Terraforming Data Science/Big Data Strategy: *building a modeling regime that works for the future*

Syed Danish Ali

1. Overview

Let's not stand on ceremony here and dive straight into the core questions: How do we develop a strategy for data science that allows us superior insight into our data? Do we make the strategy have narrow focus or broad focus? How do we see that the strategy is still relevant into the coming future of constant changes? These are some questions we tackle here to effectively terraform the big data/data science strategy for an insurance company or insuretech. While the strategy will be definitely different for other types of business, the main fundamentals will remain the same. These few un-wavering fundamentals are:

1. Apply both quantitative and qualitative analytics. No one model or method knows it all
2. Never forget risk; almost everything can be broken down into what core risks are you taking as an insurer or insuretech? Model the risks
3. Integrate your big data strategy with other crucial strategies like product strategy, business plans and so on. Make holistic sense; one strategy should complement the other
4. Put the customer and Artificial Intelligence (AI) at the core of everything that you do. Without better customer experience, the products won't sell. Without putting AI at the heart of what we do, other players especially tech disrupters will make you redundant.
5. Learn continuously; new technologies are emerging before previous ones mature. Never be afraid to try out something new and learn fast and quickly adapt. Have a futuristic vision; communicate your vision to stakeholders and invite them along for the journey.
6. Combine domain expertise with AI to overcome limitations of each (domain or AI) on their own.

The business has to further ask; what is the business focus of the company or startup of insuretech? Is it B2B where insuretech focus on improving business processes for current insurers or B2C focus? Insuretech with B2B focus automate claims, automate regulation and compliance handling, applying AI to streamline business operations and much more. For example, Lapetus startup sells Chronos software solution to insurers where Chronos allows these insurers to provide life insurance by customer taking a selfie. There are also B2C insurers like Lemonade, Trov, Teambrella and many more than approach customers directly. So is our focus B2C or B2B? And do we create something very new, like a new technology or algorithm or do we apply existing cutting edge technologies and focus on retail?

Terraforming Data Science/Big Data Strategy

Utilizing opensource technologies as much as possible should be a core aspect of any data science strategy. We are seeing the revolution brought about by opensource. No longer is it the situation that only propriety software, sitting on protected computers in some big insurer's premises has all the data and modeling software. Now we have github for coding depository where anyone can 'fork' existing code and modify them; there are Massive Online Open Courses (MOOCs) like Coursera, Udacity, edX, datacamp etc. where all the latest technologies and programming is being taught for cheap prices online; We have hadoop, R, Python, free software for database handling, web design, business intelligence and for just about anything. Instead of making, for instance, your own GLM source code, just install the GLM library in R and start applying. The opensource trend is exponential because we don't have to sit in our isolated corners and try to reinvent the wheel. Our macro strategy has to be on leveraging technology partners and utilizing SaaS to continuously deepen our modeling expertise and utilize white-labels as much are available to increase skill transfers to. We have to aim to stand on the shoulders of giants to look beyond what others are able to see faster.

It's also necessary to be model agnostic. That means that we don't rely on 1 model only and claim it to be world's best because there are thousands upon thousands of algorithms, each with their own strengths and weaknesses. Whichever model satisfies our criteria like good fit to data, good model diagnostics, takes less time etc. is preferable to us. Moreover, new and better algorithms and being built rapidly so it makes little sense to prefer one model which can turn out to be outdated over the future.

We have to focus on actionable insights and data driven decision making. This does not mean that CEO and non-data scientist are supposed to know the programming, model building and whole of data science to be able to understand the results; The data scientist will present results that are simple to understand and readily actionable for everyone to understand.

2. Durable modeling regime

There is a whole universe of actionable insights that can be arrived at through data science analysis. Actionable insights are only limited to our imagination and what questions we ask. For instance:

1. What is the probability that a person with a certain profile will buy insurance or not?
2. What is the probability that a person with a certain profile will renew his policy?
3. We start with some hypothesis; like new marketing gimmick; did it achieve its purpose? Did it increase our impact or increase profit or increase revenue?
4. What is the lifestyle behavior of particular set of customers? We can cluster them into target groups like group A has 3 similar habits, group B has 4 similar habits and so on. And then see their losses by groups; Accenture already defines 'nomads', 'hunters' and 'quality seekers'. See here. This means researching continuously and learning from consulting's vast surveys to inform our research.
5. What are the signs of unusual activity? Is it just an anomaly or a fraud or adverse selection? Is our portfolio actually made up of good risks or too many bad risks have bought our policies?
6. Have our monthly targets being met? Are our expenses being overrun? Are our rewards too costly or not making an impact?
7. KPIs like customer loyalty, expense ratio, loss ratio etc. will require constant monitoring as well.
   And so on.

