terms of results is not sufficient. To continue to improve our performance and reach more children, we must understand our path to success. The detailed performance information helps us do just that. Thanks to our digitization initiatives, the speed at which we gather data, analyze it, and send it back to the field, has reduced dramatically - paving the way for rapid course correction.

Betting on Big Data for a bigger impact

Our goal is to expand to new states and reach 16 million children over the next five years. To achieve the ambitious targets, we are leveraging advancements in Big Data. We are also considering Machine Learning (ML) to assist in our planning. We are sure that technology will help us answer the big question – what if we could reach the girls who are most in need of education faster and cheaper? For instance, our recent door-to-door survey data revealed that 50% of out-of-school girls came from just 10% of the villages. We then compared the data with readily available census and other government data to see if we could build a predictive index of the geographical whereabouts of girls most vulnerable to be kept away from school. Armed with a way of predicting where our target communities are, we should be able to lower our costs and hasten our impact significantly.

The team at Educate Girls realizes that we have a data-driven opportunity to change the course of girls' lives by tackling gender inequality and India's literacy divide. We are highly optimistic about the future as we balance a human-centric approach with exciting advancements in technology to increase girls' access to quality education.

ABOUT THE AUTHOR

Safeena Husain is the Founder and Executive Director at Educate Girls – a non-profit organization, established in 2007, that aims at tackling issues at the root cause of gender inequality in India's education system.

Safeena has spent over a decade working with grassroots projects in Latin America, Africa, and Asia and holds a graduate degree from the London School of Economics.

In Conversation with Vinod Kamat & Devsen Kruthiventi

Vinod Kamat
Head of Programmatic Technologies, Tata Trusts

Devsen Kruthiventi
Chief Technology Officer, Tata Trusts

At Tata Trusts, leveraging technology to drive innovation and 'give back to the society' is the underlying theme that defines everything we do. The organization leverages real time monitoring and data analytics to overcome challenges in social transformation and enable superior outcomes in healthcare, IT, and other domains.

Kindly describe your current role.

As technology leaders at Tata Trusts, our role involves driving technology and innovation to enhance quality of lives of communities. The focus has been aligning organizational requirements with strategic and operational part of IT team. With extensive experience in IT delivery, performance and knowledge management we have been able to leverage high quality technologies to speed and scale up programmes. This has enabled Tata Trusts design, develop and implement innovative programmes across health, education, environment and agriculture.

What are you and your organization doing to make a positive impact on the society and environment?

With our core philosophy 'giving back to the society,' our aim is to make positive difference to the society and nation. Through programs such as Internet saathi, Tata water mission, Lakhpati kisaan, Sewage treatment programme, augmented reality for crop disintegration, and water-less urinals; Tata Trusts has been able to drive society

empowerment with focus on digital innovation. In fact at Tata Trusts programs are prioritized and funded based on ranking methodology. Once the programs are choosen their effectiveness is measured through automated monitoring. Right now we have more than 800 active programs that are monitored in real-time.

What are the typical challenges you face while doing this work?

The strategic focus of Tata Trusts is making real-difference to the communities through innovation in areas of healthcare, water and sanitation, energy and environment. However, implementing these programs is a challenge given the lack of technology partners. For instance to enhance Internet and telecommunication connectivity in rural areas, we have deployed costly satellite communication. On the other hand for successful program implementation change management is crucial. Point in case is Internet Saathi program where people were reluctant to enroll. We leveraged capabilities of local population to make people understand the benefits of enrolling in the program.

What is the relevance of data and analytics in the social work that you are doing?

Success in project means being on the positive side of the stakeholder, the end-user and the organization's benchmarks. At Tata Trusts, one of the most powerful ways to measure programme success is through real time monitoring. We collect data from existing measures and conduct impact analysis of quality data. This also provides deep insight into cost of program analytics and new paths of the program.

How is a NGO/Social Enterprise using data & analytics different from the one which does not leverage the same?

Through process automation and analytics, Tata Trusts is able to provide swift and sustainable benefits to the beneficiary. As a part of our migrant empowerment initiative, we help migrants with unique identity cards. Through automation we were able to expedite migrant registration process, enabling them to find occupation for a better livelihood.

What challenges did you face while implementing a data driven approach in your organization? Was there any kind of resistance to change?

Tata Trusts' dynamic leadership and the team of millennial employees make it easy to embrace data-driven approach. We have taken various change management initiatives to help our team embrace and learn new ways of working with data. Though during our data-driven governance initiatives in villages the locals found it difficult to equip with the technology but through right training and education we have been able to iron out the resistance.

How does a NGO/SE differ from a corporate entity while taking a data driven approach?

Leadership vision in an NGO is aligned to bring sustainable change in society and environment. While data analytics dashboards used in an NGO and for profit enterprise are same, the purpose is different. NGO is driven by a collaborative motive so that their programs can provide maximum benefit to the beneficiary.

What specific advantages did you derive through a data driven approach?

Leveraging a data-driven approach, we have been able to enhance organizational decision-making. For instance we have been able to improve project planning and get improved insight into zonal meetings through data analytics. Our team processes data to align it to government development indictors to arrive at expected completion of the program. As NITI Aayog is supporting development of 85 districts, data has become eyes of the government to understand the ground realities.

What role did technology play in this transformation journey?

Tata Trusts takes an integrated approach to technology and has been adopting automation and data analytics to improve the pace of development in underdeveloped areas. We have partnered with TCS for multiple initiatives, one being leveraging TCS Digital Nerve Centre to transform healthcare in Karnataka.

Kindly describe your personal experience while pursuing this transformation journey.

It is been uplifting to experiment and identify breakthrough technology to enable social development. Our IT team at Tata Trusts has actively pursued automation and real-time monitoring for program prioritization and assessment to implement the program the way it was envisioned.

ABOUT VINOD KAMAT & DR. DEVSEN KRUTHIVENTI

Vinod Kamat is the Head of Programmatic Technologies at Tata Trusts. He is responsible for identifying and deploying cost-effective cutting-edge technologies for the social development programs driven by the organisation across multiple themes. He is also responsible for developing applications for programmatic process automations and developing data analytics capability for providing deeper insights in to the impact generated through the implemented social initiatives.

He has over 30 years of Technology experience spread across diverse business domains ranging from Electronic Media, Finance, Travel, Logistics, Satellite Communications, Telecommunications and Philanthropy.

Prior to his association with the Trusts, he was the Vice President, IT Delivery and Performance Management at Vodafone India. In the initial part of his career he served as a Class I Gazetted Officer, Doordarshan India, Mumbai under Ministry of Information and Broadcasting. He was also instrumental in setting up of VSAT Communication facility for National Stock Exchange of India for facilitating equal access stock trading.

Vinod is a Master of Technology in Telecom Networks from Indian Institute of Technology, Kanpur. He has also pursued Post Graduate Diploma from Indian Institute of Technology (IIT) Kanpur in Video Communications. He graduated as an Electronics and Communications Engineer from B.V. Bhoomaraddi College of Engineering & Technology, Hubli, Karnataka University. He is a Certified Workplace Coach from Neuro Leadership Institute, UK.

Dr. Devsen Kruthiventi, Chief Technology Officer at Tata Trusts is a Ph.D with an overall experience of 28 years, out of which he has spent 13 years in academic and research at Dept. of Mathematics and Computer Science at Sri Sathya Sai Institute of Higher Learning, prior to joining Satyam Computer services. He worked with Satyam Learning center and Business Intelligence and Data warehousing teams in various capacities and then later moved to Tata Chemicals as Head, Knowledge Management. Successfully led Tata Chemicals to become MAKE (Most Admired Knowledge Enterprise) ASIA and Make in India winner in 2012. His last assignment was with Tata projects as their CIO. He has consulted various Tata Companies in their IT and Knowledge Management journey. He has been awarded "Knowledge Management Leadership Award" in 2014. He is visiting faculty at Indian Institute of Management, Ahmedabad, GlobalNxt University, Malaysia, and ASCI.

