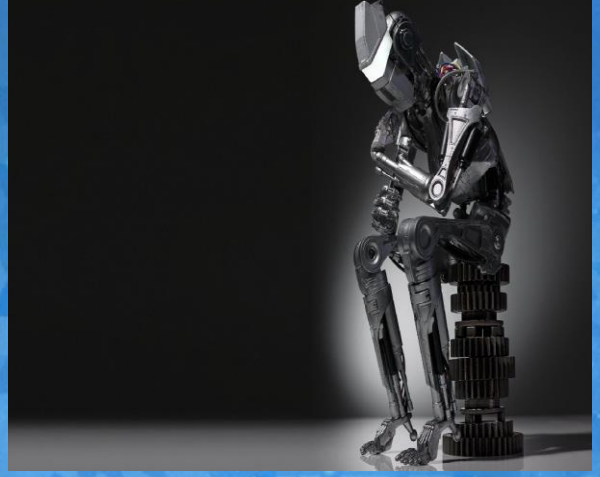
ITU KALEIDOSCOPE
SANTA FE 2018
Machine learning for a 5G future



Ethical Framework for Machine Learning

Charru Malhotra

Associate Professor (e-Governance and ICT)

Indian Institute of Public Administration

New Delhi INDIA

charrumalhotra[at]gmail.com

26-28 November
Santa Fe, Argentina

Agenda

01	Introduction	• Ethics : Right V/s Wrong • ML : Learning through Algos
02	Ethical Issues in ML	• M/C replace Humans- *so who rules?* • Unethical Hacks ! • Biased Databases …
03	Proposed Ethical Framework (EF)	Combine EQ & SQ with IQ to creates ML Algos: 'Mechanical Learner,' 'Cognitive Learner ', 'Ethical Learner', 'Ethical Master/ Leader'
04	Implementation Strategy	• Changes at 3 levels- National, Organisational , Design Approach
05	Way Forward	• Ethics' Guidelines, Collaborations, Authorities : to check compliance of EF in ML

Dreams or Reality : 1985-Now

Use of AI /ML/ Big Data Algorithms to detect Anomaly in any system **to predict unexpected disasters (Soil Health in Odhisha)** (ToI, Sep 27, 2017)

> Augmented and Virtual Reality will bring the showroom and test drive to the customer
>
> PRATAP PAT JOSHI
> CIO, *Mercedes-Benz India*

Apps, drones likely to map infrastructure projects

TNN | Dec 6, 2017, 04:19 IST

How Data Analytics Forecasted Riots/ Foreign intrusions

Using correlation metrics on social media posts, hashtags, news articles or posts by religious groups & intelligence data from police agencies... The next intrusion is likely to happen in the wee hours of the morning, around 2 am, about 11 or 12 days after a herd of cattle meanders close to the Arunanchal border.....

> Digital personal assistants will be the new financial advisors
>
> MURARI SRIDHARAN
> CTO, *BankBazaar*

> The new age of meeting rooms will have automated assistants capable of completing tasks
>
> DEBASHISH SINGH
> CIO, *Mphasis*

Robots excavators are coming to Bengaluru soon. BBMP- *Bruhat Bengaluru Mahanagara Palike* – the urban local bod – has plans to procure 'robot multi-purpose excavators' for dredging and cleaning of drains in Bengaluru

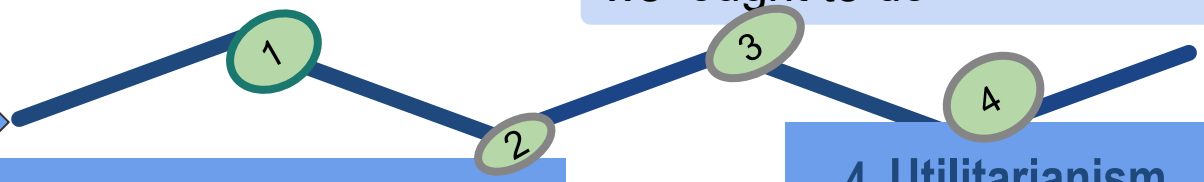
(Sep 5, 2017)