Our "Modeling Regime" and data science strategy that is suited for creating the future is broad and focuses upon the following strands and types of analytics:

1. Descriptive analysis
2. Predictive analytics
3. Unstructured data analytics
4. Big Data
5. Actuarial analytics
6. Enterprise Risk Management (ERM) modeling
7. Qualitative profiling including emerging risks

The underlying motivation for having a broad modeling regime instead of a precise compact regime is following Peter Drucker in that 'doing the right thing' is more important than 'doing things right'. Actuaries have historically over-specialized and focused on obsessive accuracy within these well-defined compartments which means that they have no time or resources available to broaden their modeling approaches. I suspect that this is particularly true of previous generations than millennial actuaries. Of course, no one actuary can do it all but together as a profession we can do the right thing without fragmenting into too many specializations that makes us lose sight of the bigger picture.

Terraforming Data Science/Big Data Strategy

There's also complexity science to know. The 21st century is the century of Complexity as per Stephan Hawking. It is the most realistic theoretical modeling framework we have yet and complexity science demolishes centuries of faulty and naïve simplistic modeling. Complexity sciences shows how very simple iterations and situations can over time lead to increasingly complex results in macro overall without the need for any middle central authority. Something very tiny becomes huge quickly and blows up the current system like financial crisis, natural catastrophes etc. (Chaos theory; fractal theory). We can include randomness as much as we want in our models to make them more stochastic, but this still underestimates the tail risks.

Essentially the data scientist will be a social psychologist applying software to know more about the customers. That means applying behavioral psychology and sociology to make customers better, healthier and more loyal.

Last but not the least, we will have to do qualitative profiling and regularly evaluate emerging risks. Emerging risks are like new products, new risks being made. Complexity science allows us to model emerging risks and extreme events better than other methods. It works very well when there is no data and when even the product and situation is open ended.

The idea behind qualitative profiling is simple; not everything that's important can be quantified. There are many dynamic qualitative features in our strategy and business decision making that can be structured through qualitative profiling.

2.1. Descriptive Analytics

Descriptive analysis helps us peer inside the nature of data, its shape to decode the meaning that the shape has. It explores what the data can tell us about customer features, preferences, likes dislikes and is mainly an observational exercise. Once we know the nature of data, then we can extrapolate those same trends into the future by assuming the past is an accurate reflection of what will come in the future. Of course this assumption is limited and so needs to be complemented with simulations and qualitative holistic understanding.

To simplify, we are giving brief snapshots into our modeling regime. The format is input-modeling-output for clearer understanding. Note that this guide is not definite; we are not saying there are only GLM or SVM because we have to be model agnostic as we have thousands of model (for instance, just one R library package 'Caret' has 147 machine learning algorithms). These snapshots just give instances of useful models that can be considered for the modeling process.

The descriptive analytics snapshot is:

Descriptive Analysis/Pattern Recognition/Exploratory Data Analysis		
Input- Structured Data Mostly	**Modeling**	**Output**
Data of customers like social-demographic data, answers to structured questions, lifestyle behavioral data, tech partners generated data that is structured like dashboard	Ensemble; Bayesian models like Naïve Bayes	statistical nature of data is identified
	Support Vector Machines	how many clients that are approached are becoming customers
	K means clustering	customer loyalty
	Principal component analysis; many more models for pattern recognition	customer churn and retention

Using research from the internet like how Accenture is categorizing customers into 'nomads', 'hunters' and 'quality seekers'.	Random forests; decision making trees class of algorithms. XG Boost	What is the underwriting portfolio's composition? Is it mostly of bad risk? Any interesting observations?
Unusual data entries	recommender systems; anomaly detection like A/B testing	Are we sustaining customer's interest or does it fizzle out after initial excitement? Anomaly detection will flag very unusual data which might mean fraud that needs to be investigated further if too many flagging are reached
	neural networks and unsupervised deep learning	How many groups are in the data? For instance, there might be 4 groups in the data; 1 group of youth, 1 of old but tech savvy, 1 of women and health conscious before getting married. These 4 groups will allow us to give targeted/personalized rewards then
Little data will be required but accurate assumptions will be required for setting up the simulations for agent based modeling. Network theory will mine social media data to see how people are connected to each other; like I am friend of his who his friend of her and so on. this will explicitly show the informal networks operating in our customers to guide our peer to peer business model	complexity science models like network theory, agent based modeling, chaos theory	How are people connected to each other? This way we can make groups based on being digital peers; emerging risks can be shown in simulation that how will they likely become in the upcoming future and impact us.
Mining qualitative data on the internet like web crawling to assess litigation chances, social media mining for seeing sentiment and opinion of your customers today, understanding notes made by underwriters and claim adjusters and other forms like performance appraisal forms		How is our brand value sentiment today? Any crisis in the making? Any angry rant from customer that should be resolved before it blows up? Internal insurance notes can give our data a context which we need

2.2. Predictive Analytics

Based on our understanding of nature of data in descriptive analytics, we can now predict what the next transactions in the future will likely have for us. That's predictive analytics; analyzing history to generalize to the future; complexity science simulations will have their own simulations and parameters and will not rely 100% on historical experience.