TECHNOLOGY & INNOVATION

Harnessing Data and Analytics to Leapfrog Innovation in Social Transformation

Balaji Ganapathy
Head of Workforce Effectiveness, Tata Consultancy Services

With the advent of the Fourth Industrial Revolution, Business 4.0™ - a thought leadership framework encompassing next-gen technology, is helping companies grow and deliver transformative outcomes through the power of digital. Enterprises are adopting Artificial Intelligence (AI), analytics, data, blockchain, cloud, cognitive computing, Internet of Things (IoT) and other technologies at a much faster pace than possible in prior innovation cycles. At the same time, the economic, political and social trends over the recent past are leading to an increasingly polarized and fractured world.

Given the societal tensions, a solution that balances the pace of growth with equitable distribution of benefits must have a foundation of trust and collaboration. The social sector commands the highest trust level among citizens, making it an ideal place to begin change. Going forward, cross-sector efforts aimed at social transformation will increasingly become the norm, with digital technologies powering not just incremental but leapfrog innovation. Various agencies will work in unison to efficiently identify and solve the world's most difficult issues, using the power of technology to customize, implement, and scale solutions at a rapid pace.

Steering social transformation: Technology meets purpose

Nonprofits play the pivotal role of being the hands-on, grassroots players, delivering valuable and often lifesaving services to beneficiaries. The scale and impact of social transformation is directly dependent on the nonprofits' ability to be effective in their beneficiary engagement and service delivery. However, most nonprofits are confronted with an existential dilemma. On the one hand, they are expected to devote a majority of their resources to core programs. On the other, tools, resources and technologies that aid decision-making are often considered as dispensable 'overheads'.

Data – gathered from beneficiaries, stakeholders, networks, and feedback loops should be at the center of every nonprofit's measurement, evaluation and reporting processes. However, most nonprofits lack digital capabilities, and data when available, is usually unstructured and inert. Quality and reliability of data are also major concerns as the social sector lacks defined data standards. What's more, there is no incentive to share data across nonprofits competing in the same space. Add to this, concerns like beneficiary rights, privacy and security, and the social sector faces a conundrum of sorts.

Despite the challenges, the power of purpose combined with the power of modern technology promises to upend conventional norms of what is considered possible in the realm of social change. In particular, the pillars of Business 4.0™ – unprecedented intelligence and mass personalization - can exponentially grow the positive impact on beneficiaries. Whether it is combating hunger, poverty, infant mortality, disease and disaster, or improving education, human rights, or the environment, intelligent decision making can serve as the catalyst for transformational change.

Leveraging key pillars of the Business 4.0 framework - data intelligence and mass personalization, businesses can successfully mine data and analytics to enable social transformation at scale.

Data intelligence takes the driver's seat

Imagine a community of social transformation stakeholders bound together not just by a deep sense of purpose but also by a shared, open, data architecture, with the ability to perform data mining, run visualizations, model behavior, and drive predictive analytics. What if we could build beneficiary personas, track operational metrics, and analyze historical results? The data could be used to model scenarios, predict success rates, and compare predictions with real world instances. Harnessing such intelligence can help prioritize resources, personalize services to beneficiaries, and maximize impact.

If this 'social science meets digital science' approach sounds more like science fiction to you, consider a few examples that prove otherwise:

- Crisis Text Line, an Indian suicide helpline, analyzed 65 million text messages to identify a path breaking insight - that the use of the word 'EMS' in a text was five times more predictive of a high risk of suicide than the word 'suicide' itself. Using this intelligence, the helpline prioritized incoming messages and was able to respond to 94% of high-risk cases in less than five minutes.

- My Choices Foundation, a non profit protecting women against abuse, is using Big Data analytics to power Operation Red Alert and help alleviate human trafficking. The solution analyzes India's census, government, education, and other data such as drought conditions, poverty level, proximity to transportation stations, to identify the places that are most at risk of human trafficking.

Data and social good: A match made in heaven

The time is ripe for powering social transformation using data intelligence and mass personalization to drive high impact social transformation. But technology alone cannot provide all the answers. We need to ask the right questions, define the problems, use appropriate tools, and validate with real world instances. For the new approach to flourish, it's equally important to build privacy, security, dignity and trust amongst beneficiaries and stakeholders. With empathy and digital embedded at its core, such a model can drive large-scale social transformation, simultaneously benefiting humanity and society.

ABOUT THE AUTHOR

Balaji Ganapathy is the Head of Workforce Effectiveness. He oversees Human Resources, Business Consulting, Diversity & Inclusion, and Corporate Social Responsibility for Tata Consultancy Services (TCS) in North America. His role includes coaching sales and business teams for growth, driving talent management and employee retention initiatives, architecting the corporate sustainability strategy, and promoting an inclusive culture. Under his stewardship, TCS is using technology innovation, thought leadership and skill-based volunteering, to impact the state of STEM education in North America - with a special focus on impacting women, girls, minorities and under-represented groups. Balaji was awarded the 2017 Charlie H. Moore Award for Leadership in Corporate Community Engagement by CECP, the CEO Force for Good, and also presented with the 2017 Million Women Mentors Male Champion of Change Award.

Balaji serves as the Chair of STEMconnector®'s STEM Innovation Task Force (SITF), Chairperson of NPower's TCC Advisory Council, and Vice-Chair of the Million Women Mentors (MWM) Leadership Council. He is also part of the Executive Committee of IMPACT 2030, World Economic Forum's Steering Committee, and the US chamber of Commerce's Education, Employment & Training Committee (EETC).

Scaling Social Transformation with Analytics

Venkatesan Sukumaran

Head - Business Analytics, Analytics and Insights, Tata Consultancy Services

As the Big Data deluge continues, with its use cases ranging from enhancing marketing and political campaigns to detecting national security threats, analytics and insights have become paramount to informed decision-making.

Though application of Big Data for social change represents a new trend, it offers tremendous potential to make a positive global impact. It can help address some of the most endemic problems ranging from poverty, disease, and ecological harm to war and famine.

Diverse data sets from mobile phones , sensors , and point-of-sale systems can not only help assess the most pressing problems facing the world but also ensure effective and targeted last mile solution delivery. Leveraging linked cross-agency data can help government to target social services more accurately and relevantly for its people. For instance, the widespread usage of Aadhaar combined with consumption and population data enabled the Indian Government to ensure savings of millions of rupees through targeted cooking gas supply as well as weed out dubious beneficiaries of such government sponsored schemes.

Let's take a look at some of the most interesting areas of social transformation where Big Data and analytics can play a definitive role in improving the lives of people.

1. Optimize reach, efficacy, and cost of public education

Education plays a critical role in any country's development, but ensuring 100% access to it is far from easy. In India, multidimensional challenges – demography, geography, income, and others stand in the way of enhancing the efficacy of education, and optimizing the cost of education delivery.

The most concerning factor for universities or educational institutions is the student dropout rate. At elementary and secondary level, historical data analytics can be used to predict the probability of dropout, enabling enhanced focus on students most likely to quit. In the absence of historical data, administering surveys and analyzing survey data to understand student sentiments and challenges can help identify key areas of focus.

Analytics can also predict what kind of path (including career) people with different education, psychographics, and capability are likely to aspire for. This can be achieved by accumulating past behavioral and demographic data, and leveraging analytics to identify homogeneous segments to begin with, and progressively moving to personalization of education. This way, communities can leverage each student's potential to benefit of the country.

The third key area where analytics can help is identifying and targeting student segments that need subsidized education. Traditionally, income has been the sole factor in determining subsidies. Leveraging analytics can help create a composite score for all the students, taking into account various relevant factors that determine the right funding.

