HCAI NOTES

MODULE-1
WHAT IS HUMAN-CENTERED ARTIFICIAL INTELLIGENCE

What is human-centered Artificial intelligence?

- Human-centered AI is an approach to artificial intelligence that prioritizes the **needs, values, and well-being of humans**. It focuses on designing AI systems that are **transparent, ethical, and aligned with human goals,** ensuring technology serves as a collaborative tool rather than replacing human input.

- Human-centered AI (HCAI) refers to the development of artificial intelligence (AI) technologies that prioritize **human needs, values, and capabilities** at the core of their design and operation

Introduction

- Researchers, developers, business leaders, policy-makers, and others are expanding the technology-centered scope of artificial intelligence (AI) to include human centered AI (HCAI) ways of thinking.

- Educators, designers, software engineers, product managers, evaluators, and government agency staffers can build on AI-driven technologies to design products and services that make life better for the users, enabling people to care for each other. However, a bright future awaits AI researchers, developers, business leaders, policy-makers, and others who build on AI algorithms by including HCAI strategies of design and testing. This enlarged vision can shape the future of technology so as to better serve human values and needs. As many technology companies and thought leaders have said, the goal is not to replace people but to empower them by making design choices that give humans control over technology.

- Researchers and developers for HCAI systems value meaningful human control, putting people first by serving human values such as rights, justice, and dignity, and supporting goals such as self-efficacy, creativity, responsibility, and social connections.

- Also write any examples

What do we mean by HCAI and what makes it different from AI?

There are many definitions, but there are two key aspects:

1) **Process**: HCAI builds on user experience design methods of user observation, stakeholder engagement, usability testing, iterative refinement, and continuing evaluation of human performance in use of systems that employ AI and machine learning.

2) **Product**: HCAI systems are designed to be super tools which amplify, augment, empower, and enhance human performance. They emphasize human control, while embedding high levels of automation by way of AI and machine learning. Examples include digital cameras and navigation systems, which give humans control yet have many automated features.

The goal is to increase human self-efficacy, creativity, responsibility, and social connections while reducing the impact of malicious actors, biased data, and flawed software.

1. **HCAI framework** that guides creative designers to ensure human-centric thinking about highly automated systems. The examples include familiar devices, such as thermostats, elevators, self-cleaning ovens, and cellphone cameras, as well as life critical applications, such as highly automated cars and patient-controlled pain relief devices. The new aspiration is to have high levels of human control AND high levels of automation.

2. **Design metaphors** suggest how the two central goals of AI research, science and innovation, are both valuable, but researchers, developers, business leaders, and policy-makers will need to be creative in finding effective ways of combining them to benefit the users. There are four design metaphors that can be used to combine the two goals of AI research:
   - **1)** intelligent agents and supertools;
   - **2)** teammates and tele-bots;
   - **3)** assured autonomy and control centers; and
   - **4)** social robots and active appliances.

1. Intelligent agents can be software-based (e.g., chatbots, recommendation systems) or hardware-based (e.g., robots, autonomous vehicles). They often rely on artificial intelligence (AI), machine learning, and data processing techniques to function efficiently.

Super tools refer to advanced AI-powered or technology-driven tools that enhance human capabilities.

EX: AI Assistants (e.g., ChatGPT, Google Assistant) – Enhance productivity with natural language processing.
Healthcare AI (e.g., AI-driven diagnostics, robotic surgeries) – Assist doctors with diagnosis and treatment.

2. **Team mates and Tele-Bots:**

Team mate is a collaboration tool where it is designed to enhance teamwork, especially for remote and hybrid teams.

EX: Microsoft Team– For messaging, file sharing, and team coordination.

Tele Bots:

A **tele bot** (short for **telecommunication bot**) is an AI-powered virtual assistant or chatbot that helps automate conversations and tasks. These bots can be used for: Call and voice Bots, Automotive Bots, Chat bots etc.

3. **assured autonomy and control centres**

It includes various fields like artificial intelligence, robotics, military operations, or cybersecurity.

4. **Social robots and active appliances.**

- A **social robot** could interact with a user and then command an **active appliance** (e.g., “Turn on the air conditioner”).
- **Active appliances** could provide data to a **social robot**, enabling it to give informed suggestions (e.g., “Your fridge is low on milk. Would you like to order more?”).

It will increase benefits for users and society in business, education, healthcare, environmental preservation, and community safety.

3. **Governance structures**

bridge the gap between widely discussed ethical principles and the practical steps needed to realize them. Software team leaders, business managers, and organization leaders will have to adapt proven technical practices, management strategies, and independent oversight methods, so they can achieve the desired goals of:
1) Reliable systems based on proven software engineering practices;

2) Safety culture through business management strategies; and
3) Trustworthy certification by independent oversight and government regulation.

1 Human Values: focus on Rights, Justice and Dignity when using the HCAI.
2 Individual Goals: refers to Self-efficacy, creativity, responsibility and social awareness.
3 Design Aspiration: refers to Reliable, Safe & Trustworthy Team, Organization, Industry & Government.
4 Threats: Human-Centered AI (HCAI) is designed to enhance human decision-making and interactions, but it also poses several threats and challenges. These threats can be categorized into ethical, security, societal, and technical risks.
5 Stakeholders: various individuals, organizations, customers, end users and institutions that influence, develop, regulate, or are affected by AI systems.

Trustworthy certification and clarity about liability comes from accounting firms that conduct independent audits and insurance companies that compensate for failures. Then there are non-governmental and civil society organizations that advance design principles, and professional organizations that develop voluntary standards and prudent policies. Further support for trustworthiness will come from government legislation and regulation, but advocates of certification and independent oversight will have to cope with resistance to regulation and "revolving door" movements in which corporate leaders make jobs in oversight organizations.