Predictive Analytics		
Input- structured data mostly	**Modeling**	**Output**
Data of customers like social-demographic data, answers to structured questions, lifestyle behavioral data, tech partners generated data that is structured like dashboard	regression on decision trees	What is the probability that a potential customer will become an actual customer?
	regression on SVM Support Vector Machines	What is the probability of customer churn and retention?
	Ensemble	What is the expected future loss for customer with particular age, gender and socio-demographic data?

	Generalized Linear Models GLMS; Generalized Additive Models GAMS. Generalized stochastic models	Is our premium price commensurate with the level of risk that we are undertaking? Is our premium underpriced or overpriced?
Using research from the internet like how Accenture is categorizing customers into 'nomads', 'hunters' and 'quality seekers'.	wide range of stochastic models like Munich chain ladder, complexity science simulations, Monte Carlo simulations etc.	Is our reserving adequate? Will we face future surplus or shortfall in reserves?
utilizing marketing data	Logistic Regression	Are our revenue, expense and profit projections accurate or not?
	Neural networks like deep learning; many sub-models within deep learning like reinforcement methods, Convolutional deep networks, Recurrent Neural Networks, hidden Markov models etc.	What is the probability that a customer will claims? (it's called frequency)

Little data will be required but accurate assumptions will be required for setting up the simulations for agent based modeling. Network theory will mine social media data to see how people are connected to each other; like I am friend of his who his friend of her and so on. this will explicitly show the informal networks operating in our customers to guide a peer to peer business model	Fuzzy logic	What is the average size of claim, once a customer claims? (it's called severity)
	Regularization models like LASSO and Elastic Net	Are we aligning our business activities with our risk appetite or are we over-doing it so that we will likely have a crises in the making soon?
	MARS regression Multivariate Adaptation Regression Splines	What group of customers are causing heavy losses? Which groups are causing low losses?
	C4.S	Price elasticity curve; how sensitive are customers to changes in our prices? For example, some people may leave if we increase prices by AED 1 but some may not leave even if we increase prices by AED 100
	Gradient Boosted Machines GBM	Is our price curve inelastic or elastic? We aim to make inelastic curve which means that customers get addicted and will not leave even after successive increases in prices
claims data;	computer vision	automated claims handling and settlement

2.3. Unstructured Data Analytics

It's well known that about 80% of data is unstructured which we could not analyze before the data science revolution. These were call scripts, video calls, telephone messages, chat typed notes, hand written notes, and data that is not tabular into neat rows and columns. Chatbots can be used as the focal point for unstructured data analytics to generate meaning and insights from the conversations that our customers have with the bot to personalize our services and remove any pain points. Unstructured data will also be analyzed internally to increase automation of routine tasks.

Chatbots can aid hugely in customer support, selling insurance, providing value added services as well in automating routine internal works in the insurance company to allow it to become an autonomous organization. An autonomous organization is able to scale rapidly and does not require increase in employees when faced with increase in revenues. All insuretech platforms aim to be autonomous organizations and not like behemoths of today filled with employees doing clerical work.

Unstructured Data		
Input	**Modeling**	**Output**
Unstructured data that cannot be properly given in rows and columns	text mining	What keywords are frequently being used?
like conversational data	sentiment analysis/opinion mining	What is the frustration level of the customer? Once a confidence level is triggered, the health coaching query is transferred to human coach for answering
social media data	Natural language processing	What are the main moods of our customers today on health coaching chatting, in social media and broadly over the internet? Is it negative or positive? What are their main points that they criticize and like?
web crawling data over the internet	parsing	What key sentences and phrases are often repeated together? What meaning does it indicate usually?
	NoSQL	
	R python text mining	
	deep dives	

	AWS and Azure Machine Learning	
voice talking by customer to our app	voice handling; deep learning; web crawling; computer vision	Opportunity to cross sell e-commerce, policy aggregation through voice interfaces.

2.4. Actuarial Modeling

This is the domain expertise in the modeling regime. Actuaries are historically taught insurance and its various nuances and if data is the new oil, then actuaries are the chief 'refiners' of this oil in making it sense and appropriate for the insurance context and task at hand. [1]

Another core element that has to be implemented is actuarial analytics and enterprise risk management (ERM). Actuarial analytics are undertaken when reserving and pricing for insurance. It also involves product development, simulations and a lot of quantitative modeling that is focus on getting the right fundamentals. Most insuretech startups lack actuarial analytics and ERM will allow us to make risk-based decision making and guide our risk appetite and business strategy. When many startups scale, you cannot have the same structure-less system, the fratboy culture of breaking things quickly and ignoring institution building and functions like HR etc. The scaling brings in a lot of problems with it. ERM will try to manage these two odds that we remain fresh and agile like a startup and not become a status quo corporate bureaucracy but that we also avoid problems arising from lack of structure and institution building.