Big Data is enabling big outcomes today – improving reach and performance of public education, agriculture, water supply management, etc. Implementing rigorous governance to ensure reliability and quality of data and analytics is therefore the need of the hour as with the great power of data, comes the greater responsibility of using it right.

2. Implement targeted agricultural initiatives

While the usage of analytics in agriculture is not new, advanced analytics paradigms can lend a whole new approach to impact the last mile and drive social engineering. This can be especially impactful in areas like weather and rainfall prediction, optimizing crop sowing patterns, and enabling the right genetic modifications for the seeds. Leveraging macroeconomic data merged with information through initiatives like Aadhaar and commercial repositories such as retail, banking, and electricity consumption can help governments deploy targeted initiatives such as raising the rural per capita income.

While agriculture forms the backbone of the rural economy, secondary employment through production activities and services can significantly improve the per capita income. Understanding the local produce and skillsets, and mapping them to local or global demand can help create a customized program for rural development. Aggregating such initiatives can create physical and virtual hubs for commercial activity, enabling economies of scale and competitive costs that lead to the creation of a new identity for the product or service.

3. Enable analytics-driven dynamic monitoring and management of water supply

The ever increasing gap between water demand and supply, thanks to urbanization, is further aggravated by the pilferage of water from existing pipelines. Unchecked drawing of water through bore-wells leading to alarming reduction in water tables and lead poisoning further add to the water woes.

Analytics can help aggregate relevant data from wired or wireless sensors on the pipelines - from source to destination. Further, manual opening and closing of valves can be fully automated and managed from the designated central location through real time visualization of the demand gap. Similarly, any pilferage or leakage in the pipelines can be detected and checked immediately. Satellite technology can further identify unauthorized bore-wells, assess ground water table levels, and pinpoint large scale diversions.

Striking the right balance between driving insights and ensuring rigorous governance is key

As data increasingly helps define social problems and develop solutions for them, various for profit and nonprofit companies are leveraging data analytics for the larger good.

However, as data sources explode, reliable data analytics is becoming a critical challenge. Besides risking privacy and personal rights, there is also the threat of using data analytics for manipulating or subverting public opinion. With great power to do social good, comes the greater responsibility of balancing it with infallible standards and commitment to ensuring propriety, and transparency.

ABOUT THE AUTHOR

Venkatesan Sukumaran heads the Business Analytics business at TCS, and leads a large analytics unit across all industry domains. Leveraging a world class team that brings together an impeccable knack to apply top notch analytical solutions to business opportunities, he continuous to scale and drive the business to new heights. Fondly known as 'Venky', he has closely been a part of the evolution of Analytics and Insights (A&I) business within TCS over the years.

Now, as Head of Business Analytics at TCS, Venky is keenly focused on fostering strategic initiatives that enhance client experience, accelerate client success and drive growth for A&I at TCS. Furthermore, he is responsible for driving operational efficiency capitalizing on market opportunities, maximizing resources, managing profitable partnership relationships and delivering increased revenues for A&I globally.

Betting on Data Analytics and Augmented Intelligence to Address India's Pressing Challenges

Mohanakrishnan P.

Head, Centre of Excellence – Data Science & Artificial Intelligence, NASSCOM

The disruptive potential of analytics and Augmented Intelligence is touching every aspect of human life and transforming how we live, work, and play. Data-driven decision making combined with Augmented Intelligence capabilities is also impacting every sector of the economy. Businesses increasingly realize how analytics and Augmented Intelligence can help optimize costs, develop effective marketing techniques, generate accurate customer profiles, and return higher profits. Most importantly, social entrepreneurs, enterprises, and the government are looking to exploit analytics and Augmented Intelligence to alleviate some of India's most pressing and complex problems in healthcare, agriculture, urbanization, and hunger management:

1. Enhancing healthcare delivery

According to a Lancet report, India ranked 154 among 195 countries in the 2017 Healthcare Index. The country is challenged by a grave shortage of specialist doctors, radiologists, technicians, and trained para-medical staff in tier two and three cities, towns, and villages. There are more than 14 lakh doctors and 28 lakh nurses in the market, of whom nearly 75% are in urban areas - compared to a significant portion of India's population which is in rural areas.

Augmented Intelligence can play a big role here by leveraging large quantities of data and information to better diagnose, predict, and personalize medicine. Tuberculosis, diabetes, chronic kidney diseases, cancer, and rheumatic heart diseases are the biggest 'killers' in India. *Figure 1* illustrates the Geography of Cancer across the country. Out of the 2,000 new women diagnosed with cancer every day, about 1,200 are detected in late stages. Augmented Intelligence can help researchers use large data sets to analyze how certain mutations and cancer proteins interact with different treatments and identify trends that can lead to better patient outcomes.

Data from the World Health Organization (WHO) shows that as many as 50 million of India's population suffer from diabetes, and patients with diabetic retinopathy are estimated to account for 34.6%. Traditional approaches require qualified optometrists and ophthalmologists to detect diabetic retinopathy. Today, machine learning algorithms are capable of analyzing large amounts of image data in a fraction of the time taken earlier to detect diabetic retinopathy.

Augmented Intelligence can further augment capabilities of specialist doctor services, thereby enabling remote access and increasing the reach and efficiency of healthcare. Early detection of medical conditions and preventive care can significantly reduce mortality caused by critical diseases.

2. Addressing agriculture productivity challenges

The agricultural industry is a prime candidate for massive disruption fueled by data. Benchmarking analysis across vast data sets like seed variety, soil type and other attributes such as weather patterns, soil composition, disease trends,

From improved healthcare, hunger management and agricultural productivity to development of smart cities – the potent combination of data analytics and the intelligence derived from it can turn India's biggest challenges into opportunities.
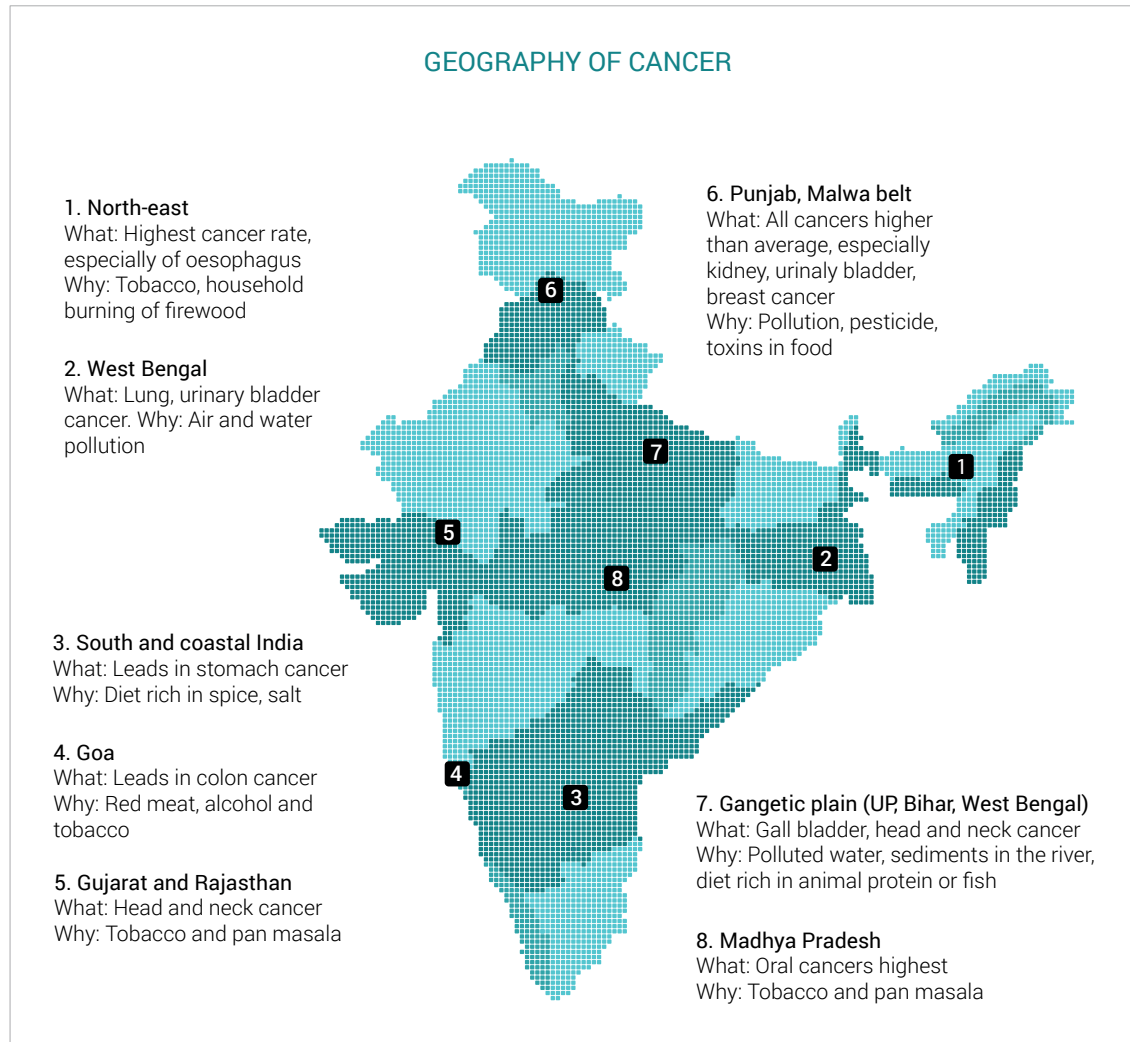
GEOGRAPHY OF CANCER