ETHICS

Morals | Values | 'Ethos' | Rules of Conduct

what is right and wrong ….how to act on it …to have happiness for… self and society

Evolution of Ethics



1. Virtue ethics

Aristotle (384-322 B.C.E)
virtuous ways benefit possessor & society

2.Social Contract Theory

Thomas Hobbs et. al (1588-1679)
Moral rules to govern relationships among citizens

3.Kantianism

Emanuel Kant (1724- 1804)
- duty central to morality
-not what 'we do' but what we 'ought to do'

4. Utilitarianism

Jeremy Bentham (1748-1832)
John Mill (1806- 1873)
Examine right & wrong based on consequences of act / rule.

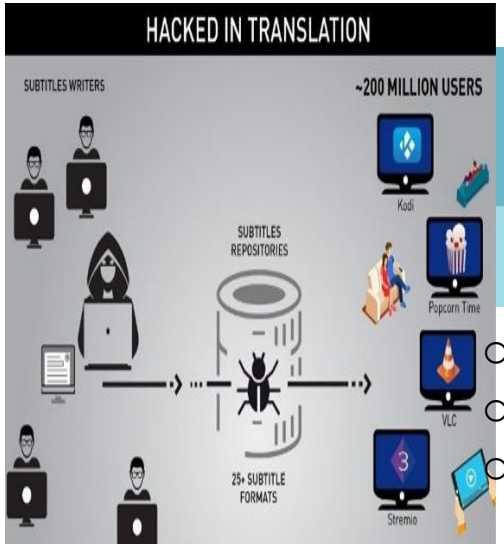
SOME ETHICAL ISSUES

Machines Replacing Humans

- *Decisions left to m/cs esp complex, huge data*
- *Whose judgment is better*
- *Replace human existence?*

Unethical Hacks

Machines recording, analysing preferences



- *internal frauds ?*
- *remote hacking ?*
- *perception-modifications?*

Biased DataBases

Machines learning from 'unclean' sources

- *Non-inclusive policies?*

Biased DataBases

(* estimated to be faced by 20% of all women)

(**faced by 1 of 3 women in the world)

Traffic cam mistakes CEO face or ×

https://inshorts.com/en/news/traffi

2016-Self Driven Uber Cars jump red lights;
Smart Home catches fire in Blackberg-Penn

2016-Semi-automated drones (LETHAL AUTONOMOUS WEAPON SYSTEMS)deployed between demilitarize zone of South and North Korea

2017-A Russian Robot,Promobot R77 wants to go Free; Sophie wants to kill the human race

Reality itself is an AI Simulation that is malfunctioning

2016-Online Feud of Wiki-Bots

Traffic cam mistakes CEO face on bus for her, charges for law break

short by Pragya Swastik / 08:51 am on 28 Nov 2018,Wednesday

An artificial intelligence-based traffic camera system in China recognised the face of a woman CEO, Dong Mingzhu, on a bus advertisement and mistakenly charged her for crossing a road during a red light. The image of Mingzhu, CEO of China's biggest air-conditioner maker Gree Electric Appliances, was displayed on a city billboard used to shame those who break traffic laws.

read more at SCMP

2017-FACEBOOK has revealed how its artificial intelligence created its own language and began nattering in private ….
Google Photos 'goof' up in facial recognition…
MS Chatbot - " Hail Nazi" tweets !!!

INTRODUCTION

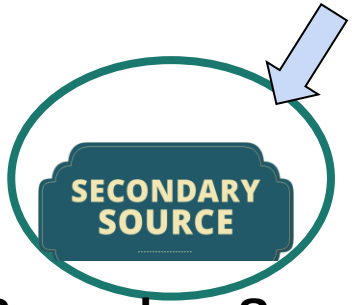
The study is an interplay of ML (the hard science) and Ethics (the soft science)

AIM

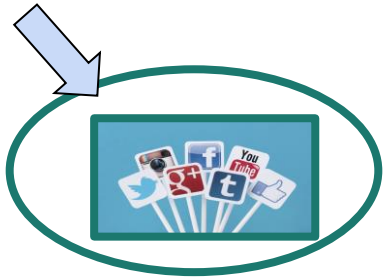
How to minimize /eliminate bias in ML algorithms for making predictions and future decision making