[1] https://www.recruiter.com/i/actuaries-of-the-future-what-role-will-actuaries-play-in-shaping-the-ai-led-insurance-companies-of-tomorrow/

Terraforming Data Science/Big Data Strategy

We can also make business plans and projections for revenue and profit as well as actuarial pricing for our products. This is crucial because for instance, actuarial pricing specifies a minimum floor that must be charged to the customers for taking on the risks. Our actuarial pricing should show that our pricing is adequate and we are not underpricing.

Actuarial Modeling		
Input	**Modeling**	**Output**
All datasets structured mention previously	Pricing	Are our prices commensurate with the level of insurance risk that we are taking on? Is our price underpriced or overpriced? What ages, gender, socio-demographic profile is causing lower losses and higher retention and vice versa
	Reserving	Are we holding adequate reserves for unexpired portion of premiums and claims that have not yet been reported?
	Product Development	making new products and evaluating their adequacy
	Evaluating financial results like underwriting results analysis, expense analysis, projections of any kind, evaluating reinsurance results and so on.	Have our actions achieved their desired results? Like has some marketing new offer resulted in 10% higher impact? Are our projections accurate? How much losses will we occur and how profit will we generate?
	building investment models like ALM Asset Liability Management Models	Making sure that our assets are matched with our liabilities by duration and nature so that our investment risk is contained.

.

2.5. ERM Modeling

Enterprise Risk Management Modeling		
Input	**Modeling**	**Output**
all of our data both structured and unstructured	capital modeling	Capital modeling quantifies all significant risks; like we need to hold $200,000 for underwriting risk, $100,000 for market risk, for financial risk, investment risk, operational risk, strategic risk and so on.
Company data	economic scenario modeling	quantifies economic rates like interest, equity rates to utilize in investment ALM asset liability management in our startup
	Own Solvency Risk Assessment	an ORSA report deals with whole risk regime and ERM system in place
	Qualitative Risk Registers	risk registers identifies what qualitative and quantitative top 20 risks we are facing in our startup and how we plan to address them
	Emerging risks analysis	What risks will become huge opportunities or crisis in the future? For instance more technology might eliminate our current markets but lead to new markets being opened. The STEEP framework for analyzing emerging risks is good; STEEP means Social, Technological, Economic, Environmental and Political nature of emerging risks.
	scenario modeling	make any scenarios and see what revenues, expenses and net profit/losses are being caused by each scenario
	stress testing	If worst case scenario actually happens in reality, will we be still solvent? Like VC withdrawal, expenses overrun, natural catastrophes, huge claims like mall fires floods terrorism events etc.? many customers lapse; number of policies traction is not being made; inflation and economic shocks occur in our target market

The advantage of ERM modeling is its holistic perspective to get to the core and de-construct situations and results into ‘risks’. Despite future being uncertain, the fundamentals can be broken down into concrete risk categories which ensure that ERM remains relevant for the coming future.

2.6. Qualitative Profiling

Not everything that is important can be quantified; hence we need to inspire our quantitative modeling with qualitative profiling; Reality will never happen what we expect and we know that all models are wrong but useful.

Maturing our forecasting foresight[2]

Emerging liability ventures into the unknown, into open ended subjectivity. To sharpen our forecasting skills in this area we utilize the profound learning provided in the essay by Werther for SOA publication. That essay aims to help financial and insurance practitioners better recognize, assess and respond to largescale, large-impact rare events (LSLIREs), occurrences often wrongly labeled as unpredictable black swans. The learning impact from recognizing LSLIREs can be readily applied to LSLIRE emerging risks and liabilities as it is the LSLIRE emerging risks that matter most and not emerging risks that will likely have little consequence. The key techniques of pattern recognition for pattern change to identify LSLIREs are:

- **Use Multiple Methods Arrayed Around the Assessment Target:** Utilizing a diverse array of models and methods to assess emerging LSLIRE can result in reasonable level of pattern stability in normal times. However, when nearing crises, these models will fail differently but around a cluster of issues. Hence, the error term in such times is most remarkable.
- **Triangulation and Patterning Emerging Change:** When we see multiple methods going wrong differently, its' now time to narrow it down. This means establishing boundaries within which some curious clusters are morphing into new manifestations. Here, socio-psychological and historical understanding is crucial as it provides a context to the change and the reasons for it.
- **Folding In and Layering the Onion:** This is simply put, deep thinking and learning. We do have a high regard for learning and the accumulation of knowledge, but we need to have a much higher regard for the cultivation of thought. Whole world's accumulation is available on Google, but it is deep human thinking which can anticipate emerging LSLIREs. We have to avoid predetermining and study what comes, essentially applying 'active imagination'. Jung advocated what he called "*active imagination*" which is described as a state of reverie in which judgment is suspended but consciousness preserved. The key point is that the more deeply, broadly and consistently over time one folds in patterns to layer the onion, the more obvious becomes any individual pattern that does not fit. They may not fit because they are wrong, our understanding is wrong or they may be a precursor of coming system change.
- **Consider a Preference for Qualitative Insights to See Change Lacunae:** Werther suggests that a qualitative, more than quantitative, perception best illuminates a pre-event period. He argues that as stable and well-patterned systems begin to change in important ways (as their syndrome changes), it is easier to recognize this pre-change period qualitatively. Quantitatively, models are slower to respond to

[2] Werther; SOA 2013; Recognizing When Black Swans Aren't: Holistically Training Management to Better Recognize, Assess and Respond to Emerging Extreme Events

changes, data is limited or not credible and too heterogeneous at this time and the model results might not make much sense at emerging stage.

- **Focus on Seeing Undergirding Socio-Psychological and Style Changes:** This means essentially attacking the source. External and internal factors are far more fluid then our underlying way of doing things, our underlying socio-psychological and philosophical makeup. When normally stable qualitative psychological patterns start consistently changing, it is time for emerging new risks to be created and come up on the surface.
- **Use an Understanding of How Things and Processes are Embedded:** Learn how everything connects to everything else and how particular configurations enable or repress options and possibilities. In this matter, network theory and its application is essential.
- **Learn to Understand How Things and Processes are Entangled:** Entanglement here refers to previously one force breaking into two different forces but still behaving relatively similar. Cultures, societies, , individuals, , organizations and so forth are entangled systems: They always carry their legacy forward. They are not random, unconstrained or free, but are shape-shifting (morphing) specifics going into their future. From knowing the historical context and deep understanding we have to arrive at the entanglement features of the systems. From what did they derive, and how does this shape them going forward?

Main shortcoming of predictive modeling is that we change only few assumptions and keep rest of them constant so that they are not dynamic enough. Of necessity, any line of disciplined inquiry focuses on certain operative variables and determinants, and freezes others. Often the ground thus frozen is that very territory which is problematical from the standpoint of emerging risks.

One very powerful technique for actuaries is to utilize quantitative models and qualitative methods simultaneously. Models and statistics create discipline and uniformity for actuaries and analysts and is a powerful source for 'herding' toward similar opinions. As Carl Jung says that "the statistical method shows us the facts but does not give us a picture of their empirical reality. Actuary can use the quantitative models to arrive at the 'normal' state of opinion and use qualitative, deep and context specific explanations to understand and explain deviations from the normal standards. One aspect of quantitative models is particularly useful before emerging LSLIRE arise is breaking down of models and increasing diverge between stories of analysts explaining these deviations. More analysts will start feeling that something is wrong but cannot identify through their mainstream models what is wrong specifically.

Hence, our modeling regime combines to form a holistic integrated way to create the future and lead to explosion of insights.

At the same time, we recognize that all models and projections are ultimately a small replica of the complex reality and is no substitute for timely management. We realize that reality will always be different than our expectations and projections. The aim hence, is not to become a better forecaster but a better evolver. We will integrate our quantitative modeling with qualitative profiling and behavioral analysis of psychology of the customers. This means a) having a close eye on emerging and futuristic trends, risks and opportunities, b) augmenting quantitative modeling with qualitative judgments c) using behavioral and sociological analysis to limit cognitive biases in ourselves as well as to keep the customers motivated and engaged.

At its heart, the failure of the current insurance systems is a failure of imagination; in what could become instead of what is only being currently done. We have to inspire people, systems, and practitioners and spark their imaginations to bring in more creative solutions to our current problems.

3. Other Key Aspects; Deeping Modeling Expertise

We have to build our modeling more complex but not our products more complex like contingent insurance, derivatives, finite and financial reinsurance etc. Financial crash of 2008 taught us that both products and modeling can be made too complex to hide bad fundamentals. We have to avoid this at all costs. Customers hate the complexity associated with buying insurance and when trying to file a claim. Simplify the experience for the user and keep the platform engaging.

Automated machine intelligence/learning (AML)

Despite the hype and glory associated with quantitative modelers like data scientists, actuaries, quants, and many others, they face a conundrum which automated machine intelligence sets out to solve. The conundrum is the gap between their training and what they should be doing compared to what they actually do. The bleak reality is most of time gets taken by monkey work (work that any monkey can do instead of an intellectually trained and competent human being) like repetitive tasks, number crunching, sorting out data, cleansing data, understanding it, documenting the models and applying repetitive programming (being spreadsheet mechanics too) and good memory to remain in touch with all of that mathematics. What should they be doing is being creative, producing actionable insights, talking with other stakeholders to bring about concrete data-driven results, analyzing and coming up with new 'polymath' solutions to existing problems.

Automated machine intelligence/learning (AML) takes care to reduce this huge gap. Instead of hiring a team of 200 data scientists, a single or few data scientists using AML can utilize fast modeling of multiple models at the same time because most of the work of machine learning is already automated by AML like exploratory data analysis, feature transformations, algorithm selection, hyper parameter tuning and model diagnostics. There are a number of platforms available like DataRobot, IBNR Robot, Nutonian, TPOT, Auto-Sklearn, Auto-Weka, Machine-JS, Big ML, Trifacta, and Pure Predictive and so on.[3][4]

Through this way AML frees up data scientists to be more human and less cyborg-Vulcan-human calculators. Machines are delegated to what they do best (repetitive tasks, modeling) and humans are delegated to what they do best (being creative, producing actionable insights to drive business objectives, creating new solutions and communicating them). [5]

Emerging Risks

[3] https://www.kdnuggets.com/2017/07/automated-machine-learning-paradigm-shift.html
[4] https://www.analyticsvidhya.com/blog/2016/05/19-data-science-tools-for-people-dont-understand-coding/
[5] https://www.datarobot.com/blog/let-data-scientists-human/

Aside from using better tools, we want to utilize a more proactive mentality as well. The traditional reactive approach to modeling has a number of limitations including that time is of the essence for new and emerging products and emerging risks/liabilities/products which are too many and too pervasive; like drone insurance, driverless cars insurance, telematics, cyber insurance, impact of genetic engineering and antibiotic resistance, personalized medicine, 3D printing, growing epidemics, impact of crypto currencies, impact of new generation millennial and disruptive fintech', impact of rapidly changing weather, impact of crowd sourcing and collective wisdom like prediction markets and others. Changing market conditions are also there; permanent lower prices of oil, increase in alternative investments like impact investing, rise in Islamic finance like Islamic insurance or takaful and so on.

If we keep waiting until credible data emerges, we will forever remain behind and less influential than others as changes are very rapidly oscillating and it seems that there will always be new emerging products and landscapes. New risks, products and liabilities are emerging and are becoming antiquated before they can even ossify. Constant revolutionizing of technology constantly keeps our social relations in everlasting uncertainty. Pre-emptive action and pro-active in the nick-of-time involvement is now perhaps the only way for us to go about dealing with the rapid fast-moving present and future. The STEEP framework for emerging risk analysis is useful for us where the acronym STEEP means Social, Technological, Economic, Environment and Political.

It is emphasized that emerging risks do not suddenly appear from nowhere and that there are always possible leading indicators, even though they may be rare and difficult to comprehend. Emerging risks are the contextual product of an evolutionary process these take time to develop and reveal themselves. A vital composition of emerging risk is the combination and integration of existing risks. The combined symptom can be assumed as an emerging risk but actually it is deeply rooted in existing risks as well; it is the potent and dynamic configuration and integration of existing risks that often give rise to new risks. One important aspect to monitor is performance of insurtechs; for instance, do we see that on-demand insurance is becoming unsustainable? Do we see lemonade's hype unfolding leading to decrease in interest in whole insuretech sector?

Blockchain

Historically and even presently, we induce trust upon transactions by recording their trail to prove the transactions' origin, destination, quantity and history. This requires huge time and efforts on part of many professions like lawyers, auditors, quality inspectors and many support functions.6 This in turn, causes humans to kill their creativity by becoming number-crunchers doing manual verifications to and fro, causes transactions to be expensive, inaccurate and expensive. Too much human suffering and Dukkha has been faced by many humans doing monotonous repetitive jobs just to create trust in these transactions. As knowledge is power, important information is kept hidden by those in power in determent to the masses. Blockchain is allowing us to potentially 'cut through all of this nonsense' of the middlemen and give power to the people through technology instead which is the chief goal of the fourth industrial revolution.
To make the process further seamless, agile, robust, invisible and as easy as a child playing, blockchain technology is used with smart contracts that execute itself when the conditions meet. This new P2P insurance model is doing away with traditional premium payment using instead a digital

[6] Blog at IOTA Organization; Carsten Stocker; Feb 22,2017; Automating Machine transactions and building trust in the fourth industrial revolution; Available at: https://blog.iota.org/automating-machine-transactions-and-building-trust-in-the-4th-industrial-revolution-d3219a157396

wallet where every member puts in their premium in an escrow-type account only to be used if a claim is made. In this model, none of the members carry an exposure greater than the amount they put into their digital wallets. If no claims are made all digital wallets keep their money. All payments in this model are done using bitcoin further reducing transaction costs. Teambrella7,8 claims to be the first insurer using this model based on bitcoin. Indeed, Teambrella is not alone. There are many blockchains based startups targeting peer to peer insurance and other areas of human activity. Some of them are:

1. Etherisc
2. Insurepal
3. AIgang
4. Rega Life
5. Bit Life and Trust
6. Unity Matrix Commons

Thus, a lot of crowd wisdom is utilized in this as the insurer 'Learns from the people, plans with the people, Begins with what they have And Builds on what they know' (Lao Tze).