1. North-east
What: Highest cancer rate, especially of oesophagus
Why: Tobacco, household burning of firewood

2. West Bengal
What: Lung, urinary bladder cancer. Why: Air and water pollution

3. South and coastal India
What: Leads in stomach cancer
Why: Diet rich in spice, salt

4. Goa
What: Leads in colon cancer
Why: Red meat, alcohol and tobacco

5. Gujarat and Rajasthan
What: Head and neck cancer
Why: Tobacco and pan masala

6. Punjab, Malwa belt
What: All cancers higher than average, especially kidney, urinaly bladder, breast cancer
Why: Pollution, pesticide, toxins in food

7. Gangetic plain (UP, Bihar, West Bengal)
What: Gall bladder, head and neck cancer
Why: Polluted water, sediments in the river, diet rich in animal protein or fish

8. Madhya Pradesh
What: Oral cancers highest
Why: Tobacco and pan masala

Figure 2

local pests and weeds, enables farmers to learn from the experience of other farmers in the region. This kind of study can determine optimum seed variety and provide insights into all aspects of farming, realizing efficiencies in large-scale crop production.

Analytics and IoT can also assist farmers in speeding up plant growth and increasing yield. It can help researchers engineer plants that can grow and thrive in harsh climates.

It is expected that farmers worldwide will turn to AI and related technologies to optimize productivity *(challenges illustrated in Figure 2)* in all areas, including crop and egg production, milk yield, poultry farming, and much more. Analytical insights also help farmers identify the optimal time to start harvesting their crops to maximize their yield. Farmers are already leveraging drones with advanced sensors and embedded deep learning technologies to survey their crops, and obtain notifications on areas that need improvement. In dairy farming, on the other hand, sensors help detect the health of animals to analyze overall animal fitness. For instance, AI makes it possible to predict when a cow is nearing heat to provide relevant veterinary care.

3. Delivering smarter cities

Typically, within the city planning function, data and analytics are not often considered. This disregards the potential of using data insights in areas such as construction zoning, public amenities design and smart policing of transportation, to improve citizen outcomes. To become 'truly smart', cities must effectively leverage Big Data, analytics, and AI.

- Planning: Often, when new areas are created or become popular, the infrastructure already in place is not scalable to sustain continued growth, hindering further improvements in the area. Sudden influx of businesses or residents can strain even basic amenities like water and electricity. Modeling and predictive analytics helps city planners visualize where growth is likely to happen and envisage the size of the development. Amenities can then be appropriately upgraded to accommodate the anticipated growth.

- Transportation: With more people flocking to major cities than ever before, crowding and transportation are expected to further deteriorate. Today, unstructured data from advancing technologies such as smart cards, GPS vehicle locations, mobility tracking apps, images and videos, as well as data from social media, is available to transportation planners for better decision making.

- Analytics can identify patterns to help planners focus their efforts on high priority needs of commuters. With Big Data collected in real time, a city can visualize an accurate picture of the number of people riding its buses and metros, the routes that are on time, drivers' response to changing conditions and so on. The result: ability to optimize operations and quickly add vehicles where needed.

- Public security: Threat to public security, either from criminals, terrorists or riot-like situations and civil unrest, can be addressed using analytics. An analyst can deploy pattern recognition techniques and connect the dots by combing through huge data sets. Financial crime solutions focus on capabilities such as transaction monitoring and Know-Your-Customer (KYC) functionality, rules engines, and anomaly detection.

Similarly, by amalgamating different neural network models to process images, speech and text to extract meaningful information, police can identify criminals and terrorists in real-time.

INDIA'S AGRICULTURE SECTOR: PRODUCTIVITY CHALLENGES

At **169.6 million hectares,** India's cultivated land mass is the largest in the world.[23]

The Government of India's top research institute reports that nearly **60% of agricultural land is at risk** because of fertilizer misuse, poor cropping practices and soil nutrient deficiencies.[24]

India uses 13% of the world's exrtacted water. and **87% is used for irrigation. Expanding irrigation** has been a key strategy for increasing productivity; the proportion of arable land under irrigation increased from 20% to 35% from 1981 to 2013.[25]

The country is faced with the prospect of **declining rainfall during the monsoon,** India's prime growing season for rainfed agriculture.

Irrigation water use efficiency is very low 35-40% efficiency in surface irrigation such as flooding or canals, and 65-75% efficiency when pumping groundwater. These **unsustainable practices** are depleting the country's aquifers.[27]

55% of the **population** is engaged in agricultural production. As farms are divided among family members, average **farm size today** (1.16 hectares ? 2.87 acres) **is half what it was 40 years ago.**[26]

Unemployment among agricultural workers **rose from 9.5%** in 1993-1994 **to 15.3%** in 2004-2005.[28]

Government **subsidies to farmers** for fertilizer, electricity and irrigation **increased more than eightfold** between 1990-1991 and 2006-2007. Areas receiving the highest subsidies regularly underperform those with lower subsidies.[29]

Government **subsidies for** buying and distributing **food grains** to low-income and disadvantaged households **grew from 2.2%** of agricultural GDP during the 1990's **to 5%** in the 2000s,[30] crowding out investments in agricultural education, rasearch, technology and extension.

India's Ministry of Agriculture reports that from 2005 to 2007, **30%** of harvest and post-harvest **economic losses came from the fruit and vegetable sectors,** although that sector comprised only 13.6% of total production.[31]

Global Harvest
INITIATIVE   **2014 GAP Report**

---

4. Better hunger management:

According to a UN report, India is home to 190.7 million undernourished people and 38.4% of children under five in India are stunted. What's more, approximately 50% of agricultural product and food produced is wasted even before it reaches the needy. The 'Million Meals' project by Akshaya Patra, an NGO, is a great example of how disruptive technologies such as AI, IoT, and blockchain can address significant challenges in mass meal production, efficient delivery, and measuring quantity and quality of food (6).

What data analytics and AI mean for the society

Analytics and AI is a potent combination – one that is powering a new approach to solving societal challenges by putting the common man on the same platform as businesses. Successfully leveraging the power of these combined technologies requires an innovative approach that democratizes data access and creates new models for public policy development and deployment.