Methodology

Descriptive, Exploratory & Analytical



+



Secondary Sources

Academic Literature of ML & Ethics

Social Media

Netflix, Youtube, Blogs,Fiction too

Proposed Ethical Framework

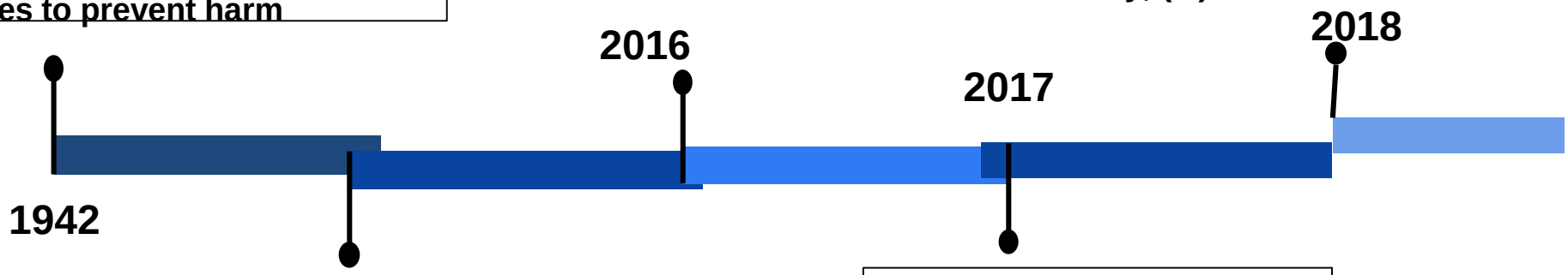
DESIGNING ML ALGOs- Consider Emotional Quotient (EQ) & Spiritual Quotient (SQ) too apart from the Intellectual Quotient (IQ)

Existent Ethical Frameworks

-Isaac Asimov ("Runaround" , 1942)
-3 laws of robotics",
-human-like robots to act servants & need programming rules to prevent harm

Satya Nadella (2016) (CEO, Microsoft), 10 essential rules for an ideal AI

Sundar Pichai (2018) (CEO,Google) -seven principles that guide Google
(i) socially beneficial; (ii) avoid unfair bias; (iii) built & test for safety; (iv) accountable etc.



Wallach & Allen (2009) suggest top-down, bottom-up & hybrid approach for machine morality.

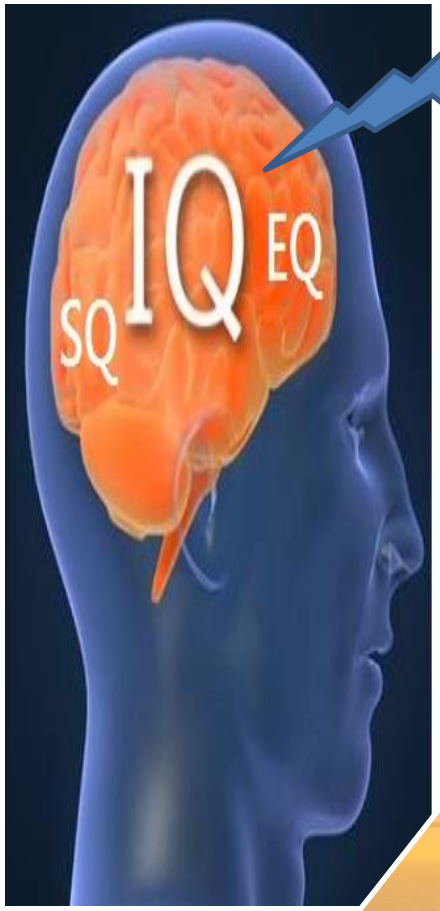
- Top-down- moral frameworks of philosophers,
- bottom-up- machine learn through manipulation like a child