Instead of an actuary maximizing profit for the shareholders, sitting isolated from ground realities, lacking skin in the game, and have far less access to awareness (i.e., data) of people relative to their peers, this peer to peer empowers the crowd and taps in into their wisdom (instead of wisdom from books) which is far better. There are also no unfair pricing practices here like rating based on gender, pricing optimization which charges you higher if you are less likely to shift to another insurer and vice versa. The giant insurer cannot know you more than your peers, it's as simple as that.

These same peer to peer insurance can be carried out on non-blockchain based distributed ledgers too like IOTA, Dagcoins and Byteballs with additional technological benefits of these new ledgers over current blockchain. These digital tokenization startups have the promise to radically reinvent business models where transactions, pooling and just about anything gets done for the community and by the community in an automated fully trust worthy manner with no oppressive middlemen like governments, capitalist businesses, social institutions and so on. Peer to Peer Insurance is just one part of the whole program.

Smart contracts have built in conditions with them which are automatically triggered when the contingency happens and claims get paid instantly. The huge need for labor force with high qualifications but essentially doing clerical work is removed altogether to build a sleek autonomous organization of the future. The oppressing middleman of 'shareholders' are avoided which means that consumer interests are acted upon by providing convenience, low prices and good customer support. In this peer to peer setting, the benefits goes to the community instead of the shareholder. IoT provides the main source of data to these pools to develop protocols when to release claim payment and when not to. The same tokenization means that anyone anywhere can have access to the insurance pool instead of being limited by geography and regulations.

[7] Teambrella company website available at: https://teambrella.com/
[8] Teambrella: A Peer to Peer Insurance System. Available at: https://teambrella.com/WhitePaper.pdf

To see the explosive impact it can have, imagine replacing centuries old Lloyd's insurance market. A startup, TrustToken9,10 is trying to create a trust economy to carry out transactions USD 256 trillion, which is the value of all real-world assets on earth. The current transactions take place in outdated models with limited transparency, liquidity, trust and a lot of problems. Carrying out these transactions using digital ledgers like blockchain is far more lucrative through the potential of tokenization. Tokenization is the process through which real world assets are converted into digital tokens. TrustToken is making the bridge between digital and real worlds through tokenizing real world assets in a way that is acceptable in real world too and is 'legally enforced, audited and insured'. This is done through creation of 'SmartTrust' contract that guarantees ownership with legal authorities in the real world, and also implements any necessary action when contracts are broken, including reposing, charging criminal penalties and much more. A decentralized TrustMarket is available for all stakeholders to gather and negotiate the prices, services and TrustTokens are the signals and rewards parties receive for trustworthy behavior, to create an audit trail and to insure the assets.

Whether TrustTokens are able to carry out sound insurance is a matter up for debate but we can already see this in the centuries old Lloyd's market. In Lloyd's market, buyers and sellers of insurance and underwriters gather together to carry out insurance. An administration of Lloyd's funds monitors their various syndicates and provides capital adequacy to absorb the shocks that come from insuring too. TrustMarket has the potential to become the modernized version of Lloyd's market but it is too early to determine its precise success. TrustToken can open up the economy and create better value and lesser costs and corruption in real world assets, especially in real estate, insurance and commodities that create too much power in the hands of the very few.

That means that these blockchain peer to peer insuretech should deeply study Lloyds' insurance market as they aim to replicate this and not become a traditional insurance company.

Machine Teaching

Machine teaching, is perhaps the most exponential trend that we are seeing which can allow M2M economy to prop up exponentially from humble beginnings to become a dominant feature of our everyday lives. Imagine! Machines not only transacting with each other and other platforms like servers and humans but also teaching each other. This has already happened with Tesla Model S's autopilot feature. The human driver acts as the expert teacher to the car but the cars share these data and learning between themselves radically improving their experience in extremely short time. Now one IoT device is not an isolated device that will have to learn everything from scratch on its own; it can leverage the mass learning learned by other similar IoT devices worldwide as well. This means that intelligent systems of IoT trained by machine learning are not just becoming smarter; they are getting smarter faster over time in exponential trends.