ABOUT THE AUTHOR

Mohanakrishnan P. heads the Centre of Excellence for Data Science (DS) & Artificial Intelligence (AI) at NASSCOM. He is responsible for augmenting the DS & AI ecosystem in the country with a mandate to develop India as an innovative country and position it amongst the top three technology destinations globally. He has over 20 years of expertise in developing strategies for outsourcing and offshoring for large IT-BPM organizations. Mohanakrishnan is a Cost and Management Accountant and an alumnus of Indian Institute of Management (IIM), Ahmedabad. He also holds a Bachelor's degree in Commerce from St. Joseph's College of Commerce, Bangalore.

References:

- Healthcare in India, Wikipedia article, Accessed February 2019, *https://en.wikipedia.org/wiki/Healthcare_in_India*

- 'Call for Action: Expanding cancer care in India', Research report, Ernst & Young Global Limited, Published July 2015, Accessed February 2019, *http://www.ey.com/Publication/vwLUAssets/EY-Call-for-action-expanding-cancer-care-in-india/$FILE/EY-Call-for-action-expanding-cancer-care-in-india.pdf*

- Geography of Cancer, Insights Active Learning, Accessed February 2019, *https://www.insightsonindia.com/wp-content/uploads/2015/05/Geography-of-Cancer-India.jpg*

- 'India is the diabetes capital of the world!', Article, Times of India, Published January 2016, Accessed February 2019, *https://timesofindia.indiatimes.com/life-style/health-fitness/health-news/India-is-the-diabetes-capital-of-the-world/articleshow/50753461.cms*

- 'India home to 23.4% of world's hungry, 51% women are anemic: UN report', Article, LiveMint, Published September 2017, Accessed February 2019, *https://www.livemint.com/Politics/8BBA9K4GHvpSvXR0ps602O/India-home-to-234-of-worlds-hungry-51-women-are-anemic.html*

Utilities and Social Transformation: Exploring the Art of Possible with Data and Analytics

Peter Jackson

Director - Group Data Sciences, Legal and General Group plc

My university days have long passed. A reflection on my time there and my career since graduation, however, reveals a close link to data. The gradual shift in my focus from technology to data is the result of the emerging recognition that the insights derived from data analytics is the key to tapping into huge opportunities.

After leaving university, I became a Business Analyst, which was the perfect foundation for a Chief Data Officer - a role that needs to understand how a business operates along with the business strategy and goals. In the mid 1990's, I recognized that the World Wide Web was going to have a significant impact on business transformation and became a back-end developer, building teams and a business to deliver integrated CMS and CRMs as a service. It was not long before I could see that the true business value and transformation did not rest in the technology but in how the technology was used. Subsequently, I transitioned into consultancy for digital strategies. This is when I had another epiphany - the value for business was not in digital platforms or even in how they are used, but in the data within the digital platforms. Long story short, my career evolved into developing and delivering data strategies.

I spent several years in consultancy delivering data strategies for fast moving commerce goods and financial services, focused on enhancing customer acquisition and retention, delivering customer experience and delight, and analyzing transactional data. It was not until 2016 that I was offered the opportunity to join The Pensions Regulator in the UK as their first Head of Data. It is here that I saw the huge opportunities and potential for data and analytics to transform an organization, and deliver social benefit by regulating pensions for the UK government.

In 2017, the idea of tapping into opportunities for business and social transformation using the data collected by a water utility, brought me to Southern Water - as their first Chief Data Officer. Water is a vital resource for society and public health, and safe, clean and reliable water supply is a basic requirement to stimulate economic and social development. Water utilities collect huge amounts of data: customer, meter, asset, operational, and safe water and waste water quality data. Collecting the right data, at the right time, storing the data safely as well as governing, managing and analyzing the data in the right way, can deliver not only operational benefits but also secure a resilient water supply.

Here are two examples of real and immediate social transformations that can be supported by data and analytics in the water sector. One - ensuring resilience of water supply into the future. Two - understanding water supply and demand. Data analytics and predictive modeling provide deep insights to understand the pressure on demand and supply - not only tomorrow but also over the medium and long term.

The pressures of population growth, economic development, rising standards of living, agricultural evolution, land use, and changing weather and climate patterns can be modeled and understood better through data and analytics.

Leveraging data and analytics in the water sector promises two-fold benefits - ensuring resilient water supply in the future and understanding the current and future pressures on demand and supply vis-à-vis urbanization, population growth, economic development, climate evolution, etc. Utilities organizations often have humongous data, mining which can help ensure optimal utilization of natural resources, enabling innovation and social transformation.

The insights from the analytics are critical to driving social transformation by enabling safe and resilient water supply. Predictive and prescriptive analytics also offer opportunities to improve the quality and standard of asset management and reduce the cost to maintain and support a safe water and waste water network.

The role of the Chief Data Officer in unlocking the potential in the data is paramount. Most organizations have a Chief Financial Officer to manage and realize the value in an organization's financial assets. Similarly, organizations need a data professional, a Chief Data Officer, to manage and exploit the value in an organization's data assets. When data is managed and utlilized strategically, it can drive business change. Many sectors have been disrupted by 'data-driven' businesses where data fuels business change and innovation, and delivers social transformation.

Since November 2017, Southern Water has been leading the The DataWell initiative across the water sector in the UK to embrace the release of open data sets for the benefit of the water sector and society. We believe that open data sets and data sharing lead to increased transparency and trust, enabling innovation and social transformation.

ABOUT THE AUTHOR

Peter Jackson is Director, Group Data Sciences at Legal and General. Previously Peter was Chief Data Officer at Southern Water and prior to that Head of Data at The Pensions Regulator (TPR), which regulates the pensions and automatic enrolment in the UK. Before joining TPR Peter spent 17 years providing data strategy consultancy across the not for profit sector, financial services and FMCG, working with large multi-national organisations and blue chip brands. Peter is a specialist in Data Strategy, Data Technologies, Master Data Management Strategies, Data Governance Frameworks, GDPR and Data Science Strategies. Peter is the co-author of 'The Chief Data Officer's Playbook' published by Facet November 2017. Peter's second book 'Data-Driven Business Tarnsformation' is due to be published in March 2019.

How the Internet of Things is Ushering in Social Good

Prateep Misra
Head of Technology and Engineering and Chief Architect, Tata Consultancy Services

IoT, one of the most exciting technologies of our time, is transforming human lives, businesses and societies at large. It has unleashed a deluge of innovative solutions that are being developed and deployed globally. Some of the global trends in the IoT space include:

Emergence of global IoT hackathons - These include both 'public' hackathons arranged by universities, nonprofits and governments, as well as those conducted by technology companies looking to attract top technical talent to develop their IoT products or harvest ideas from existing employees.

Strong venture capital funding for IoT startups – Venture capital funding for IoT startups is growing exponentially across geographies. IoT startups in the US raised around USD 1.5 billion in funding in 2017. The growth of IoT startup funding in other geographies such as India, Europe, and APAC is also very strong. Corporate venture funds are also active in IoT venture funding – especially in Industrial IoT.

Spurt in patenting activity - IoT and related technology based solutions, including embedded systems, analytics, and cloud are witnessing a massive growth in patent publishing. From 2007 to 2017 there were about 129,000 patent publications with over 35,000 patent grants. In 2016 alone, there were more than 22,000 patent applications. A new patent category for IoT has been created in many jurisdictions, with China being the most active jurisdiction. Companies such as Samsung, LG, and Qualcomm are some of the most prolific patent publishers and patent holders.

IoT: Revolutionizing society, not just industry

IoT has become mainstream in both industrial and consumer IoT segments. Several major companies, especially those in manufacturing, consumer goods, energy and utilities, aerospace, automotive, healthcare and life sciences, have invested in industrial IoT programs. According to McKinsey, by 2025, the economic impact of IoT on various sectors is estimated to be USD 11 trillion. This means even a 1% improvement in operational efficiency due to IoT would lead to significant savings. Realizing its potential, global corporations, standards organizations, and industry bodies have come together to form the Industrial Internet Consortium to enable standardization and adoption of IoT in businesses.