Take any examples like Elevetor, self driving cars, camera etc

The three fresh ideas are covered. They are the foundation for achieving the aspirations, goals, and human values shown in Figure which is a compact overview of this book The stakeholders participate in every aspect, while the threats from malicious actors, bias, and flawed software remain prominent in stakeholder minds.

The three fresh ideas are the HCAI framework, design metaphors, and governance structures.
For more explanation refer text book

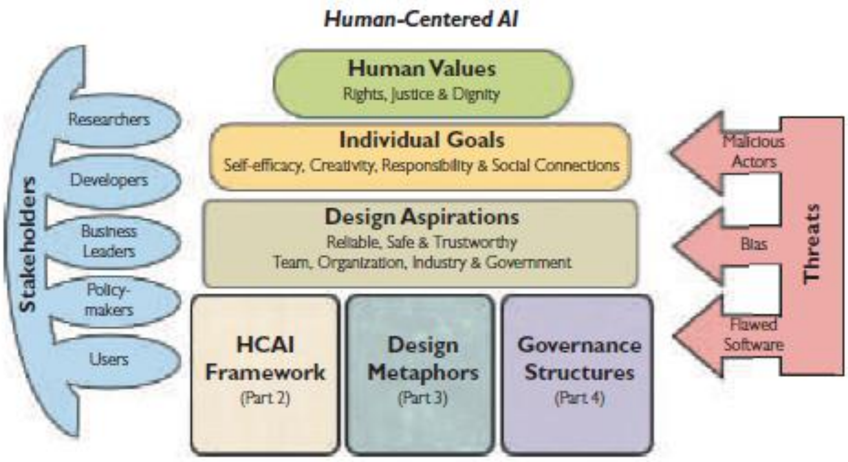

Fig 1.2 The three ideas of this book support the aspirations, goals, and human values, while recognizing the needs of stakeholders and the dangers of threats.

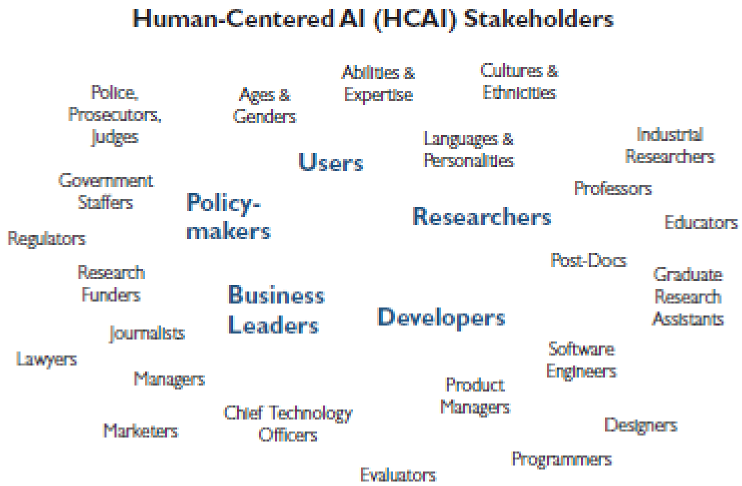

Fig 1.3 HCAI stakeholders with core professionals who are researchers, developers, business leaders, and policy-makers.

Are People and Computers in the Same Category?

Contrast between AI and HCAI advocates is the issue of whether people are in the same category as computers or if they are distinct. In contrast, many HCAI sympathizers believe that there is a vast difference: “People are not computers. Computers are not people.”

- Humans have bodies. Having a body makes you human. It puts us in touch with pain and pleasure, with sadness and joy. Crying and laughing, dancing and eating, love-making and thinking are all parts of being human. Emotions and passions are worth celebrating and fearing.
- Human emotions go far beyond the seven basic emotions that Paul Ekman described as universal: anger, contempt, disgust, enjoyment, fear,

sadness, and surprise.His work, which has been used in the AI community, oversimplifies the complexity of human emotions and their facial expressions.

- One direction for richer views of emotion is to do what many sentiment-analysis programs do, which is to assume that there are many more emotions.
- While there is a large body of work on how computers can detect human emotional states and then respond to them, many researchers now question how accurate these can be. Even if it were possible, the idea of enabling social robots to express emotions in facial features, body language, and spoken language is troubling. Deceptive practices, whether banal or intended, can undermine the very trust that designers seek to build.
- Emotional reactions by computers may be useful in entertainment or game applications, which may be enough to justify the research, but for most applications users want to get their tasks done with minimal distraction. Some users may be annoyed by or distrust computers that pretend to express emotion.
- A more promising and reliable strategy is sentiment analysis, which analyses text in social media posts, product reviews, or newspaper headlines. These aggregate data analyses, not attempt to identify the current emotions of an individual, can show differences in language usage by men and women, Democrats and Republicans, ethnic groups, or socioeconomic clusters.
- Sentiment analysis can also show changes over time, for example, to show that newspaper headlines have become increasingly negative.
- Mimicking or imitating a human by computer is an enjoyable pursuit for some people, but a technology designer's imagination could be liberated by using other inspirations. More ambitious goals lead to valued innovations such as the World Wide Web, information visualization, assistive technology, Wikipedia, and augmented reality. These innovations extend human abilities to enable more people to be more creative more often.

Another question: what value is there in building computers that look and act like people?

- There is a large community of people who believe that human-like, also called anthropomorphic, humanoid, or android, computers are the way of the future.
- This community wants to make social robots with human faces, arms, legs, and speech capabilities that could move around in a human world, maybe as older adult caretakers or disaster response robots. This notion has lead to a long history of failures. Advocates say that this time is

different because computers are so much more powerful and designers are so much more knowledgeable.

- