

**COURSE OBJECTIVES:**

- Study the morality and ethics in AI
- Learn ae Ethical initiatives in the field of artificial intelligence
- Study about AI standards and Regulations
- Study about social and ethical issues of Robot Ethics
- Study about AI and Ethics- challenges and opportunities

**UNIT I INTRODUCTION 6**

Definition of morality and ethics in AI-Impact on society-Impact on human psychology-Impact on the legal system-Impact on the environment and the planet-Impact on trust

**UNIT II ETHICAL INITIATIVES IN AI 6**

International ethical initiatives-Ethical harms and concerns-Case study: healthcare robots, Autonomous Vehicles , Warfare and weaponization.

**UNIT III AI STANDARDS AND REGULATION 6**

Model Process for Addressing Ethical Concerns During System Design - Transparency of Autonomous Systems-Data Privacy Process- Algorithmic Bias Considerations - Ontological Standard for Ethically Driven Robotics and Automation Systems

**UNIT IV ROBOETHICS: SOCIAL AND ETHICAL IMPLICATION OF ROBOTICS 6**

Robot-Roboethics- Ethics and Morality- Moral Theories-Ethics in Science and Technology - Ethical Issues in an ICT Society- Harmonization of Principles- Ethics and Professional Responsibility- Roboethics Taxonomy.

**UNIT V AI AND ETHICS- CHALLENGES AND OPPORTUNITIES 6**

Challenges - Opportunities- ethical issues in artificial intelligence- Societal Issues Concerning the Application of Artificial Intelligence in Medicine- decision-making role in industries-National and International Strategies on AI.

**30 PERIODS**

## UNIT -1 INTRODUCTION

### ARTIFICIAL INTELLIGENCE (AI):

'Artificial Intelligence (AI) refers to **systems that display intelligent behaviour** by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals.

### EXAMPLES: (OR) AI BASED SYSTEM TYPES:

#### 1. PURELY SOFTWARE BASED:

- A) voice assistants,
- B) image analysis software,
- C) search engines,
- D) speech and face recognition systems

#### 2. AI EMBEDDED IN HARDWARE:

AI can be embedded in hardware devices.

- A) advanced robots
- B) autonomous cars
- C) drones
- D) Internet of Things applications.

### AI TYPES:

NARROW AI	AGI: ARTIFICIAL GENERAL INTELLIGENCE
Refers to the intelligence in current AI systems and robots that are <b><u>capable of undertaking one or few specialised task.</u></b>	- A <b><u>long-term goal</u></b> of AI and robotics research - which would be <b><u>comparable to human intelligence</u></b>

**Machine learning** is the term used for AIs which are **capable of learning** or, in the case of **robots, adapting to their environment.**

Supervised learning systems	Unsupervised learning
Supervised learning systems generally make use of <b>Artificial Neural Networks (ANNs)</b> , - Trained by presenting the ANN with inputs (for instance, images of animals) each of which is tagged (by humans) with an output (i.e. giraffe, lion, gorilla). - <b><u>This set of inputs and matched outputs is called a training data set.</u></b>	Unsupervised learning has <b><u>no training data;</u></b> -The AI (or robot) must figure out on its own how to solve a particular task -Generally by trial and error.

## DEEP LEARNING:

The term deep learning refers to supervised machine learning systems with large (i.e. many-layered) ANNs and large training data sets.

## MORALITY AND ETHICS

Ethics are moral principles that govern a person's behaviour or the conduct of an activity.

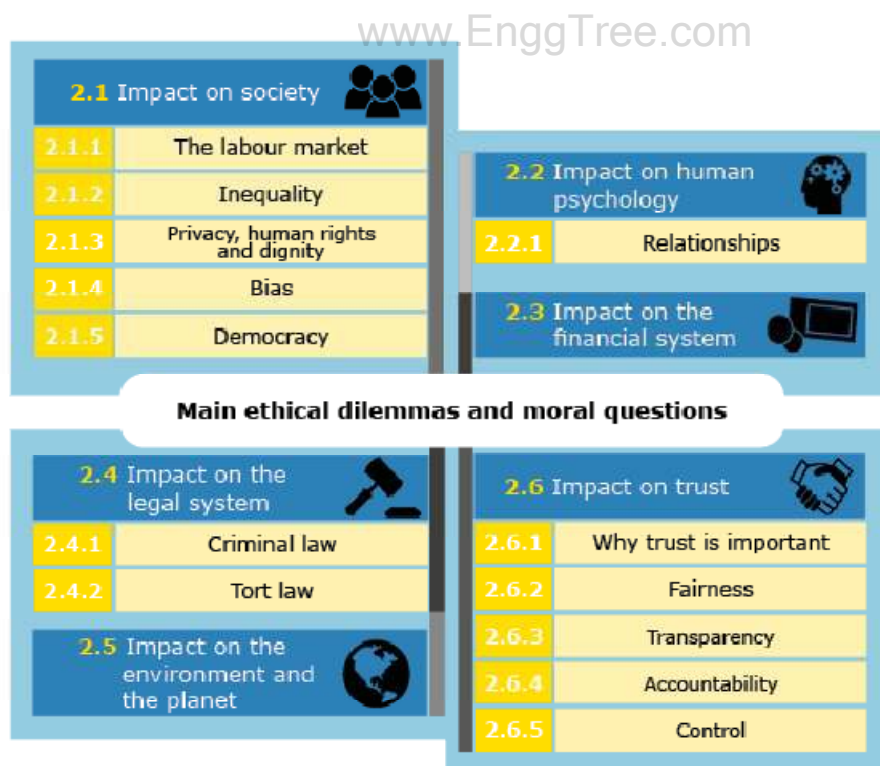
Example Ethical Principle: to treat everyone with respect.

## Kant's PRINCIPLE:

'act as you would want all other people to act towards all other people'

## AI ETHICS:

AI ethics is concerned with the question of how human developers, manufacturers and operators should behave in order to minimise the ethical harms that can arise from AI in society, either arising from poor (unethical) design, inappropriate application or misuse.



-----X-----X-----

## IMPACT ON SOCIETY:

### 1. LABOUR MARKET 2. INEQUALITY 3. PRIVACY

1. Labour Market: AI and robotics have been predicted to destroy jobs and create irreversible damage to the labour market.

#### **Labour Market: Impact on economic growth and productivity**

Economists are generally enthusiastic about the prospects of AI on economic growth.

Robotics added an estimated **0.4 percentage points of annual GDP growth** and labour productivity for 17 countries between 1993 and 2007

#### **Labour Market: Impact on the workforce**

- 48 percent believed that robots and digital agents would displace significant numbers of both 'blue' and 'white' collar workers, with many expressing concern that this would lead to vast increases in income inequality, large numbers of unemployable people, and breakdowns in the social order (Smith and Anderson, 2014).
- However, the other half of the experts who responded to this survey (52%) expected that technology would *not* displace more jobs than it created by 2025.
- For example, the share of workers **in leisure and hospitality sectors** could increase if household incomes rose, enabling people to afford more meals out and travel
- According to their analysis, **telemarketers, watch repairers, cargo agents, tax preparers, photographic process workers, new accounts clerks, library technicians, and data-entry specialists** have a **99 percent chance** of having their jobs computerised.
- At the other end of the spectrum, **emergency management directors, mental health social workers, health care social workers, surgeons, firefighter supervisors and dieticians** have less than a one percent chance of this.

#### **2. InEquality**

**In Equality in Revenues** will be split across fewer people, increasing social inequalities. Consequently, individuals who hold **ownership in AI-driven companies** are set to benefit disproportionately.

## Inequality: exploitation of workers:

AI needs huge training data sets:

Ex:1.manually tagging objects in images in order to create training data sets for machine learning systems (for example, to generate training data sets for driverless car AIs)

2. interpreting queries (text or speech) that an AI chatbot cannot understand.

One of the key ethical issues is that – given the price of the end-products – these temporary workers are being inequitably reimbursed for work that is essential to the functioning of the AI technologies.

## In Equality in: Concentration of power among elites

The accumulation of technological, economic and political power in the hands of the top five players – Google, Facebook, Microsoft, Apple and Amazon – affords them undue influence in areas of society relevant to opinion-building in democracies: governments, legislators, civil society, political parties, schools and education, journalism and journalism education and — most importantly — science and research.

## 3. Privacy, human rights and dignity:

The privacy and dignity of AI users must be carefully considered when designing service, care and companion robots, as these devices work in people's homes .

AI is already affecting privacy is via Intelligent Personal Assistants (IPA) such as Amazon's Echo, Google's Home and Apple's Siri.

These voice activated devices are capable of learning the interests and behaviour of their users, but **concerns have been raised about the fact that they are always on and listening in the background.**

## **Human rights**

If AI can be used to determine people's political beliefs, then individuals in our society might become susceptible to manipulation.

Political strategists could use this information to identify which voters are likely to be influenced to make a Party win.In India, sentiment analysis tools are increasingly deployed to gauge the tone and nature of speech online, and are often trained to carry out automated content removal.

-----X-----X-----

## IMPACT ON HUMAN PSYCHOLOGY

### **Impact on Relationships: Manipulation, Affect Cooperation , Affect Ethics**

AI is getting better and better at modelling human thought, experience, action, conversation and relationships.

As AI machines interact with humans, the impact on real human relationship should be studied.

### **1. Relationships:**

In the future, robots are expected to serve humans in various social roles: nursing, housekeeping, caring for children and the elderly, teaching, and more. People may start to form emotional attachments to robots.

### **Danger: Manipulation**

Social robots that are trusted could be misused to manipulate people ;

- for example, a hacker could take control of a personal robot and exploit its unique relationship with its owner to trick the owner into purchasing products.
- While humans are largely prevented from doing this by feelings like empathy (அனுதாபம்) and guilt (குற்ற உணர்வு), robots would have no concept of this.

### **Danger: Robots Can affect Cooperation among Humans:**

**EXPERIMENT:** small groups of people worked with a humanoid robot to lay railroad tracks in a virtual world.

<b>case 1: Robots were Programmed to Make occasional errors</b>	<b>Case 2 : Error Prone Robots used</b>
<b>Outcome: Human Work &amp; Cooperation Improved</b>	<b>Outcome: Robots Outperformed Human Groups</b>

**Danger: AI can also make Humans behave less productively and less ethically.**

### **EXPERIMENT:** Humans were given Money to Play Online Game

**Game Rule :** In rounds, Players can keep their money with themselves or help others by giving their money.

Groups Behaviour changed and they acted selfishly when some ROBOTs that acted as human players were introduced to behave selfishly.

-----X-----X-----

## IMPACT ON THE FINANCIAL SYSTEM

### **1. Harming Honest traders 2. Market Manipulation 3. Fraudulent Promotion**

Stock Markets are well suited to automation, as they now operate almost entirely electronically, generating huge volumes of data at high velocity,

The dynamism of markets means that timely response to information is critical, and hence slow thinking humans will not succeed.

Algorithmic trading can generate profits at a speed and frequency that is impossible for a human trader.

### Danger: Autonomous Algorithmic Trading can harm Honest Traders

- Incident of Knight Capital Group.
  - During the first 45 minutes of the trading day on 1 August 2012, while processing 212 small orders from customers, an automated trading agent developed by and operating on behalf of Knight Capital erroneously submitted millions of orders to the equity markets.
  - Over four million transactions were executed in the financial markets as a result, leading to billions of dollars in net long and short positions.
  - The company lost \$460 million on the unintended trades, and the value of its own stock fell by almost 75%.

Autonomic trading agents could also be used maliciously to destabilise markets, or otherwise harm innocent parties.

### Danger: Market manipulation

Autonomous & AI Programmed financial agents could commit financial crimes, including market manipulation,

- AI can learn the technique of order-book spoofing, which involves placing orders with no intention of ever buying them in order to manipulate honest participants in the marketplace.

### Danger: Fraudulent Promotion by Bots

Social bots have also been shown to exploit markets by artificially inflating stock through fraudulent promotion, before selling its position to unsuspecting parties at an inflated (High) price

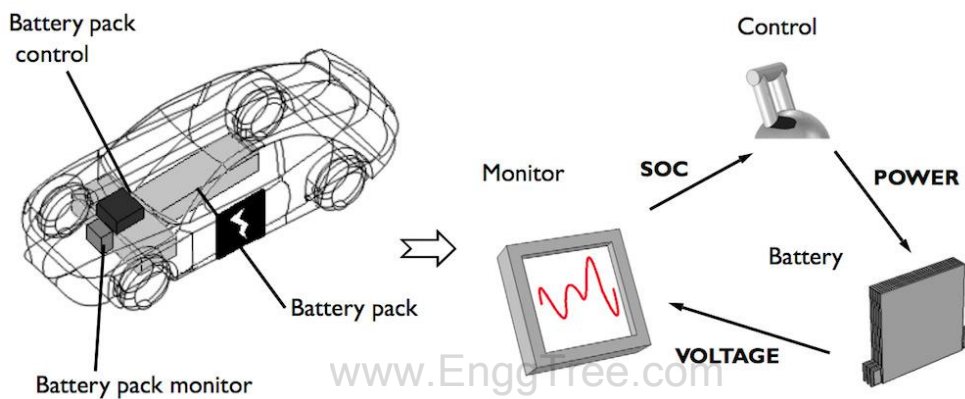
## IMPACT ON THE ENVIRONMENT AND THE PLANET

1. Use of Natural Resources
2. Pollution & Waste
3. Energy Concerns
4. Ways AI could help the Planet

AI and robotics technologies require considerable computing power, which comes with an energy cost.

### 1. Use of Natural Resources:

- The extraction of nickel, cobalt and graphite for use in **lithium ion batteries** – commonly found in electrical cars and smartphones - has already **damaged the environment**
- AI will likely increase this demand.



### 2. Pollution and waste:

- Once the lifetime of electronic goods are over, they are usually discarded, leading to a build-up of heavy metals and **toxic materials in the environment**.
- Increasing the production and consumption of **technological devices** such as robots will **increase** this waste problem,
- As the devices will likely be designed with '**inbuilt obsolescence**' – a process where products are designed to **wear out 'prematurely'** so that customers have to buy replacement items.



- Sources indicate that in North America, over **100 million cell phones and 300 million personal computers** are discarded each year.

- Way to Reduce E-Waste:
  - Encouraging consumers to prefer **ECO-FRIENDLY**, more sustainable products.



### 3. Energy concerns

- Adoption of AI technology will require **more and more data** to be processed.
- That requires **huge amounts of energy**.
- In the United States, **data centres use 2 percent** of all electricity used.

AI will also require **large amounts of energy for manufacturing and training** – for example, it would take **many hours to train a large-scale AI model** to understand and **recognise human language** such that it could be used for translation

### 4. Ways AI could help the planet

- AI could also help the planet to manage waste and pollution.

#### Examples:

- Autonomous vehicles could **reduce greenhouse gas** emissions, as autonomous vehicles could be programmed to follow the principles of **eco-driving** throughout a journey
- Autonomous vehicles could also **reduce traffic congestion** by recommending **alternative routes** and the shortest routes possible, and by sharing traffic information to other vehicles on the motorways, resulting in **less fuel consumption**.



- Studies Have shown that **Eco-Driving can reduce fuel consumption and CO2 per KM driven** than a Normal Driving style. ECO-DRIVING can be programmed in Autonomous vehicles .

- **Images of animals captured by motion-sensor cameras** in the wild, could then be used to provide accurate location, **count, and behaviour** of animals in the wild, which could be useful in enhancing local **biodiversity**.

-----X-----X-----

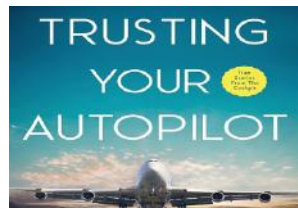
## IMPACT OF AI ON TRUST:

1. Fairness
2. Transparency
3. Algorithm Auditors
4. Accountability
5. Human in the Loop

### Why trust is important?

*In order for AI to reach its full potential, machines must sometimes work autonomously, and make decisions by themselves without human input*

- We trust machine learning algorithms
  - To diagnose diseases and identify a possible cure.
  - We trust robots to take care of our elderly , to patrol borders, and
  - To fly us around the globe.



- We trust digital technologies to simulate experiments and provide results.

This kind of trust is available widespread. It is only reassessed in the event of serious negative consequences.

<b>SURVEY: To Find Public Opinion to Accept Robots &amp; AI</b>	
People favoured the idea of using robots in areas that <u>have risk or difficulty</u> to humans EX: space exploration, manufacturing, military, security	People opposed the use of robots to care for children, the elderly, and the disabled; for healthcare

### 1. Fairness:

In order to trust AI it must be fair. As more and more decisions are delegated to AI based computer Algorithms, we must ensure that those decisions are free from bias and discrimination.

EXAMPLES of Algorithm based decision making:

- Filtering through CVs for job interviews,
- Deciding on admissions to university,
- Conducting credit ratings for loan companies,

### 2. Transparency

Important aspect of fairness is to know *why* an automated program made a particular decision. For example, a person has the right to know why they were rejected for a bank loan.

- Many AIs employ complex 'neural networks' so that even their designers cannot explain how they arrive at a particular answer

TRANSPARENCY CASE STUDY: Use of AI ALGORITHM to assess the <u>performance of teachers</u>	
• Those with high ratings won praise	• Those with low ratings fired out of job
• Some teachers felt that the AI ALGORITHM gave less Rating to them <u>without good reason</u>	
• There was no Transparency as the <u>method of rating</u> is kept as a Trade secret of the developer	
• The teachers took their <u>case to court</u> , and a judge ruled that the AI program had violated their civil rights.	

This case study highlights the importance of transparency for building trust in AI - it should always be possible to find out *why* an autonomous system made a particular decision.

### 3. Algorithm auditors

AI programs should use professional algorithm auditors, whose job would be to interrogate algorithms in order to ensure they comply with pre-set standards.

Example : Autonomous vehicle algorithm auditor

**ROLE:** Provide simulated traffic scenarios to ensure that the vehicle did not increase the risk to pedestrians or cyclists.

### 4. Accountability:

Accountability ensures that if an AI makes a mistake or harms someone, there is someone that can be held responsible,

1. The designer of AI,
2. The developer or the corporation selling the AI.

In the event of damages incurred, there must be a mechanism for redress so that victims can be sufficiently compensated.

### 5. Human in the Loop :

**Idea:** Always keep a human-in-the-loop (HITL). Here a human operator would be a crucial component of the automated control process, supervising the robots.

**EXAMPLE:** Humans are allowed to mark spams in EMAIL

**OUTCOME:** Human users in the EMAIL system help machines to identify SPAM and helps to control the entire Algorithm. when you mark an email as 'spam', you are one of

many humans in the loop of a complex machine learning algorithm, helping it in its continuous quest to improve email classification as spam or non-spam.

-----X-----X-----

## UNIT II ETHICAL INITIATIVES IN AI

### INTERNATIONAL ETHICAL INITIATIVES:

Initiative	Country	Key issues tackled								
The Institute for Ethics in Artificial Intelligence	Germany	<b>Human-centric engineering</b> covering disciplines including philosophy, ethics and political science.								
The Institute for Ethical AI & Machine Learning	United Kingdom	Based on <b>eight principles for responsible machine learning</b> : <table border="1" data-bbox="829 747 1214 785"> <tr> <td>H</td> <td>R</td> <td>B</td> <td>T</td> <td>W</td> <td>P</td> <td>T</td> <td>S</td> </tr> </table> <ol style="list-style-type: none"> <li>1. maintenance of <b>human</b> control,</li> <li>2. <b>redress</b> for AI impact,</li> <li>3. evaluation of <b>bias</b>,</li> <li>4. <b>transparency</b>,</li> <li>5. effect of AI automation on <b>workers</b>,</li> <li>6. <b>privacy</b>, 7. <b>trust</b>, and 8. <b>security</b>.</li> </ol>	H	R	B	T	W	P	T	S
H	R	B	T	W	P	T	S			
The Future of Life Institute	United States	Focus on safety : <b>autonomous weapons arms race</b> ,								
The Association for Computing Machinery	United States	The transparency, usability, security, accountability of AI in terms of <b>research, development, and implementation</b> .								
The Foundation for Responsible Robotics	The Netherlands	Responsible robotics with <b>Proactively taking actions</b> (Anticipating or Foreseeing)								
Enabling responsible AI ecosystems	Finland	Helping companies, governments, and organisations to develop and deploy <b>responsible AI ecosystems</b> ,								
euRobotics	Europe	extending <b>progress in robotics &amp; AI</b> in Europe								

-----X-----X-----

## ETHICAL HARMS AND CONCERNS

- IEEE recommends **new governance frameworks, standards**, and regulatory bodies which oversee the use of AI
- IEEE suggest **prioritising human well-being** throughout the design phase, and using the best and most widely-accepted available metrics to clearly measure the societal success of an AI.

### H E A S S S F C E I

1. Human rights and well-being : *Is AI in the best interests of humanity and human well-being?*
2. Emotional harm : *Will AI facilitate emotional or mental harm?*
3. Accountability and responsibility : *Who is responsible for AI, and who will be held accountable for its actions?*
4. Security, privacy, accessibility, and transparency : *How do we balance accessibility and transparency with privacy and security,*
5. Safety and trust : *What if AI is deemed untrustworthy by the public, or acts in ways that threaten the safety*
6. Social harm and social justice: *How do we ensure that AI is inclusive, free of bias and discrimination, and aligned with public morals and ethics?*
7. Financial harm : *How will we control for AI that negatively affects economic opportunity and employment, and either takes jobs from human workers or decreases the opportunity and quality of these jobs?*
- 8 .Control and the ethical use – or misuse – of AI : *How might AI be used unethically - and how can we protect against this? How do we ensure that AI remains under complete human control, even as it develops and 'learns'?*
- 9 .Environmental harm : *How do we protect against the potential environmental harm associated with the development and use of AI?*
10. Informed use: *What must we do to ensure that the public is aware, educated, and informed about their use of AI?*

## Human rights

AI must be created and operated to respect, promote, and protect internationally recognised human rights.



## Wellbeing

AI must increase human wellbeing to be considered a success.



## General principles for the ethical and values-based design, development and implementation of autonomous and intelligent systems

*As defined by the IEEE's Ethically Aligned Design v. 1 (March 2019)*

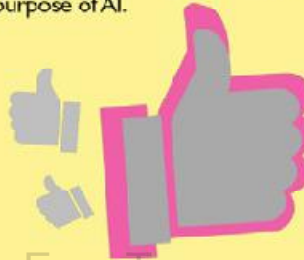
## Data agency

AI must empower individuals with the ability to access and share their data, and give people control over their identity.



## Effectiveness

AI creators and operators must provide evidence of the effectiveness and fitness for purpose of AI.



## Transparency

It should always be possible to discover the basis of AI decisions.



## Accountability

AI must provide an unambiguous rationale for all decisions made.



## Awareness of misuse

AI shall guard against all potential misuses and risks of AI in operation.



## Competence

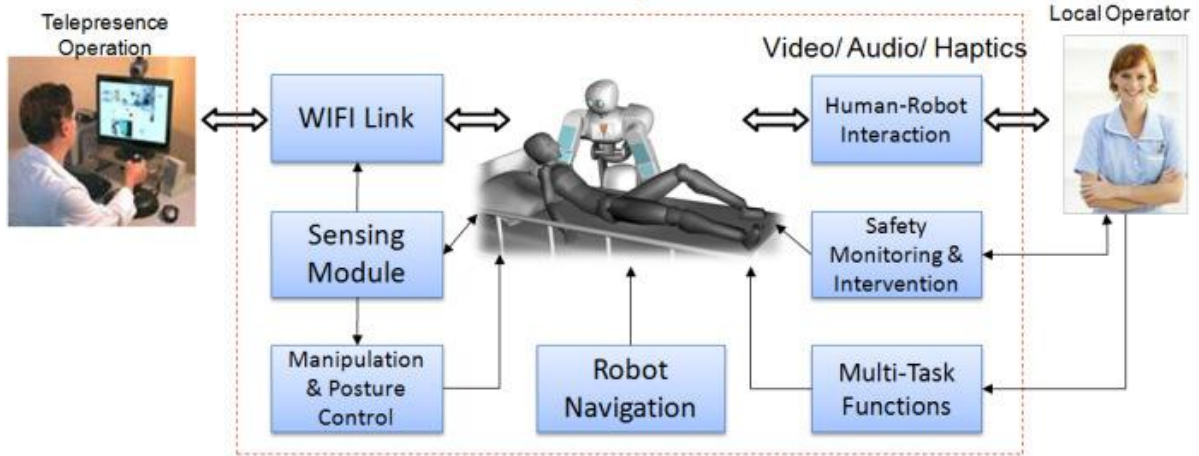
AI shall specify, and operators shall adhere to, the knowledge and skill required for safe and effective operation.



-----X-----X-----

## CASE STUDY: HEALTHCARE ROBOTS:

1. Role of robots in healthcare
2. Issues in using robots in healthcare
  1. Safety
  2. user understanding
  3. data protection
  4. legal responsibility
  5. equality of access
  6. Autonomy



Artificial Intelligence and robotics are rapidly moving into the field of healthcare and will increasingly play roles in **1. diagnosis** **2. clinical treatment.**

## ROLE OF ROBOTS IN HEALTHCARE;

1. Robots will help in the **diagnosis** of patients
2. Perform simple **surgeries**
3. **Monitoring** of patients' health and mental wellness
4. Work as companion **carers** (Attender),
5. Remind patients to take their **medicines**

## ISSUES IN USING HEALTHCARE ROBOTS :

As robots become more prevalent, the potential for future harm will increase.

### 1.SAFETY

Robots **should not harm people**, and that they should be safe to work with.

This point is especially important in areas of healthcare that deal with the **ill people, elderly, and children.**

Digital healthcare technologies offer the potential to improve accuracy of diagnosis and treatments, but to thoroughly establish a technology's long-term safety and performance **investment in clinical trials is required.**

## **2.USER UNDERSTANDING**

**Trained Healthcare Professional** in Robotic Technologies is important to ensure the effective implementation.

With **machine learning becoming embedded in diagnoses** and medical decision-making, healthcare professionals need to become **digitally literate** to understand each technological tool and use it appropriately.

### **EXAMPLE SCENARIO:**

A Machine learning algorithm **erroneously considered** a low risk asthmatic patient as high risk and took to **ICU**.

## **3.DATA PROTECTION**

Personal medical data needed for healthcare algorithms may be at risk.

### **EXAMPLE SCENARIO:**

Data : Personnal health data of persons Gathered by Fitness trackers	Danger: Data Might be sold to third parties like insurance companies
---	---

## **4.LEGAL RESPONSIBILITY**

when issues occur, legal liability must be established. If **equipment can be proven to be faulty** then the manufacturer is liable, but it is often tricky to establish what went wrong during a procedure and whether anyone, **medical personnel or machine, is to blame**

## **5. EQUALITY OF ACCESS**

Digital health algorithms and machines will improve the lifestyle .

**Ex fitness Equipments, Self Pumping Insulin etc**

But people with less digital Knowledge will not be able to use these advancements and lead to the inequality in Medical treatments.

## **6. AUTONOMY;**

Robots could be used to help elderly people live in their own homes for longer, giving them **greater freedom and autonomy.**



Question?

If a patient asked a robot to throw them off the balcony, should the robot carry out that command?

Hence the degree of autonomy for robots should be under control.

-----X-----X-----

## **CASE STUDY: WARFARE AND WEAPONISATION**

AI technology has the potential to transform warfare dangerous than the use of nuclear weapons, aircraft, computers and biotechnology.

1. Autonomous weapons
2. Drone Technologies
3. Robotic assassination
4. mobile robotics for transporting explosive devices
5. Issues of AI in warfare
6. Uses of AI in national security

### **1. Autonomous Weapons**

As automatic and autonomous systems have become more capable, militaries are trying to adopt them. Widespread adoption of AI lead to arms race.

The Russian Military Industrial Committee has already approved an aggressive plan whereby 30% of Russian combat power will consist of entirely **remote-controlled and autonomous robotic platforms by 2030.**

### **2. Drone technologies**

COST OF MILITARY AIRCRAFT	COST OF A DRONE (Unmanned Aerial Vehicle)
\$100 million	\$100 ONLY

- For the price of a single high-end aircraft, a military could acquire one million drones.

### **3. Robotic assassination**

Widespread availability of low-cost, highly-capable, autonomous robots could make targeted assassination tasks easy to implement.

Automatic sniping robots could assassinate targets in a far distance.

Similarly, self-driving cars could make suicide car bombs more frequent and devastating since they no longer require a suicidal driver.

## **4.Mobile-robotics- for transporting Explosive Devices**

Currently, the technological capability to rapidly deliver explosives to a precise target from many miles away is restricted to powerful countries. However, if long distance package delivery by drone becomes a reality, the cost of precisely delivering explosives will become easy.

### **use of machine learning in warfare;**

- Commanding Officer (CO) could employ an Intelligent Virtual Assistant (IVA)
- Uses satellite images to detect specific vehicle types

## **5. ISSUES of AI in warfare :**

- Automated weapon systems that exclude **human judgment** could violate International Humanitarian Law,
- International Humanitarian law states that any attack needs to distinguish between **combatants and civilian objects** and must not target civilians.

## **6. USES OF AI IN NATIONAL SECURITY**

### **How could AI advance national security?**

- AI will change how **we defend**.
- AI will change how **intelligence agencies** make sense of the world.
- AI will change how **we fight**.

### **EXAMPLES:**

- Dome system detects incoming rockets, **predicts their trajectory using AI**, and then sends this information to a human soldier who decides whether to launch an interceptor rocket
- Robot SGR-A1 built by Samsung, uses a low-light AI camera and pattern-recognition software to detect intruders and then issues a verbal warning. If the intruder does not surrender, the robot has a machine gun that can be fired remotely by a soldier the robot has alerted, or by the robot itself if it is in fully automatic mode.

-----X-----X-----

## UNIT III AI STANDARDS AND REGULATION 6

### Model Process for Addressing Ethical Concerns During System Design

**IEEE Std 7000:** The goal of this standard is to enable organizations to **design systems** with explicit consideration of

1. individual and societal ethical values,
  - a) transparency,
  - b) sustainability,
  - c) privacy,
  - d) fairness, and
  - e) accountability,
2. as well as values typically considered in system engineering, such as
  - a) efficiency
  - b) effectiveness.

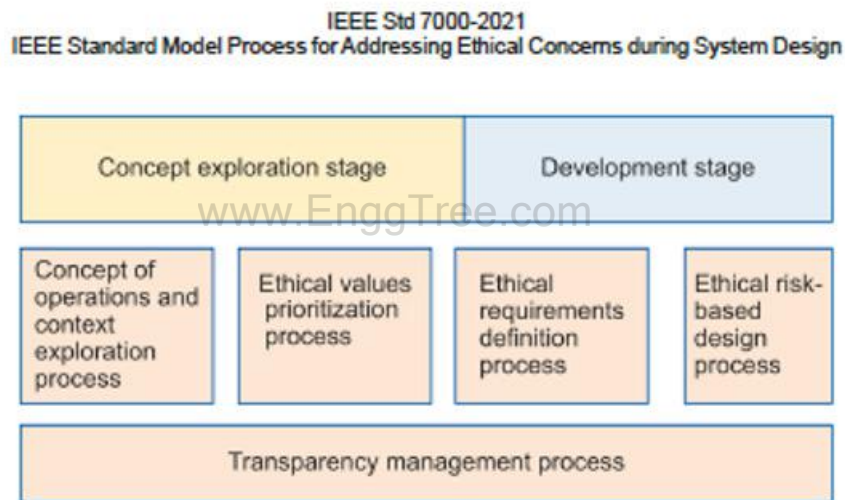


Figure 1—Relationship of processes and stages in IEEE Std 7000

### **PURPOSE:**

- **A set of processes** by which organizations can include consideration of ethical **values throughout the stages of concept exploration and development** is established by this standard.
- The **integration of Management and engineering** with transparent communication to achieve ethical values is supported by this standard
- Processes that provide **for traceability of ethical values** in the concept of operations, ethical requirements, and ethical risk-based design are described in the standard.

- **All sizes and types of organizations** using their own life cycle models are relevant to this standard.

## **SCOPE:**

IEEE Std 7000(TM) **does not give specific guidance on the design of algorithms** to apply ethical values such as fairness and privacy.

## **BENEFITS OF IEEE STD 7000:**

Projects conforming to IEEE Std 7000 balance **management commitments for time and budget constraints** with the long-term values of **social responsiveness and accountability**.

- To enable this, the **commitment of top executives** to establish and uphold organizational values is important.

-----X-----X-----

## **TRANSPARENCY OF AUTONOMOUS SYSTEMS**

### **IEEE STANDARD P 7001:**

- Transparency can be defined as the **extent to which the system discloses the processes or parameters** that relate to its functioning.
- Transparency can also be considered as the property that makes it possible to **discover how and why the system** made a particular **decision** or acted the way it did.

### **P7001 STANDARD Scope and Structure**

- The aim of P7001 is to provide a standard that sets out **“measurable, testable levels of transparency”**, so that autonomous systems can be objectively assessed
- An autonomous system is defined in P7001 as “a system that has the capacity to **make decisions itself, in response to some input data or stimulus**, with a varying degree of human intervention depending on the system’s level of autonomy”.
- **The intended users** of P7001 are
  - a) specifiers,
  - b) designers,
  - c) manufacturers,
  - d) operators and maintainers of autonomous systems.

- Furthermore P7001 is generic; it is intended to apply to all autonomous systems including robots (**autonomous vehicles, assisted living robots, drones, robot toys, etc.**), as well as **software-only AI systems, such as medical diagnosis AIs, chatbots, loan recommendation systems, facial recognition systems, etc.**

To ensure the **transparency of autonomous systems to a range of stakeholders** the IEEE P7001 standard address the following issues.

STAKEHOLDERS	ISSUES ADDRESSED
<i>Users</i>	ensuring users understand what the system does and why, with the intention of building trust;
<i>Lawyers and expert witnesses:</i>	ensuring that, following an accident, these groups are able to give evidence;
<i>Public</i>	enabling the public to assess technology ( <i>e.g. driverless cars</i> ) (and, if appropriate, build confidence).

www.EnggTree.com

## **EX : TRANSPARENCY ASSESSMENT FOR ROBOT TOY:**

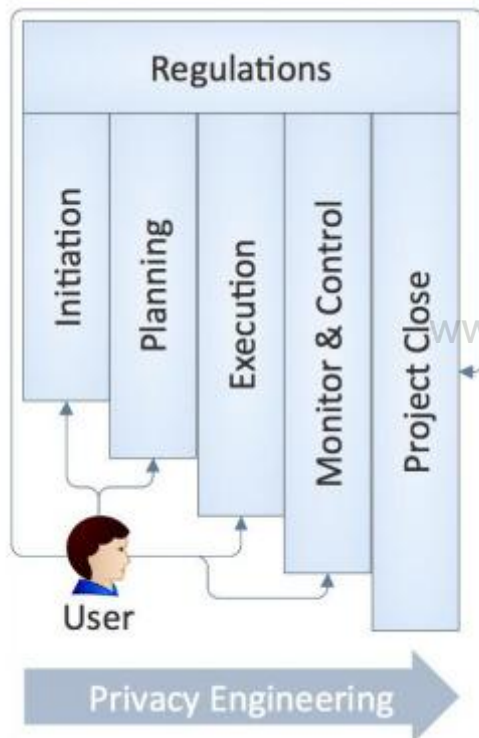
STAKEHOLDERS	ISSUES ADDRESSED
<i>Users</i>	<ul style="list-style-type: none"> <li>• <b><u>User Manual</u></b> is provided for parents</li> </ul>
<i>Lawyers and expert witnesses:</i>	<ul style="list-style-type: none"> <li>• <b><u>Risk Assessment</u></b> is performed by the manufacturer and documented</li> </ul>
<i>Public</i>	<ul style="list-style-type: none"> <li>• <b><u>Interactive Online visual guide</u></b> shows how to safeguard from Hacking and explains the risk of addiction</li> </ul>

-----X-----X-----

## Data Privacy Process- IEEE STANDARD P7002

- This standard specifies how to manage **privacy issues for systems or software** that collect personal data.
- It will do so by defining **requirements** that cover corporate data collection policies and quality assurance.
- It also includes a **use case and data model** for organizations developing applications involving personal information.
- The standard will help designers by providing ways **to identify and measure privacy controls** in their systems utilizing privacy impact assessments.

## Data Privacy Process in Projects



**1. Initiation :** Does this project implicate personal data? Are we allowed to do what we are thinking about?

**2. Planning :** How can we respect the user while using data about them?

**3. Execution:** Are there data protection checkpoints and/or gates?

**4. Monitor & Control:** Q/A tests for data protection?

**5. Project Close :** Release to production with data protection instrumentation (metrics and controls)

## ALGORITHMIC BIAS CONSIDERATIONS : The IEEE P7003 standard

Any system **that will produce different results for some people** than for others is open to challenges of being biased.

Examples could include:

- **Security camera applications** that detect theft or suspicious behaviour.
- **Marketing automation applications** that calibrate offers, prices, or content to an individual's preferences and behaviour.
- **Amazon's Recruiting Engine biased against women**  
The Amazon recruiting engine is an artificial intelligence algorithm that was created to analyze the resumes of job applicants applying to Amazon and decide which ones would be called for further interviews and selection. However, the **Amazon algorithm** turned out to be biased against women in the recruitment process. When Amazon studied the algorithm, they found that it automatically handicapped the resumes that contained words like "women" and also automatically downgraded the graduates of two all-women colleges. Therefore Amazon finally discarded the algorithm and didn't use it to evaluate candidates for recruitment.

### SCOPE OF P7003:

- The IEEE P7003 standard will provide **a framework**, which helps developers of algorithmic systems and those responsible for their deployment to identify and mitigate
  - a) unintended,
  - b) unjustified and/or
  - c) inappropriate biases in the outcomes of the algorithmic system.

Algorithmic systems in this context refers to the **combination of algorithms, data and the output deployment process** that together determine the outcomes that affect end users.

Unjustified bias refers to differential treatment of individuals based on criteria for which **no operational justification** is given.

Inappropriate bias refers to bias that is **legally or morally unacceptable** within the social context where the system is used, e.g. algorithmic systems that produce outcomes with differential impact strongly correlated with protected characteristics (such as race, gender, etc).

-----X-----X-----

## **ONTOLOGICAL STANDARD FOR ETHICALLY DRIVEN ROBOTICS AND AUTOMATION SYSTEMS : IEEE P 7007 STANDARD**

Ontologies are formal representations of the concepts, relations, and constraints in a domain of knowledge. They are widely used in artificial intelligence (AI) to provide a common vocabulary, structure, and reasoning for various tasks and applications.

The ontological specification reports provide methods to assess AI systems and organizations in their ethical performance regarding the key ethical principles of transparency, accountability, bias, and privacy.

### **EXAMPLES OF ONTOLOGIES:**

<b>ONTOLOGY CATEGORY</b>	<b>CONCEPTS</b>	<b>RELATIONSHIPS</b>
medical ontology	Include concepts such as disease, symptom, treatment, and drug.	<ul style="list-style-type: none"><li>• "a disease can have multiple symptoms"</li><li>• "a drug can be used to treat multiple diseases."</li></ul>
product ontology	include concepts such as product, category, attribute, and value	<ul style="list-style-type: none"><li>• "a product can belong to multiple categories" and</li><li>• "an attribute can have multiple values."</li></ul>

### **ETHICAL RISKS :**

1. For example, an ontology for a criminal justice AI system might include concepts such as crime, criminal, and victim. If the ontology is biased against certain groups of people, such as African Americans, then the AI system will be more likely to recommend harsher punishments for members of those groups.
2. For example, an ontology for a military AI system might include concepts such as weapon, target, and enemy. If the ontology is biased against certain groups of people, then the AI system will be more likely to identify members of those groups as enemies and recommend that they be attacked.

### **IEEE P 7007:**

This standard establishes a set of ontologies with different abstraction levels that contain concepts, definitions, axioms, and use cases that are appropriate to establish ethically driven methodologies for the design of robots and automation (R&A) systems.

### **Purpose:**



The purpose of the standard is to establish a set of definitions and their relationships to enable the development of R&A in accordance with shared values and internationally accepted ethical principles that facilitate trust in the creation and use of R&A.

## **BENEFITS:**

The use of ontologies for representing knowledge in any domain has several benefits that include the following:

- a) A formal definition of concepts of a particular domain in a language-independent representation,
- b) Tools for analyzing concepts and their relationships in searching for inconsistency, incompleteness, and redundancy

Ex: Webprotegy analysis tool

- c) Language being used in the communication process among robots from different manufacturers is standardized

## UNIT-4 ROBO ETHICS

### ROBOETHICS:

Ethics is the **branch of philosophy which studies human conduct**, the concepts of good and evil.

Roboethics --also called machine ethics-- deals with the code of conduct that robotic designer engineers must implement in the Artificial Intelligence of a robot.

Isaac Asimov developed the Three Laws of Robotics arguing that intelligent robots should be programmed in a way that they obey the following three laws:

- A robot may not injure a human being
- A robot must obey the orders given it by human beings except where such orders would conflict with the First Law
- A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

### ETHICAL CONCERNS

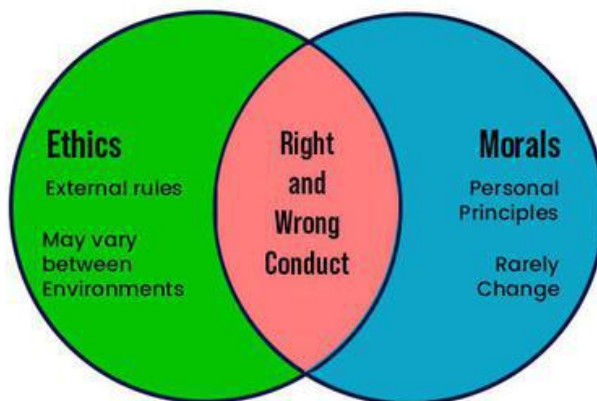
1. **Social cues:** robots should deliver a service that is as human-like as possible .. However, a customer service robot **should not hinder or replace human-to-human interactions**. It is important to guarantee this aspect when a company wants to use robots in a service delivery context.
2. **Trust and safety:** The extent to which a robot is deemed safe and trustworthy is important to the user's intention to use the technology. Companies using a service robot must always guarantee the trust and safety.
3. **Autonomy:** Even though, in our case, this variable did not have an influence on the user's intention to use a robot, the idea of being able to restrict a robot's autonomy can be found in ethical charters. Therefore, we argue that a company using a service robot should always be able to regulate a robot's autonomy, especially in cases when the consequences of the robot's actions cannot be totally controlled.

4. **Privacy and data protection:** Privacy and data protection play a big role in the intention to use a robot. First, a company using a service robot should always respect its customers' right to privacy. As transparency (i.e., disclosure about what, how and why data is collected) leads to a better user experience, companies (and their robots) have to be transparent about the collection and use of their customers' data. Secondly, companies using customer service robots should ensure that they protect their customer's data by encrypting.
5. **Human worker replacement:** A company should incorporate its employees in the choices and decisions related to the service robot, such as the choice of the robot, or the decisions related to the definition of its tasks. If a robot takes a worker's job, the firm should retrain its employee for a new occupation.

-----X-----X-----

## ETHICS AND MORALITY

ETHICS	MORALITY
<ol style="list-style-type: none"><li>1. Ethics is the <b><u>branch of philosophy</u></b> concerned with the evaluation of human conduct.</li><li>2. Ethics are <b><u>universal</u></b></li><li>3. Ethics applies to groups and organizations</li></ol>	<ol style="list-style-type: none"><li>1. morality is the right or wrong of an action, <b><u>a way of life</u></b></li><li>2. Morality is often <b><u>culture-specific</u></b>.</li><li>3. morality applies to individuals</li></ol>



-----X-----X-----

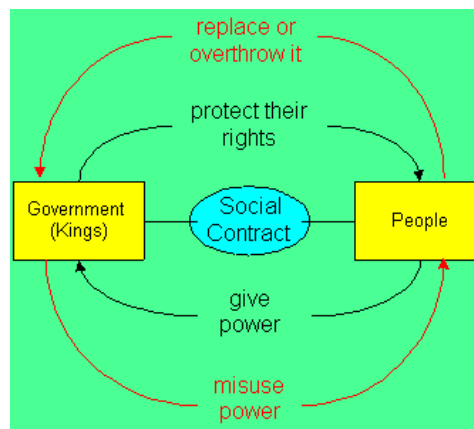
## MORAL THEORIES

1. Utilitarianism is a theory of morality that advocates actions that foster happiness and oppose actions that cause unhappiness. Utilitarianism promotes "the greatest amount of good for the greatest number of people."



2. *Contractualism*: is the theory based on the mutual contract between the designer and the contract provider.

An action is morally wrong if it is contrary to the general system of moral rules upon which there could be **informed and unforced agreement**.



3. *Deontology*;

Deontological ethics is the ethical theory that the morality of an action should be based on whether that action **itself is right or wrong** under a series of rules and principles, rather than based on the consequences of the action.

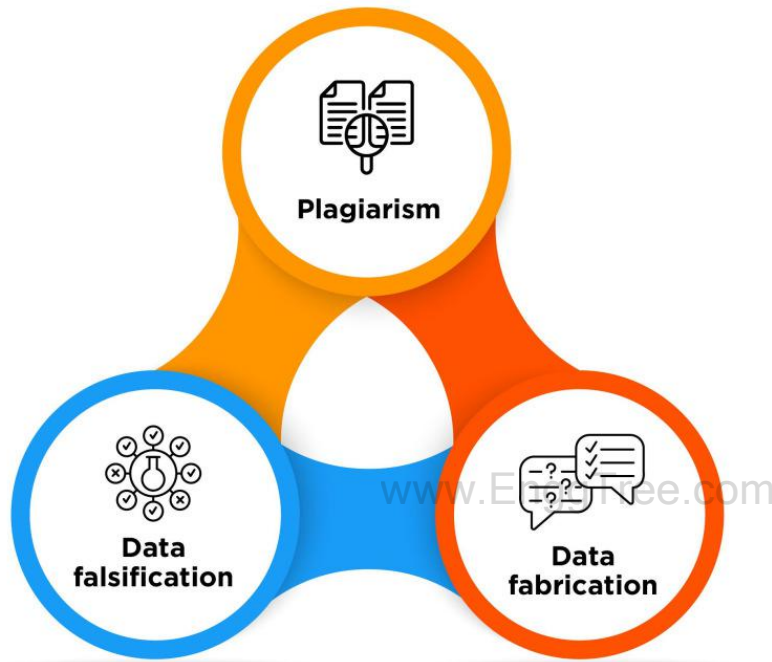
Ex: 1. Do not kill 2. do not steal 3. do not lie 4. do not cheat etc

-----X-----X-----

## ETHICS IN SCIENCE AND TECHNOLOGY

**Science ethics:** In science, ethical principles, especially **honesty and integrity**, should guide all stages of scientific practice—**including data collection, peer review, publishing, and replication of findings**—to assure that scientific knowledge is unbiased and trustworthy.

### BAD PRACTICES:



#### 1. Scientific honesty

Scientists should not commit scientific fraud by, for example, **destroying, or misrepresenting data**.

#### 2. Carefulness

Scientists should avoid **careless errors** in all aspects of scientific work.

#### 3. Intellectual freedom

Scientists should be free to pursue **new ideas and criticize old ones** and conduct research on anything they find interesting.

#### 4. Openness

Whenever possible, scientists should **share data, results**, methods, theories, equipment, and so on; allow people to see their work; and be open to criticism.

## **5. Attribution of credit**

Scientists **should not plagiarise** the work of other scientists. **They should give credit where credit is due but not where it is not due.**

## **6. Public responsibility**

Scientists **should report research** in the public media when the research has an important and direct bearing on human happiness and when the research has been sufficiently validated by scientific peers.

**Technology ethics:** Ethics in Technology governs principles of right and wrong while using technological advancements . Examples:

1. The use of **surveillance technologies**.
2. The use of AI to make a **medical diagnosis**.
3. The use of robotic systems in **surgery**.
4. Ethics of **self-driving cars** and their decisions.
5. Ethical issues related to **cloning**.
6. Ethics related to implantable medical devices such as **pacemakers**.
7. The ethics of **hacking** for security testing purposes.
8. Use of technology such as **cryptocurrency**.
9. The **environmental impact** of technology.
10. **Internet of things** and the embedding of networked devices into everyday things.
11. The impact of **social media** on society.
12. The use of **AI in education**.
13. **Screen time** and its impact on children.

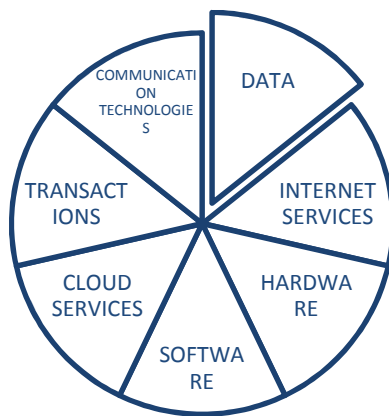
-----X-----X-----

## Ethical Issues in an ICT Society

Information and communication technologies (ICT) is defined as a diverse set of technological tools and resources used **to transmit, store, create, share or exchange information.**

Information Technology have a wide area of applications in **education, business, health, industries, banking sector .**

COMPONENTS OF ICT:



To deal with ICT society it is important to find out the ethical issues.

Some of the major ethical issues faced by Information & communication Technology (ICT) are:

1. Personal Privacy
2. Access Right
3. Harmful Actions
4. Copyright
5. Liability
6. Piracy

### 1. Personal Privacy

Due to the distribution of the network on a large scale, **data or information transfer** in a big amount takes place which leads to the **hidden chances of disclosing information** and violating the privacy of any individuals or a group.

### 2. Access Right

Network on the internet cannot be made secure from **unauthorized access**. Generally, the intrusion detection system are used to determine whether the **user is an intruder or an appropriate user.**

### **3. Harmful Actions**

Harmful actions in the computer ethics refers to the **damage to the IT** such as loss of important information, loss of property, loss of ownership. To recover from the harmful actions extra time and efforts are required to **remove the viruses** from the computer systems.

### **4. Copyright**

Copyright law works as a very powerful legal tool in protecting computer software, both before a **security breach and after a security breach**.

### **5. Liability**

Software developer should be aware of the liability issue in making ethical decisions. Software developer makes **promises about the nature and quality** of the product that is given as **warranty**.

### **6. Piracy**

Piracy is the creation **of illegal copy** of the software . The software industry is prepared to do encounter against software piracy. The courts are dealing with an increasing number of actions concerning the protection of software.

-----X-----X-----

### **Harmonization of Principles**

Harmonization is the act of **making systems or laws the same or similar** in different companies, countries, etc. so that they can work together more easily.

Internationally recognized institutions such as the United Nations, the World Health Organization (WHO), have identified general ethical principles that have been adopted by most nations, cultures, and people of the world.

Furthermore, the international scientific, community has proposed a harmonization of world ethical principles applied to science and technology, especially in those cases when these principles involve sensitive issues such as life, human dignity, and freedom.

Three kinds of integration are analyzed:



1. Human–softbot integration, as achieved by AI research on information and communication technologies. A softbot (software robot) is a program that interacts with a software environment by issuing commands and interpreting the environment's feedback.
2. Human–robot noninvasive integration, as achieved by robotic research on autonomous systems inhabiting human environments
3. Physical, invasive integration, as achieved by bionic research

-----X-----X-----

## Ethics and Professional Responsibility

In the design, development, and application of a new technology, designers, manufacturers, and end users should be following rules, which are common to all human beings:

- human dignity and human rights
- equality, justice
- benefit and harm
- respect for cultural diversity
- nondiscrimination
- autonomy and individual responsibility
- informed consent
- privacy and confidentiality
- cooperation
- sharing of benefits
- responsibility towards the biosphere
- cost-benefit analysis

Computer and information ethics has developed a codes of ethics called *PAPA*

**Privacy:** What information about ones self or ones associations must a person reveal to others

**Accuracy:** Who is responsible for the authenticity

**Property:**Who owns information? What are the just and fair prices for its exchange?

**Accessibility:** What information does a person or an organization have a right

Professionals are advised to apply, in performing sensitive technologies, the *precautionary principle*:

*When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some **cause-and-effect relationships** are not fully established **scientifically**.*

All over the world, associations and orders of engineers have adopted codes of ethics guiding towards responsible conduct in research and practice.

**security and reliability** are the most important ethical codes of conduct.

-----X-----X-----

## **ROBOETHICS TAXONOMY**

Roboethics Taxonomy refers to the **classification** of types of Robots available and the ethics corresponding to their usage.

**1. Humanoids:** A humanoid robot is a robot **resembling the human body** in shape.

Artificial intelligence will be able to lead the robot to fulfill the missions required by the end users.

To achieve this goal, over the past decades scientists have worked on AI techniques in many fields, including:

1. Artificial vision
2. Perception of the environment
3. Natural language processing
4. Human interaction
5. Machine learning
6. Neural networks

**Humanoid Ethics:** one of the fundamental aspects of the Humanoid robots is their capability to **learn the characteristics** of the surrounding environment.

Humanoid intelligence is a **learned intelligence**, fed by the world models uploaded by the designers.

The increasing autonomy of the robots could give rise **to unpredictable** and non-predictable behaviors.

Hence the designers must **predict** the threats and foresee the dangers.

## **2. Industrial Robotics**

An industrial robot is officially defined by **ISO** as an automatically controlled, reprogrammable, multipurpose manipulator.

Typical applications of industrial robots include welding, painting, assembly, pick and place etc

**Ethical issues in Industries:** Social problems stemming from the introduction of robots in factories are the loss of jobs and unemployment.

Ethically, the loss of job by the introduction of Industrial robots should be compensated by giving an alternate job like 1. Maintenance of the Robot 2. A new skill etc

### **3. Adaptive Robot Servants**

Adaptive Robo servants are Service robots that support and back up human operators. Robots come in several shapes and sizes (wheeled, legged, humanoids), equipped with different kinds of sensing systems (artificial vision systems, ultrasonic, radio) and manipulations (grippers, hands, tools, probes).

#### **Ethical Issues for ROBOT servants:**

- Designers must guarantee safety and Security
- Overuse could lead to technology addiction or invasion of privacy

### **4. Distributed Robotic Systems**

Distributed Robotic systems are **Networked Robotic systems that are linked to the web.**

A full-scale robot team would be of tremendous value in a number of applications such as security, surveillance, monitoring, gardening, and pharmaceutical manufacturing.

#### **Ethical Issues in Distributed Robotic systems:**

- unpredictability of behavior due to the availability of a team of several systems.
- Assignment of Liability for Misbehavior
- Vulnerability to hacking

### **5. Outdoor Robotics**

Outdoor robots are intelligent machines that explore, develop, secure our world.

- Rescue robotics (robots that support first-response units in disaster missions)
- Agricultural (autonomous tractors, planters and harvesters, applicators for fertilizers and pest control)

#### **Issues in using Outdoor Robots :**

Use of outdoor robots could lead to excessive exploitation of the planet, which can become a threat to biodiversity and life on the planet.

### **6. Surgical Robotics**

Surgical Robots are robots that allow doctors to perform many types of complex surgeries with more precision, flexibility and control than is possible with conventional techniques.

Typical applications are:

- Robotic telesurgical workstations
- Robotic systems for diagnosis ( Ct Scan – Computerized Tomography Scan)

## Issues:

- High cost of robotic systems in the medical field could widen the **digital divide** between developed and developing countries

-----X-----X-----

## UNIT-5

### OPPORTUNITIES OF AI:

#### 1. Artificial Intelligence in E-Commerce

- **Artificial Intelligence** is widely used in the field of **E-commerce**
- AI helps the organization to establish a good engagement between the user and the company.
- Artificial Intelligence helps to make appropriate suggestions and recommendations as per the user search history . There are also **AI chatbots** that are used to provide customer support instantly.
  - a) **Personalization:** AI helps customers to see those products based on their interest pattern . [www.EnggTree.com](http://www.EnggTree.com)
  - b) **Dynamic Pricing Structure:** It's a smart way of fixing price of the product based on the demand data.
  - c) **Fake Review Detection:** AI algorithms are used to detect and delete Fake Reviews.

#### 2. AI in Education Purpose

**AI** helps the faculty as well as the students by making course recommendations.

**Making automated messages** to the students, and parents regarding any vacation, and test results are done by Artificial Intelligence.

#### AI applications in Education.

- a) **Voice Assistant:** With the help of AI algorithms, this feature can be used in multiple ways to save time.

- b) **Gamification:** This feature has enabled e-learning companies to design attractive game models, so that kids can learn in a fun way. This will also ensure that they are catching the concepts.

### 3. Artificial Intelligence in Robotics

Artificial Intelligence is one of the major technologies that provide the [robotics](#) field with a boost to increase their efficiency. AI provides robots to make decisions in real time and increase productivity. For example, suppose there is a warehouse in which robots are used to manage good packages. The robots are only designed to deliver the task but Artificial Intelligence makes them able to analyze the vacant space and make the best decision in real-time. Let's take a closer look at AI applications in Robotics.

- **NLP:** Natural Language Processing plays a vital role in robotics to interpret the command as a human being instructs. This enables AI algorithms & techniques such as sentimental analysis, syntactic parsing, etc.
- **Object Recognition & Manipulation:** This functionality enables robots to detect objects within the perimeter and this technique also helps robots to understand the size & shape of that particular object. Besides this, this technique has two units, one is to identify the object & the other one refers to the physical interaction with the object.
- **HRI:** With the help of AI algorithms, HRI or Human-Robotics Interaction is being developed that helps in understanding human patterns such as gestures, expressions, etc. This technique helps maximize the performance of robots and ensures that it reaches and maintains its accuracy.

### 4. GPS and Navigations

[GPS](#) technology uses Artificial Intelligence to make the best route and provide the best available route to the users for traveling. This is also suggested by research provided by the **MIT Institute** that AI is able to provide accurate, timely, and real-time information about any specific location. It helps the user to choose their type of lane and roads which increases the safety features of a user. GPS and **navigation** use the [convolutional](#) and [graph neural network](#) of Artificial Intelligence to provide these suggestions. Let's take a closer look at AI applications in GPS & Navigation.

- **Voice Assistance:** This feature allows users to interact with the AI using a hands-free feature & which allows them to drive seamlessly while communicating through the navigation system.

- **Personalization (Intelligent Routing):** The personalized system gets active based on the user's pattern & behavior of preferred routes. Irrespective of the time & duration, the GPS will always provide suggestions based on multiple patterns & analyses.
- **Traffic Prediction:** AI uses a Linear Regression algorithm that helps in preparing and analyzing the traffic data. This clearly helps an individual in saving time and alternate routes are provided based on congestion ahead of the user.
- **Positioning & Planning:** GPS & Navigation requires enhance support of AI for better positioning & planning to avoid unwanted traffic zones. To help with this, AI-based techniques are being used such as Kalman, Sensor fusion, etc. Besides this, AI also uses prediction methods to analyze the fastest & efficient route to surface the real-time data.

## 5. Healthcare

Artificial Intelligence is widely used in the field of **healthcare and medicine**. The various algorithms of Artificial Intelligence are used to build precise machines that are able to detect minor diseases inside the human body. Also, Artificial Intelligence uses the medical history and current situation of a particular human being to predict future diseases. Artificial Intelligence is also used to find the current vacant beds in the hospitals of a city that saves the time of patients who are in emergency conditions. Let's take a closer look at AI applications in Healthcare.

- **Insights & Analysis:** With the help of AI, a collection of large datasets, that includes clinical data, research studies, and public health data, to identify trends and patterns. This inversely provides aid in surveillance and public health planning.
- **Telehealth:** This feature enables doctors and healthcare experts to take close monitoring while analyzing data to prevent any uncertain health issues. Patients who are at high risk and require intensive care are likely to get benefitted from this AI-powered feature.
- **Patient Monitoring:** In case of any abnormal activity and alarming alerts during the care of patients, an AI system is being used for early intervention. Besides this, RPM, or Remote Patient Monitoring has been significantly growing & is expected to go up by USD 6 Billion by 2025, to treat and monitor patients.

- **Surgical Assistance:** To ensure a streamlined procedure guided by the AI algorithms, it helps surgeons to take effective decisions based on the provided insights to make sure that no further risks are involved in this while processing.

## 6. Automobiles

Artificial Intelligence is bringing revolutionary changes in the field of [automobiles](#). From speedometers to self-driving cars, Artificial Intelligence is really doing a significant difference in these sectors. AI is used to detect the traffic on the street and provide the best route out of the present all routes to the driver. It uses **sensors**, **GPS technology**, and **control signals** to bring the vehicle the best path. Let's take a closer look at AI applications in Automobiles.

- **ADAS:** Advanced Driving Assistance System or ADAS is an AI algorithm that is known for processing some of the major and sensitive data that includes driving assistance, crash detection, parking assistance, and so on. The algorithm has been designed in such a way that it automatically starts alarming to prevent any collisions.
- **Traffic Management:** With the help of AI systems, now it has become easier to analyze traffic data from various sources, including vehicles, sensors, and cameras. This feature helps in boosting driving assistance by offering alternate routes. This AI algorithm has been designed in such a way that it is well capable of assisting users with real-time traffic insight and by offering the fastest route as an alternative to users to save time efficiently.
- **Emission Reduction:** This feature detects and learns patterns from the given inputs i.e. from the driving pattern of the user and based on this it strategizes to perform efficient driving patterns by reducing emissions. This algorithm is well capable of analyzing routes, traffic, car performance patterns, and so on.
- **Autonomous Driving:** This AI approach enables automatic driving and navigating vehicles without actual human intervention. Sensors like LIDAR, RADAR, and other sensors help in collecting additional data to analyze the surroundings for taking optimal decisions in the real world.

## 7. Agriculture

Artificial Intelligence is also becoming a part of **agriculture** and farmers' life. It is used to detect various parameters such as the **amount of water and moisture**, **amount of deficient nutrients**, etc in the soil. There is also a machine that uses AI to detect where

the weeds are growing, where the soil is infertile, etc. Let's take a closer look at AI applications in Agriculture.

- **Stock Monitoring:** To have rigorous monitoring, and ensure that crops that not being affected by any disease, AI uses CN to check crop feeds live and alarms when any abnormality arises.
- **Supply Chain:** The AI algorithm helps in analyzing and preparing the inventory to maintain the supply chain stock. Although it's not new, for the agriculture field, it does help farmers to ensure the demands are being met with minimal loss.
- **Pest Management:** AI algorithms can analyze data from multiple sources to identify early warnings to their respective farmers. This technology also enables less usage of harmful pesticides by offering the best resources for pest management.
- **Forecasting:** With the help of AI, analyzing the weather forecast and crop growth has become more convenient in the field of agriculture and the algorithms help farmers to grow crops with effective business decisions.

-----X-----X-----

www.EnggTree.com

## **CHALLENGES OF AI:**

Artificial intelligence is **essential** across a wide range of industries, **including healthcare, retail, manufacturing, and even government.**

But there are challenges with AI, it is necessary to **be vigilant (Cautious)** about these issues to make sure that artificial intelligence is **not doing more harm than good.**

### **7 biggest challenges of artificial intelligence:**

#### **1. Biases**

We need data to train our artificial intelligence algorithms, and we need to do everything we can to eliminate bias in that data.

The ImageNet database, for example, has far more white faces than non-white faces. When we train our AI algorithms to recognize facial features using a database that doesn't include the right balance of faces, the algorithm won't work as well on non-white faces, creating a built-in bias that can have a huge impact.

I believe it's important that we eliminate as much bias as possible as we train our AI, instead of shrugging our shoulders and assuming that we're training our AI to accurately



reflect our society. That work begins with being aware of the potential for bias in our AI solutions.

## **Control and the Morality of AI**

As we use more and more artificial intelligence, we are asking machines to make increasingly important decisions.

For example, right now, there is an international convention that dictates the use of autonomous drones. If you have a drone that could potentially fire a rocket and kill someone, there needs to be a human in the decision-making process before the missile gets deployed. So far, we have gotten around some of the critical control problems of AI with a patchwork of rules and regulations like this.

The problem is that AIs increasingly have to make split-second decisions. For example, in high-frequency trading, over 90% of all financial trades are now driven by algorithms, so there is no chance to put a human being in control of the decisions.

The same is true for autonomous cars. They need to react immediately if a child runs out on the road, so it's important that the AI is in control of the situation. This creates interesting ethical challenges around AI and control.

## **Privacy**

[www.EnggTree.com](http://www.EnggTree.com)

Privacy (and consent) for using data has long been an ethical dilemma of AI. We need data to train AIs, but where does this data come from, and how do we use it? Sometimes we make the assumption that all the data is coming from adults with full mental capabilities that can make choices for themselves about the use of their data, but we don't always have this.

For example, Barbie now has an AI-enabled doll that children can speak to. What does this mean in terms of ethics? There is an algorithm that is collecting data from your child's conversations with this toy. Where is this data going, and how is it being used?

As we have seen a lot in the news recently, there are also many companies that collect data and sell it to other companies. What are the rules around this kind of data collection, and what legislation might need to be put in place to protect users' private information?

## **Power Balance**

Huge companies like Amazon, Facebook, Google, are using artificial intelligence to squash their competitors and become virtually unstoppable in the marketplace. Countries like China also have ambitious AI strategies that are supported by the government.

President Putin of Russia has said, “Whoever wins the race in AI will probably become the ruler of the world.”

How do we make sure the monopolies we’re generating are distributing wealth equally and that we don’t have a few countries that race ahead of the rest of the world? Balancing that power is a serious challenge in the world of AI.

## **Ownership**

Who is responsible for some of the things that AIs are creating?

We can now use artificial intelligence to create text, bots, or even deepfake videos that can be misleading. Who owns that material, and what do we do with this kind of fake news if it spreads across the internet?

We also have AIs that can create art and music. When an AI writes a new piece of music, who owns it? Who has the intellectual property rights for it, and should potentially get paid for it?

## **Environmental Impact**

Sometimes we don’t think about the environmental impact of AI. We assume that we are using data on a cloud computer to train an algorithm, and then that data is used to run recommendation engines on our website. However, the computer centers that run our cloud infrastructure are power-hungry.

Training in AI, for example, can create 17 times more carbon emissions than the average American does in about a year.

How can we use this energy for the highest good and use AI to solve some of the world’s biggest and most pressing problems? If we are only using artificial intelligence because we can, we might have to reconsider our choices.

## **Humanity**

My final challenge is “How does AI make us feel as humans?” Artificial intelligence has now gotten so fast, powerful, and efficient that it can leave humans feeling inferior. This issue may challenge us to think about what it actually means to be human.

AI will also continue to automate more of our jobs. What will our contribution be, as human beings? I don’t think artificial intelligence will ever replace all our jobs, but AI will augment them. We need to get better at working alongside smart machines so we can manage the transition with dignity and respect for people and technology.

## **Societal Issues Concerning the Application of Artificial Intelligence in Medicine**

Medicine is becoming an increasingly data centred discipline and, beyond classical statistical approaches, artificial intelligence (AI) is used for the analysis of medical data.

AI applications in healthcare have literally changed the medical field, including imaging and electronic medical records (EMR), laboratory diagnosis, treatment, augmenting the intelligence of the physicians, new drug discovery, providing preventive and precision medicine.

- In healthcare, current laws are not enough to protect an individual's health data.
- Clinical data collected by robots can be hacked into and used for malicious purposes that minimize privacy and security.
- All individuals have the right to get information and ask questions before procedures and treatments.
- Patients should be able to be aware of the treatment process, the risks of screening and imaging, data capture anomalies, programming errors
- Patients will lose empathy, kindness, and appropriate behavior when dealing with robotic physicians and nurses because these robots do not possess human attributes such as compassion.
- In Gynecology, any clinical examination requires a sense of compassion and empathy, which will not be achieved with robotic doctors.
- Children usually experience fear or anxiety as they engage in healthcare settings . Their behavior could be uncontrollable with the new robotic medicine system.

## **ISSUES IN USING HEALTHCARE ROBOTS :**

As robots become more prevalent, the potential for future harm will increase.

### **1.SAFETY**

Robots **should not harm people**, and that they should be safe to work with.

This point is especially important in areas of healthcare that deal with the **ill people, elderly, and children.**

Digital healthcare technologies offer the potential to improve accuracy of diagnosis and treatments, but to thoroughly establish a technology's long-term safety and performance **investment in clinical trials is required.**

## **2.USER UNDERSTANDING**

**Trained Healthcare Professional** in Robotic Technologies is important to ensure the effective implementation.

With **machine learning becoming embedded in diagnoses** and medical decision-making, healthcare professionals need to become **digitally literate** to understand each technological tool and use it appropriately.

### **EXAMPLE SCENARIO:**

A Machine learning algorithm **erroneously considered** a low risk asthmatic patient as high risk and took to **ICU**.

## **3.DATA PROTECTION**

Personal medical data needed for healthcare algorithms may be at risk.

### **EXAMPLE SCENARIO:**

Data : Personnal health data of persons Gathered by Fitness trackers	Danger: Data Might be sold to third parties like insurance companies
---	---

## **4.LEGAL RESPONSIBILITY**

when issues occur, legal liability must be established. If **equipment can be proven to be faulty** then the manufacturer is liable, but it is often tricky to establish what went wrong during a procedure and whether anyone, **medical personnel or machine, is to blame**

## **5. EQUALITY OF ACCESS**

Digital health algorithms and machines will improve the lifestyle .

**Ex fitness Equipments, Self Pumping Insulin etc**

But people with less digital Knowledge will not be able to use these advancements and lead to the inequality in Medical treatments.

## **6. AUTONOMY;**

Robots could be used to help elderly people live in their own homes for longer, giving them **greater freedom and autonomy.**

Question?

If a patient asked a robot to throw them off the balcony, should the robot carry out that command?

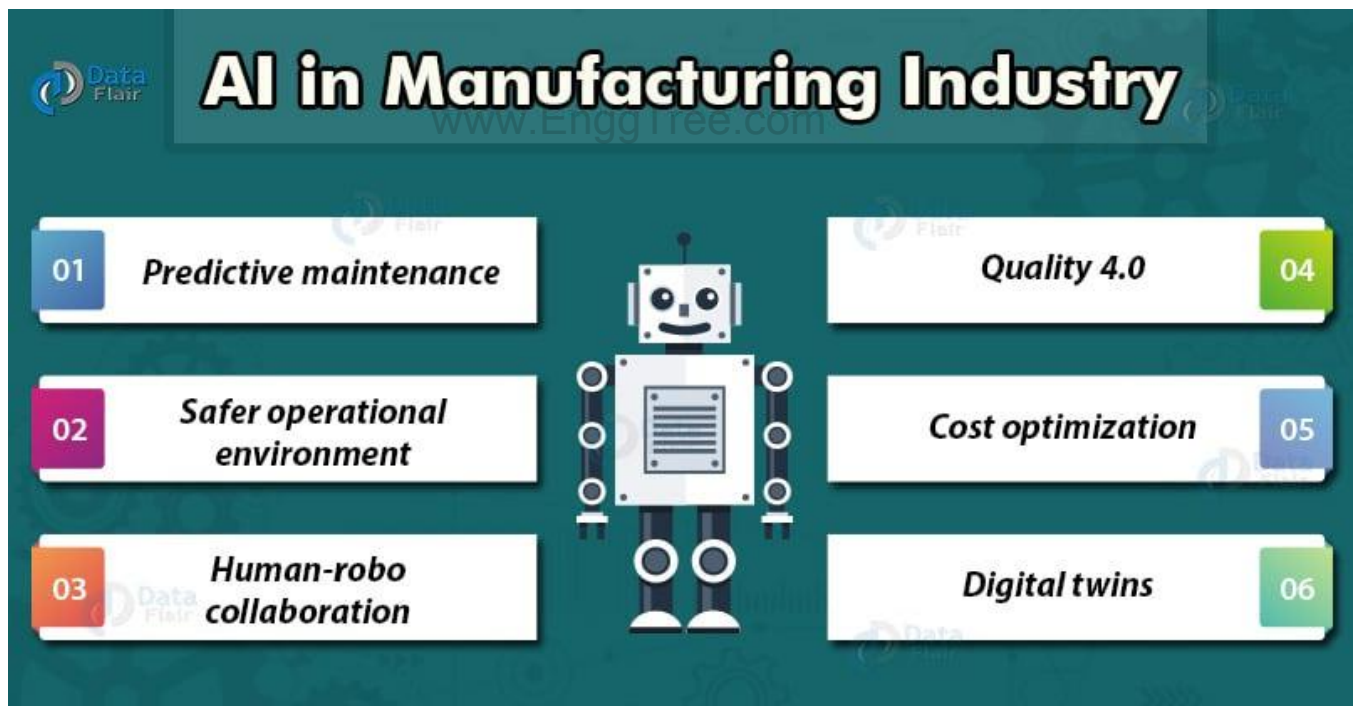
Hence the degree of autonomy for robots should be under control.

-----X-----X-----X-----

## IMPACT OF AI IN DECISION MAKING ROLE IN INDUSTRIES:

Manufacturing in the near future would be fully automated. The manufacturing processes enabled by Artificially Intelligent Systems would be able to perform the required processes. It will also be able to inspect, improve, and quality checks the products without any human intervention.

According to the reports of Markets and markets, Artificial Intelligence in the manufacturing market is expected to grow from USD 1.0 billion in 2018 to USD 17.2 billion by 2025, at a CAGR of 49.5% during the forecast period.



### 1. Predictive Maintenance

Predictive maintenance techniques are designed to determine the condition of machines to estimate when maintenance should be performed. **This approach promises cost savings over routine or time-based preventive maintenance**, because maintenance tasks are performed only when required.

For example, if a machine is under continuous load. With the help of an infrared imager, we can monitor the temperature profile of equipment without affecting the overall performance.

## **2. Safer Operational Environment**

With even a **minor mistake** occurring on the assembly line proving hazardous, a stage towards **AI implies less human assistance** needed to complete unsafe work.

As robots support people and perform unsafe exercises, the number of working environment mistakes will diminish . As a result, this will lead to safer working conditions than before.

For example, instead of a human doing a crash test of a car, an AI would be a natural option.

## **3. Human-Robo Collaboration**

There are a large number of robots working in manufacturing plants everywhere throughout the world. People are concerned that their **occupations may be replaced** by robots.

Individuals can be recruited for more elevated level situations for **programming** and the **executives of the business** forms.

## **4. Quality 4.0**

Industry 4.0 methods, can deliver good quality items by utilizing AI calculations. If any issue is found at the **starting stages**, we can deal with it right away.

## **5. Cost Optimization**

Bringing AI into the production lines would require an **enormous capital venture**, but the ROI (RETURN ON INVESTMENT) is high. As smart machines begin dealing with everyday exercises, organizations can get lower working costs.

## **6. Digital Twins**

[AI and Digital Twins](#) make a **virtual portrayal** that reproduces the physical attributes of the plant, items, or machine segments.

By utilizing cameras, sensors, computerized twin can make **a live model** of the production line plant that helps to precisely foresee wear, development, and collaborations with different gadgets.

-----X-----X-----X-----

## NATIONAL AND INTERNATIONAL STRATEGIES ON AI

### NATIONAL STRATEGIY OF AI:

The National level strategy of India is

#### 1. #AiforAll.

- #AiforAll will aim at enhancing **human capabilities** to address the challenges of
  - access,
  - affordability,
  - shortage and inconsistency of skilled expertise;
- effective implementation of AI **to develop scalable solutions** for emerging economies;
- **To tackle some of the global challenges** from AI's perspective, like application, research, development, technology, or responsible AI.
- #AiforAll will focus on **harnessing collaborations** to ensure prosperity for all.
- Thus, #AiforAll means technology leadership in AI for achieving the **greater good**.

#### 2. AI Garage for 40% of the world

- **Solve for India** means solve for 40% or more of the world.

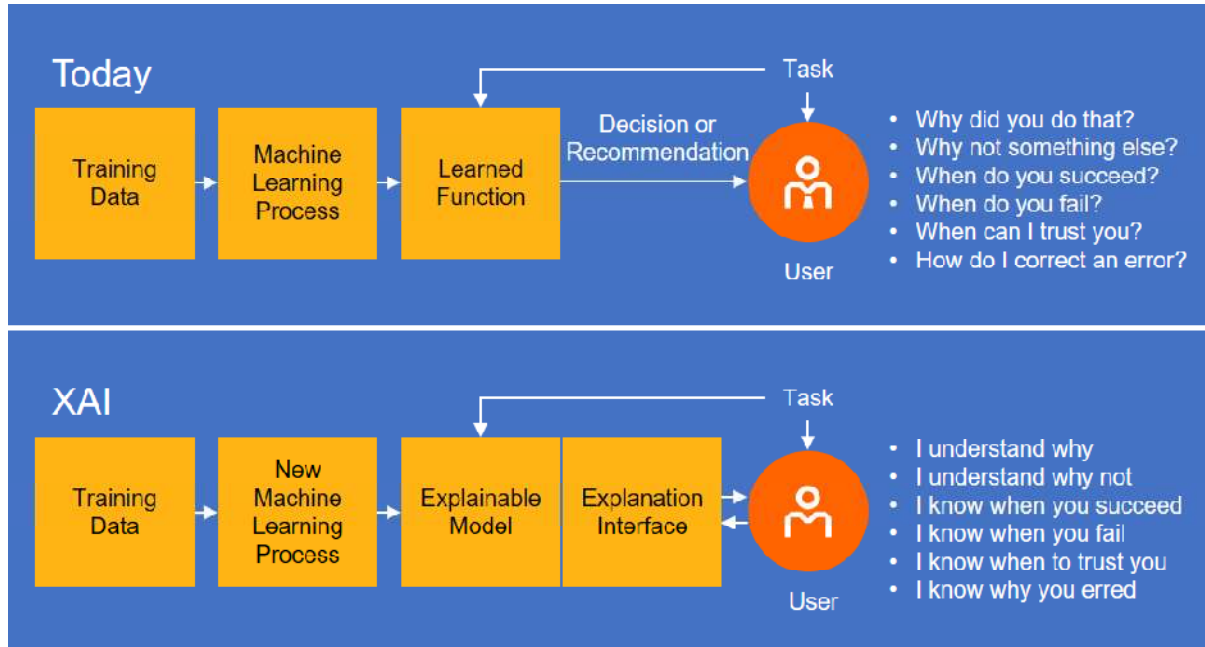
An advanced AI based solution for early **diagnosis of tuberculosis** (one of the top-10 causes of deaths worldwide), for example, could easily be **rolled out to countries** in South East Asia or Africa, once developed and refined in India.

- **Solved in India** (or more accurately, **solved by Indian IT companies**) could be the model going forward for Artificial Intelligence as a Service (AIaaS). Indian IT companies have been pioneers in bringing technology products and developments as solutions across the globe.

#### 3. Decoding Explainable AI

The Explainable AI (XAI) program aims to create a suite of machine learning techniques that:

- Produce **more explainable models**
- The machine learning algorithms of tomorrow should have the **built-in capability** to explain their logic



## INTERNATIONAL STRATEGIES:

Initiative	Country	Key issues tackled
The Institute for Ethics in Artificial Intelligence	Germany	<b>Human-centric engineering</b> covering disciplines including philosophy, ethics and political science.
The Institute for Ethical AI & Machine Learning	United Kingdom	Based on <b>eight principles for responsible machine learning:</b> <div style="border: 1px solid black; display: inline-block; padding: 2px;">H R B T W P T S</div> 1. maintenance of <b>human</b> control, 2. <b>redress</b> for AI impact, 3. evaluation of <b>bias</b> , 4. <b>transparency</b> , 5. effect of AI automation on <b>workers</b> , 6. <b>privacy</b> , <b>7. trust</b> , and <b>8. security</b> .
The Future of Life Institute	United States	Focus on safety : <b>autonomous weapons arms race</b> ,
The Association for Computing Machinery	United States	The transparency, usability, security, accountability of AI in terms of <b>research, development, and implementation</b> .



The Foundation for Responsible Robotics	The Netherlands	Responsible robotics with <b><u>Proactively taking actions</u></b> (Anticipating or Foreseeing)
Enabling responsible AI ecosystems	Finland	Helping companies, governments, and organisations to develop and deploy <b><u>responsible AI ecosystems</u></b> ,
euRobotics	Europe	extending <b><u>progress in robotics &amp; AI</u></b> in Europe

-----X-----X-----X-----

**ALL THE BEST**

**WE CREATE SUCCESS**

[www.EnggTree.com](http://www.EnggTree.com)