While IoT usage in the industry impacts people's lives, albeit indirectly, IoT has transformed the lives of common people directly as well. Here are four successful IoT projects that have positively impacted the society.

- Swasthya Slate is a low cost medical diagnostic device that performs a battery of 33 tests in less than 45 minutes. The device is being used in more than 1000 "street clinics" in Delhi to serve the needs of low income citizens. It performs common blood tests, heart rate, urine, blood pressure, and glucose checks, apart from tests to diagnose diseases such as malaria, dengue, hepatitis, and more. Clinics offer this free service with Swasthya Slate providing instant and automated diagnostic tests.

Moving beyond industrial impact, the Internet of Things (IoT) is now transforming societies and quality of life – in a cost-effective and scalable manner. From enabling better healthcare delivery to agricultural outcomes, environmental protection and care for senior citizens – IoT is emerging as the new frontier of social good.

- Nano Ganesh is a mobile phone-based remote controller for water pumps used in agriculture. Manufactured by an Indian company, Ossian Agro Automation. The simple and inexpensive device provides a low cost water pump monitoring and control system. Thousands of farmers across India are able to save labor, time, and fuel with this device.

- Safecast is an open network that enables ordinary citizens to volunteer, collect environmental data, and make it available for public dissemination as well as scientific analysis. The Safecast hardware design is open source and allows anyone to create a Safecast device. The data is also made available for free. It was used to effectively collect radiation data in Japan following the Fukushima nuclear disaster. Japanese citizens were provided a low cost Geiger counter and a Safecast device to collect radiation data at a very granular level and post it to a central cloud database. The device empowered ordinary citizens to contribute towards collecting and sharing high quality radiation and air quality data on an equal footing with the scientists. More than 2000 volunteers collect over 2 million measurements per day using Safecast.

- Shine Seniors is an assisted living project in Singapore that aims to help senior citizens confidently spend their senior years in a single location. It provides real-time alerts to community care givers in case of emergencies. IoT devices and sensors installed in the homes of senior citizens collect and transmit data that are used to generate alerts in case any abnormalities are detected in data patterns. The data is collected in a non-intrusive manner that respects the privacy needs of individuals, creating a vital safety net for senior citizens who live by themselves.

Impact of IoT: Changing lives and the world

Clearly, IoT is making a significant impact on both industry as well as the society. On-going investments in IoT devices and related technologies are driving data collection at an unprecedented scale, enabling real-time visibility into various aspects of human lives, the environment, city spaces, infrastructure, factories, and businesses processes. The best and brightest technical minds are engaged in developing IoT technology and applications for monitoring and reducing air pollution, improving water conservation, and feeding the hungry, thereby transforming lives and benefiting citizens around the world.

ABOUT THE AUTHOR

Prateep Misra is Head of Technology and Engineering and Chief Architect of the TCS IoT Platform - Connected Universe Platform (TCUP). He has over 25 years of experience in software engineering, research and technology consulting, particularly in the areas of embedded systems, RFID, IT infrastructure, real-time analytics, and cloud computing.

References:

- Venture Funding In IoT Startups Continues To Go Up, Published March 2, 2018, Accessed Feb, 2019, *https://news.crunchbase.com/news/venture-funding-iot-startups-continues-go/*

- Relecura IP Intelligence Report, Published May 2017, Accessed Feb 2019, *https://relecura.com/reports/IoT_IP_Landscape_Commercialization_May2017.pdf*

- How The Swasthya Slate Is Revolutionizing Healthcare, And Why It Steers Clear Of The US, Published Nov 2014, Accessed Feb 2019, *https://www.forbes.com/sites/stephenwunker/2014/11/22/how-the-swasthya-slate-is-revolutionizing-healthcare-and-why-it-steers-clear-of-the-united-states/#4c7e7eaf477e*

- Nano Ganesh, Wikipedia, Accessed Feb 2019, *https://en.wikipedia.org/wiki/Nano_Ganesh*

- Helping Seniors Shine On, TCS Website, Accessed Feb 2019, *https://www.tcs.com/helping-seniors-shine-on*

- Open Environmental Data For Everyone, Accessed Feb 2019, *https://blog.safecast.org/*

How the Notion of 'Distance' in Analytics Can Help Solve Complex Business Problems

Dr. Bimal Kumar Roy
ASU, Indian Statistical Institute, Kolkata, India

Dr. Sourav Sen Gupta
SCSE, Nanyang Technological University, Singapore

Employing an innovative approach to the concept of distance in analytics can enable profound business implications by opening up new ways of exploring and interpreting raw data.

Data in the context of data science and business analytics assumes an underlying structure. In certain cases, the structure of data is implied in the context of the problem, while in others, it is crucial to explicitly define the structure of the data at the outset. Often, in data science, the form and shape of the data plays a bigger role than the content.

Data science deals with four problems in particular – predicting continuous variables, predicting categorical variables, discovering patterns in data, and highlighting anomalies in data. Each of these maps to a problem in computational optimization, and knowing the context and structure of the data not only helps in identifying patterns but also in optimizing outcomes.

Importance of analyzing and comprehending the topology of data

In business analytics, the most common form of data belongs to a hybrid vector space, with numerical as well as categorical dimensions. This enables users to embed the well-known topology of a multi-dimensional vector space over numeric dimensions and regular cubic partitions on the categorical ones. Every benefit of a real vector space, especially that of continuous gradients, is inherently assumed for problems such as regression. Similarly, problems in classification implicitly assume the benefit of partitions in the data space to employ the tools and techniques of piece-wise regression and range search with tree-based models.

The commonly utilized models in business analytics are simply adaptations of standard optimization techniques customized to the space and structure of the data being considered. The advent of new contexts in data science has ushered in new topologies in data, some of which are transforming the way we think of analytics.

While the applications of data science in the context of society have been around for over a century, the notion of society in the context of data science is relatively new. It began with the inception of the World Wide Web, which created a space for peers to generate, curate, and utilize data for day-to-day activities. The boom in social networks and online communities further propelled the notion of data into the context of peer-to-peer societies. This naturally transformed the data landscape into network-based topologies, with graphs and complex networks dominating the arena of data science, sentiment analysis, recommendation systems, and online advertisements. Let's take a look at graph-based topology of data to understand the significance of various notions of distance.

Optimizing complex topologies

In the classical space of data, finding patterns and clusters depended on the statistical notion of 'within and between' variances, which often relied heavily on the predominant notion of distance in the space. Even in a standard real vector space, choosing cosine similarity as the notion of distance, instead of the regular Euclidean distance, would

completely change the form and shape of clusters. Similarly, choosing standardized statistical distance instead of the regular Euclidean distance provides a new notion of scale and skew to a standard vector space. The effect of considering correlation-respecting notions of distance in a vector space was postulated by the pioneer of Statistics in India – Prasanta Chandra Mahalanobis.

While Euclidean distance preserves the standard notion of orthogonal axes and cubic topology in a real space, cosine similarity provides a fresh angular topology on the same while standardized statistical distance introduces a notion of scale and skew to the topology. In contrast, Mahalanobis distance essentially generalizes the notion of standardized statistical distance to a space with non-orthogonal set of axes, and provides crucial insights into clustering data represented by a correlated set of features.

However, the advent of complex networks in data science - resulting from the World Wide Web and social networks have given rise to a broad spectrum of graph-based notions of distance. Clustering on a network often targets community detection in a graph, allowing one to choose two critical distance measures. One while defining the edges in the graph and the other while connecting two smaller communities to form a larger one. The first notion of creating edges in a graph can follow any topology imposed on the data, creating a uniquely diverse possibility for the data space.

Consider the network of roads on the city map of Kolkata, for instance, where one may define a graph by considering the regular paths as edges and intersections as vertices. The weights on the edges however, is open for interpretation – one may choose either the distance, the time to travel, the traffic density, or any combination of these parameters to define the weight on the edges. This creates a multitude of graph topologies based on the same network. Each such definition will provide a new notion of shortest distance between two nodes in the graph and will lend itself to a fresh interpretation of algorithmic optimization.

In a slightly different setting, the distance between two profiles on a social network may be defined by the minimum number of hops between them (as in LinkedIn), the number of common associations shared between two profiles (as in Facebook), or the number of shared messages (as in Twitter). Once again, each such notion of distance provides a unique abstraction of the network into a graph, which in turn leads to solving a case-specific problem in analytics. The notion of clustering seamlessly migrates from statistical compactness in a real vector space to tightly wound cliques on a graph in each such case, where communities are formed based on proximity of the abstraction, instead of the actual proximity. The abstraction of a data space into a complex network, with varying notions of distance, is probably one of the most influential transitions in the domain of data science and analytics today.

Diverse notions of distance in abstract topology: Business implications

Significant practical impact of the diverse notions of distance in the abstract topology of a graph has been observed in contexts never imagined before such as:

- Suggesting connections on social networks (Facebook, LinkedIn).
- Recommending products used by similar identities (Netflix, Amazon).
- Finding political communities in online opinion platforms (Twitter, Blogs) and mining their sentiments for predicting outcomes of elections.
- Finding the shortest path on a complex public transport map (Google, Baidu).
- Optimizing allocation of resources in catering to customers spread across a neighborhood (Uber, Ola).
- Identifying links between nodes in a network (Instagram, WhatsApp).
- Identifying targeted profiles on the cryptocurrency network (Bitcoin, Monero).

It is of critical importance in modern data science to elevate practical business problems from their natural space of features to a rich topological abstraction that supports complex computing tools and techniques. Nothing is more impactful in this context than an innovative approach to the notion of distance, which can provide a fresh new topology for the data scientist to embed and explore raw data.

ABOUT THE AUTHORS

Dr. Bimal Kumar Roy is the Head of R.C. Bose Centre for Cryptology and Security at the Indian Statistical Institute (ISI), Kolkata. He was previously the Director of ISI from 2010-2015 and has served as a visiting professor at many universities across the world. Dr. Roy's research areas include Cryptology, Information Security, and Data Science and he has received many prestigious awards, including the Padma Shri from Govt. of India, the IBM Faculty Award, the INSA Teachers' Award, and the Reliance-NASI Platinum Jubilee Award. He holds bachelor's and master's degrees in statistics from the ISI and a Ph.D. from University of Waterloo, Canada.

Dr. Sourav Sengupta is a lecturer at Nanyang Technology University, Singapore. His current research interests focus on evolving areas in Computer Science, including bitcoins and blockchain, machine learning, and analytics on encrypted data. He has delivered lectures at prestigious universities, conferences, and workshops around the globe. Dr.Sengupta holds a Bachelor's degree in Electronics and Telecommunication from Jadavpur University, Kolkata, Master's degree in Mathematics from the University of Waterloo, Canada, and a Ph.D. in Cryptology from Indian Statistical Institute, Kolkata.

In Conversation with Anita Kane

Anita Kane
City Data Officer, Pune Municipal Corporation (PMC)

Pune Municipal Corporation's open data initiative aims to tap the potential of municipal datasets to drive data-driven decision making, and seed the data culture within and beyond the Corporation.

Could you describe your current role at PMC and how you began your journey with the organization?

Sure. As the city data officer (CDO) for the Pune Municipal Corporation, I am responsible for three things:

• Driving and leading the organization's open data initiative

• Contributing through capacity building of the departmental data officers

• Improving the data culture within PMC

Incidentally, I am the first CDO for any Urban Local Body (ULB) in India. The role had no set frameworks and guidelines to start with. Because it was a first-of-its-kind role, it required tremendous drive and creativity along with patience to work in systems with significant bureaucratic and political undertones. But it was also a unique leadership role within the city administration. I work with the senior city leadership to tap the potential of municipal datasets to drive data-driven decision making, and seed the data culture within and beyond the Corporation.

The story behind my appointment began in 2016. Tata Trusts tied up with the World Council for City Data (WCCD), an international organization based in Toronto, and enabled the ISO 37120 certification. Three cities came forward and Pune was one of them. In early 2017, Pune obtained the Platinum Level certification for ISO 37120 - a certification for smart cities. Pune did a great job of reporting more than 90 indicators (out of the 100 indicators that are recognized by ISO) in order to obtain this certification. In terms of taking this initiative forward, Kunal Kumar, the then Commissioner of PMC, envisioned the need for a City Data Officer (CDO). He expressed his desire to Mr. Ratan Tata that Tata Trusts fund the position of CDO to take PMC's data program forward. In due course, I was selected and brought on board through a structured interview process.

Has Tata Trusts partnered only with the city of Pune for sponsoring the role of CDO or is this across multiple cities?

Currently, Pune is the only city where Tata Trusts has deputed a CDO. Tata Trusts views this as an opportunity to deepen its relationship with cities. Surat and Jamshedpur were the other cities that Tata Trusts established a relationship with in 2017, and is currently working on moving the relationship to the next level.

It is worthwhile to note that within a year after the deputation of the first CDO in Pune, the Smart Cities Mission included the appointment of a CDO in the Human Resource guidelines for Smart Cities. In 2018, Tata Trusts established a relationship with five more cities. Considering all this, it is very likely that they will depute CDOs in these five cities.

What are you and your organization doing to make a positive impact on the society or the environment in general? For instance, what is Tata Trusts doing, what are you as the CDO doing, and what is the data ecosystem at PMC doing?

Tata Trusts has been engaging with the Ministry of Housing and Urban Affairs as well as Niti Aayog in sharing the best practices around the CDO's role. I have been very closely associated with them through this journey. The inclusion of a CDO in the HR guidelines, that I talked about earlier, is a contribution to the government. Tata trusts is also working very closely with the Ministry of Housing and Urban Affairs in creating training modules for capacity building of CDOs as well as departmental data officers. These are 15-20 hour training programs in the form of videos.

Coming to the question on my contribution, let me give you two examples. The first thing that I did after I joined PMC was to enrich the open data portal. This included collecting data and then publishing it on the open data portal. My focus was on improving the data quality, as well as the depth and breadth of the datasets. Secondly, I engaged with the academia, talked to them about what was available, and invited them to co-create opportunities to serve our city. Through this engagement, the students of the Department of Statistics at Pune University came forward with five domains/datasets for analysis. The topics were around water body pollution, noise and air pollution, transportation, and health. The key findings were discussed with the senior leadership of PMC. The papers were also published in conferences such as IEEE Punecon.

Though I was encouraged by the response from the Statistics Department, I was still thinking about how the adoption of open data could become a "Tsunami" of sorts. Through my interactions, I had found that the understanding of the concept of open data was very poor amongst the general community. The community required hand-holding to relate the data with possible solution areas.

Given the focus on data science in today's curricula, I structured a series of hackathons for the top educational institutions in the city. These hackathons have elements of team work as well as competition and focus on data extraction and cleaning of data sets in domain areas such as education, transportation, property tax. Problem statements are defined, related datasets are identified, and there is some hand holding to get the students. Guidance is provided by a small team of

experts. These hackathons resulted in additional "clean" datasets for the PMC. Sponsored by IEEE Pune several hackathon sessions have been held at Symbiosis, COEP, SPPU Department of Statistics. A few more have been lined up.

A little bit about a sample outcome of these hackathons: education data sets now have class-wise enrollment data for more than 1200 schools in Pune. In addition, it includes information on school facilities, teacher availability and so on. This data could be used to develop an interactive visualization that helps parents select the right school for their wards. This is the social impact aspect of it. I am looking forward to more outcomes such as this one in the future.

Are there any specific priority areas that you are focusing or are you looking at data across all sectors?

I am looking at data across all sectors. However, I am focusing on priority areas that are aligned with citizens' top priorities. For instance, water supply, sewage, solid waste management, education and primary health services. Roads, traffic and public transport is another important aspect that I am looking at. Availability of reliable data is important too. Not all of the available data is clean and reliable. There is a manual collection interface that can lead to errors. Education and healthcare sectors have fairly good data. Currently, I am not doing anything special in less critical areas such as cultural heritage or labor offices, while we do publish whatever data we have in these areas.

What are some of the typical challenges that you face in your day to day work?

One challenge is that the data exists in silos, i.e. the systems do not talk to each other. For example, the database at the building permission department does not seamlessly talk to the property tax department that actually taxes the properties. So the data exists in silos and it's not possible to perform good analysis with such data. The second challenge is related to data collection. Most of the data is coming through manual collation. So we deal with two limitations here; one is the frequency at which the data can be collected, and the other is data quality and reliability. Let me explain this challenge a little bit. We have 10 sewage treatment plants in Pune and they are all located close to the river. While a mobile app was created to provide reports on the pollution levels of the inlets and outlets, the data is collected through a manual interface today. This means extensive validation of data can be a big challenge. The new STP systems that are expected to be built in 2019-20 will capture data directly with IOT enabled devices, ensuring cleaner data.

The third challenge is 'people and change management'. Today, the data culture is missing. Let's say a good quality environment status report has been developed by the environmental officer. But two things happen; he just stores the final report and the charts but not the raw data. So it is important to sensitize all stakeholders about how to effectively use and store data. The other aspect is that usually PMC has consultants who come in gather a lot of data, crunch it to and produce the final output. But nobody in-house takes the ownership of the data or tries to completely understand what the consultants have done. So the data ultimately remains with the consultants and not within PMC. The third aspect is regardless of where the data resides, there is still a lot to be done to inculcate the culture of actually using the data for decision making. The citizens can now come forward and ask questions supported by data. Let's say you live in a ward which has good tax collection percentage. You now have the capability to check on whether the quality of services or the funding that the corporation is providing is in fact proportionate to the revenues they are making in that ward.

The fact that Pune obtained a Platinum Certification, does that mean the city was at a higher level of maturity where data is concerned? Is there a general appreciation for the need for data and analytics?

Pune was not immune to the various quality issues and the data culture issue that I mentioned earlier. But Pune did have some capabilities in place. The fact that it could report more than 90 indicators for WCCD certification is not a trivial matter. Understanding the metrics, calculating the different indicators, sampling, projecting and so on is fairly complex data analysis. So Pune had a lot of things in place along with a better understanding of the data ecosystem.

As for appreciation for data and analytics, I think senior professionals have it. But it has not percolated down the hierarchy. However, Central Government's initiatives such as Livability Index and Swachh Survekshan are creating an awareness for data.

Is there anything else you wish to talk about when it comes to the relevance of data in your work?

Yes, data analytics is extremely relevant to our work. We have very good quality data on vector borne disease incidence. There is data available on chronic water logging spots too. So there's potential to build a predictive system for vector borne diseases. Smart Cities Mission is now working on building a predictive model around women's safety. So yes, I am excited about the many good things achieved as well as those that are yet to come.

Under Swachh Bharat, they run an annual survey and Pune is ranked somewhere between 10 and 15. So there must be different set of parameters used for this report... is there an action plan to change things based on the analysis of the data or are the initiatives run routinely in accordance with established process?

Good question. I was just in a meeting with the Commissioner of Pune who is taking the Swachh Survekshan survey very seriously. The discussion revolved around the weak areas we need to focus on based on data analysis. A part of the discussion, also revolved around the Livability Index, where Pune was ranked number one. The idea is to analyze last year's survey and take necessary actions to keep Pune in the top spot.

How are non-profits and social enterprises that use data analytics different from the ones that do not use it?

These days, most social enterprises that are decent sized are conscious about data. Take for example Janwani, which is responsible for program managing the road cleaning activities of 50% of 'Prabhags' in PMC. Janwani's project manager has an impressive dataset that allows him/her to identify chronic garbage spots. Janwani reports the interventions undertaken to address these chronic areas and reports the outcomes on a weekly basis. The cycle plan data and the WISE index created by the Centre for Environmental Education is another example of data being used by non-profits. Shelter Associates is another well known social enterprise that uses data. The problem is these bodies don't talk to each other and are averse to bringing their data into the open data realm. The need of the hour is a platform where the non-profits can share and exchange information, best practices, and even extra supplies.

You built a career in the corporate world and now you are working in a different space. When it comes to taking a data-driven approach, how does a public entity differ from a corporate entity?

This question reminds me of some of the data related issues that were discussed when I was in the corporate world about 12-15 years ago. The challenges revolved around siloed data, not having a single source of truth, etc. The world of public entities is grappling with these issues only now. So they are approximately 10 years behind the corporate world. For instance, the Property Tax department is currently carrying out a GIS enabled survey of all the properties in PMC. The survey is almost complete. It shows that a few hundred crores of additional tax could be potentially billed by the PMC. But the department must perform 'x' number of steps to validate the survey outcomes before it can raise an internal "demand" to take action based on the results. Unfortunately, since all the business processes of the Property Tax department are not automated, it is very difficult to track progress.

What role did technology play in this transformational journey?

Technology is the primary enabler in transformation. For example, the open data portal is technology-enabled. We are all waiting for more technology-enabled initiatives to go live - like the fully-automated sewage treatment plants that are IoT enabled - in order to get more granular and reliable data at a higher frequency. On a similar note, the Toilet Board Coalition, has partnered with PMC to enable smart sanitation. One aspect of this is about creating business models that will make public toilet management self-sustainable. There are three types of sensors that are being piloted in public toilets. There is the footfall sensor, based on which advertisement revenue can be generated. Then there is the ambient air quality sensor which could be used to generate alerts for the next cleaning cycle. The third type of sensors will actually be installed inside the toilets to detect early indications of an epidemic or the signature of viruses. This can also help monetize the data. All of this has to be done without compromising the privacy of the individuals. With increasing technology enablement, data governance and policy are becoming extremely important. This is another key focus area for me, currently

Is this kind of data maturity a prerequisite for getting a city recognized as a smart city? Is it an important parameter?

Yes. The Smart Cities Mission is now using a term called "Data Smart Cities" instead of just smart cities! There is a checklist that cities must meet before they are deemed as smart cities. While it includes very basic things like a double entry accounting system, it also asks for a certain level of maturity – such as a plan to implement 100% metering of water. Once you qualify as a smart city, two things are mandated: an integrated command and control center and a citizen portal. So data and technology-enablement is mandatory.

Can you describe your personal experience in this transformational journey.

Moving to PMC was a big change from the corporate world that I was used to. But having been a member of the Pune civil society for 30 years, I thought this was an excellent opportunity for me to contribute to the municipal corporation. After a long career in the corporate world, this seemed like the right area to get into. Getting accepted at PMC, was not simple though. The approach that I took was, "I am one of you". And this worked for me. I am enjoying the work very much, despite the many challenges. It's very exciting to be the first CDO as well as contribute to the nation's transformation journey. I am thankful to Tata Trusts, TCS and PMC for giving me this unique opportunity.

ABOUT ANITA KANE

Anita Kane is City Data Officer in Pune Municipal Corporation and a Principal Consultant with Tata Consultancy Services. She is currently leading and driving the open data initiative for the Pune Municipal Corporation. She is an IT professional with over 30 years of experience. She specializes in the Banking and Financial Services industry segment and has held responsible positions at TCS in the areas of Business Development, Solution Design and Delivery. Anita is an alumna of Indian Institute of Technology, Bombay and has a passion for teaching. She has been an active sports person in her younger days. She was the national girls Table Tennis Champion in 1976 and represented India too.