This 'Machine Teaching' has huge advantages in that it lowers the training time required, bypasses the need to have massive training data and allows machines to learn by themselves to improve the user experience. This Machine Teaching can be sometimes collective like self-driving cars sharing

[9] Futurism.com; Kristin Houser; Jan 8, 2018; We need to expand access to USD 256 trillion in real world assets. Available at: https://futurism.com/need-expand-access-256-trillion-real-world-assets/

[10] TrustToken's website: https://www.trusttoken.com/

and learning together in sort of a collective hive mind, or it can be adversarial like two machines playing chess against itself, one machine acting as the fraud and the other machine as the fraud detector and so on. The machine can also teach itself by playing simulations and games against itself without need of any other machine. AlphaGoZero has done exactly that. AlphaGoZero did not use any training data and played against itself and then defeated the AlphaGo which was the AI that had defeated the world's best human Go players (Go is complicated Chinese chess). The feeling that chess grandmasters had of watching AlphaGoZero play was like an advanced alien super-intelligent race playing chess.

The applications from this are staggering; hyperloop (very fast train) based tunnel pods communicating with each other, autonomous ships, trucks, whole fleets of drones running on swarm intelligence and the living city learning from itself through smart grid interactions. This along with other innovations occurring in the fourth industrial revolution of Artificial Intelligence can eradicate current health problems, many social problems like absolute poverty and allow us to colonize Moon and Mars.

4. Final Words

We should attempt to creatively synthesize the many pluralistic approaches as well as focus more on synergistic interpretation of findings of these pluralistic researches.

We can recognize that though we cannot precisely predict black swans but forecasting emerging liabilities and their ratemaking can be a professional-character building experience in itself where we train to be better evolvers rather than better predictors alone.[11]

We can highlight that while recognizing facts (in form of quantitative analysis), it seems as if we only tend to scratch its surface as data, on its own, highlights results; whereas there are plenty of processes that culminate in data generation as well as modeling methodology in the first place. There is an incredible depth once when we start looking beyond the facts into fact-making itself; and this is where expert judgment and qualitative profiling can prove invaluable to guide the modeling exercise.

Agile Risk culture is foremost for any modeling exercise because complex systems like financial and insurance sector are not solely run by quantitative numbers, but by the underlying human psychology as well. It is up to the risk culture to not antagonize in binary opposites like complex/simple, good/bad etc., but to reach the middle ground to converge communication and mentalities between different stakeholders.

In the end, it is useful to keep a few sobering meditations in mind[12,13]:

- We suffer too profoundly even from small data glitches.
- Better than many complicated equations are few statements that give clarity to shareholders
- The experience of all deep datasets is slow. They must wait long until they know what has fallen into their depths. Machine learning can lower that waiting time.
- Generally, there is either over-reliance on data and models or negligible reliance on them. We have to be familiar with the golden mean that resides between two vices. So here our data and modeling orientation should be in between the extremes of reliance on only opinions and only data and models.
- Unless one considers intention—philosophy, cognitive system, behavioral bias, etc.—used in building data, models and expert's analysis, and implications, one can be missing the big picture already.
- Provide historical data to limit the amount of work required for attaining a context for the data but data should be adjusted to reflect current conditions, not historical circumstances.[14]

[11] Mills, Allan: Society of Actuaries (2010): Complexity Science: an introduction and invitation for actuaries
[12] Data Science Central; June 20, 2016: Syed Danish Ali: Modeling Meditations. Available at: https://www.datasciencecentral.com/profiles/blogs/modeling-meditations
[13] Data Science Central; June 20, 2016: Syed Danish Ali: Inspiring Imagination in data science qualitative profiling. Available at: https://www.datasciencecentral.com/profiles/blogs/inspiring-imagination-in-data-science-qualitative-profiling
[14] Philippos Papadopoulos April 2015 The Zen of Modeling

- Focus on developing a ratemaking plan, not numerical premium and projections only.[15]
- Know your context.[16]
- Beware of qualitative shifts[17]
- Know how the model results will be used[18]
- Do not anthropomorphize models. Anthropomorphism is the tendency to characterize animals, objects, and abstract concepts as possessing human-like traits, emotions, and intentions. Models are not reality or real human social behavior. At best models are idols[19]; at worst a distraction and cause for herding.

This report sought out to formulate a modeling regime that is broad, comprehensive and agile for the future. This framework can be utilized in deploying models while ensuring that the bigger picture remains intact and insight. The real implementation of this regime can only come from a change in mentality and approach from 'doing things right' to 'doing the right thing'.

[15] Ibid
[16] Werther; SOA 2013; Recognizing When Black Swans Aren't: Holistically Training Management to Better Recognize, Assess and Respond to Emerging Extreme Events
[17] Ibid
[18] Ibid
[19] Wilmott, P. & Derman, E, 2009. "The Financial Modelers' Manifesto