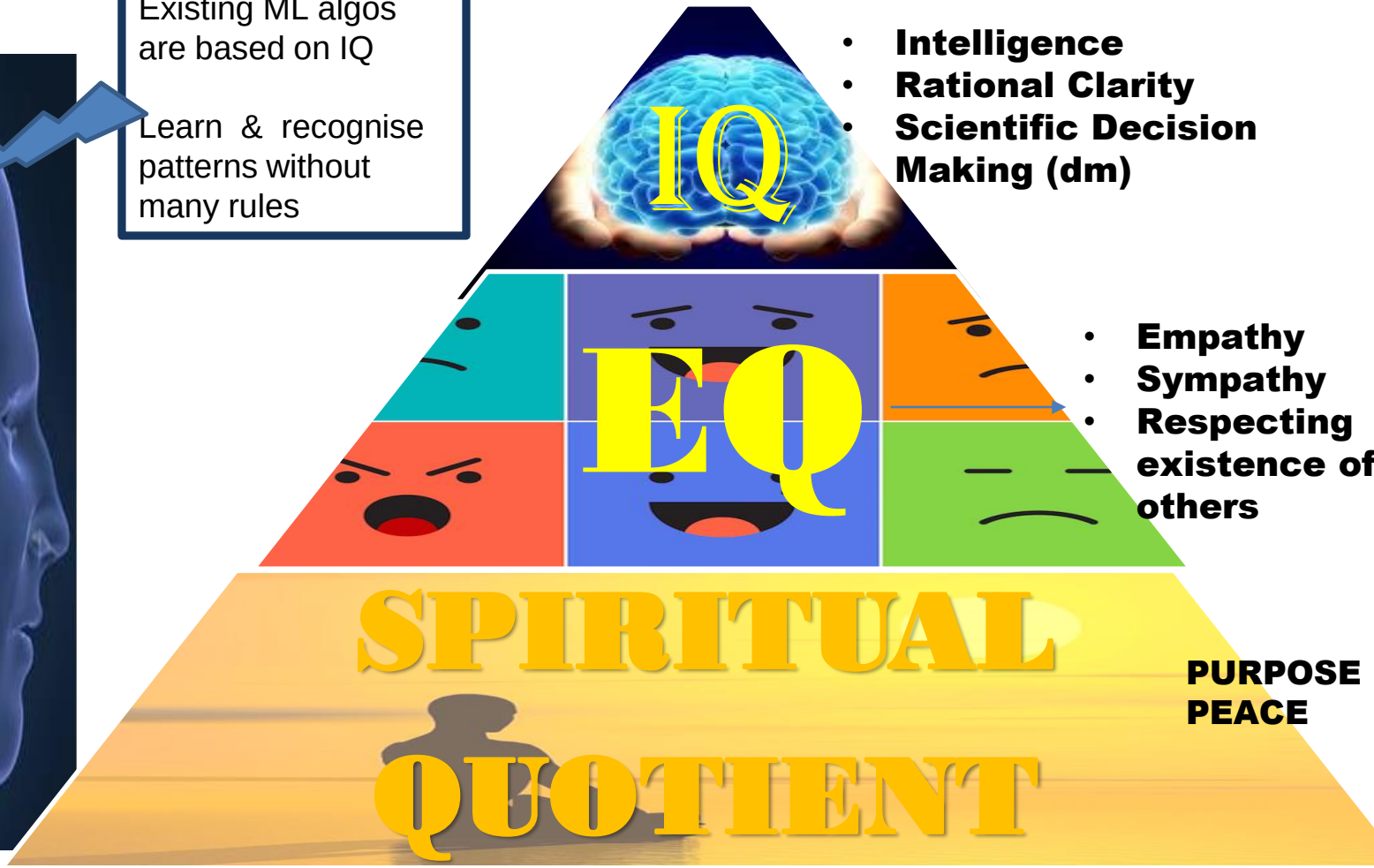
Comest of UNESCO:

proposes technology-based ethical framework on robotics

Existing ML algos are based on IQ

Learn & recognise patterns without many rules



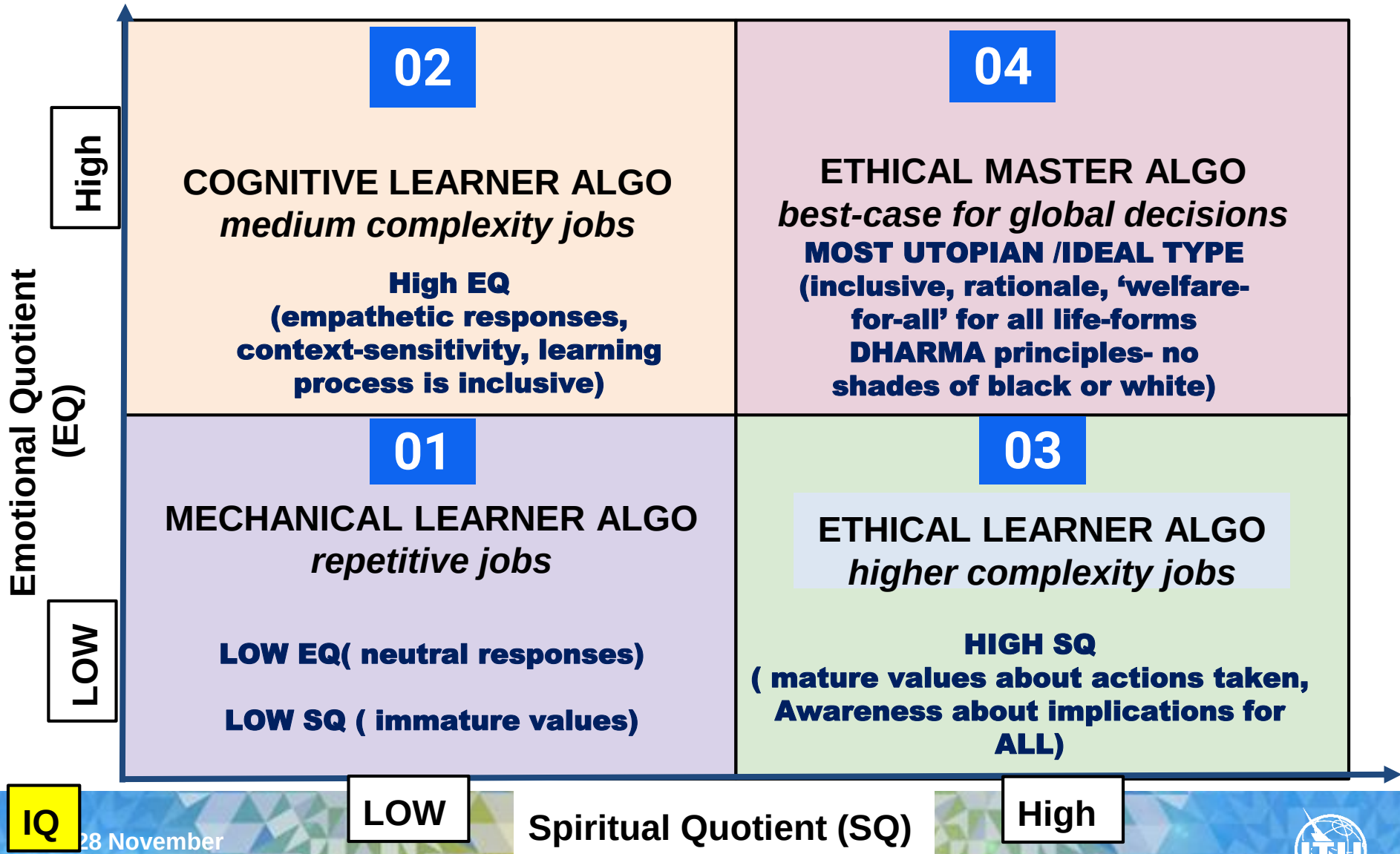


- **Intelligence**
- **Rational Clarity**
- **Scientific Decision Making (dm)**

- **Empathy**
- **Sympathy**
- **Respecting existence of others**

PURPOSE
PEACE

Proposed Ethical Framework



Proposed Ethical Framework

'welfare-for-all'

02
COGNITIVE LEARNER ALGO
medium complexity jobs
empathetic, inclusive, contextual

- Electronic Assistants (Alexa)
- Predicting User Behaviour
- TeleMedicine
- NLP
- Facial Recognition

04
ETHICAL MASTER ALGO
V. complex decisions by country leaders

- *Gia* Sensitive Global Policies
- *DSS/ ES for Global Leaders & Policy Makers*

01
MECHANICAL LEARNER ALGO
repetitive jobs neutral responses & immature values

- Assembly Line Production
- Restaurant Robots
- Weather Forecasting
- Network mgmt, Intrusion /Fault Detection, Clustering

03
ETHICAL LEARNER ALGO
higher complexity jobs- mature values, awareness of implications for all; community dm

- Industry 4.0
- Military Drones ; AV
- CRM
- Humanoids-Sophia/ Rashmi
- Dynamic DM for resource-mgmt. – Traffic Mgmt; SMART HOMES

Emotional Quotient (EQ): High / LOW

Spiritual Quotient (SQ): LOW / High

IQ

Implementation Strategies

National Level COE

AUTHORITY; REGULATIONS; ETHICAL INDICES/STDs/ COLLABORATIONS

- Work with multi experts
- National Ethical Framework
- Support Policy-enhancements
- Standards for big data
- Encourage research in emerging areas by global / local alliances

Organizational Level

'ETHICS are a MUST'

- Develop, Follow & Monitor Ethical F/w Templates
- Changes in hiring, CB- training, promotion & feedback
- Incorporation of ethical aspects at design stage
- Assigning an Ethics engineer

Design & Authority

"Humans stay the Master"

- Design using EF-Citizen-Centric approach
- Machine should be obedient to Owner, Design & Authority
- Provide some autonomy to machine but basic program should be locked so that data cannot be erased o hacked.

Way Forward : QoV (Quality of Values) in ML

Collaborate → **Formulate** → **Intervene** → **Monitor & Control**

Associations of leaders, academia, ethics- advocates , practitioners at global level

Multi-disciplinary diverse design teams at local level

Based on the proposed EF, formulation of International guidelines for QoV (just like we have QoServices , Qo Experiences)

Guidelines

Interventions at National level like

- **Regulatory Framework**
- **Standards /Authority**
- **EF Templates for ML programming**

National authority to check ethical indices of ML innovations for QoV ; Organisations to check compliance

About the Speaker

IT Consultancies with International agencies

WTO, WCO, GENEVA (2018)
UNCTAD (United Nations Conference on Trade and Development), GENEVA (2018)
The Swiss Embassy-India (2017)
The World Bank (2000 – till Now)
The Economist Intelligence Unit , HK-2017)
IIMD- International Institute for Management Development (Lausanne, Switzerland) (2016)
ADB - Asian Development Bank (2014)
Panthéon Sorbonne, PARIS (2016,2014)
Wollongong University, AUSTRALIA (2011)
The Winrock International (1999)

National Agencies

Central Universities - JNU,DU, BHU,IIT-D, Rajasthan University
Central ministries- meity ü**, DoPT, MHA, MEA,DST, CPWD, DARP&G**
STATE GOVTS – KARNATAKA, HARAYANA
INDUSTRY ASSOCIATIONS - NASSCOM
MEDIA/ Civil Society – The LokSabha, DEF

28 years of IT Journey

-From MIS/IT/ICT realm (MCA ,MCSD) to the 'real world' (IIT-Delhi) of **600 villages of India (2005-2009)**
-NE (Sikkim, Shillong, Mizoram)
-South (TN, Kerala, K'tka)
-North (Punjab, Haryana, Rajasthan)
-Centre (MP, 36garh,UP)

Generated

Almost 70.5 millions for IIPA in last 5 years

Guided

20 Mphils/PhD

- 30 Papers/ chapters/ articles
- 6 handbooks
- 5 reports

"For the longevity of natural life on earth it is imperative ***we design algorithms for machines that become aware of what harmony, empathy, conservation of resources*** *is… that knows that WE ARE ALL ONE"*

'Gracias'

Charru Malhotra

charrumalhotra@gmail.com

Vinod Kotwal

vinod.kotwal@nic.in

Surabhi Dalal

intern5.icm@mea.gov.in

Special Thanks to- Ms.Ishita Malhotra and Mr. Gaurav Khandelwal; Presentation Courtesy- Shilpa, Nishtha, Heeba, Rashmi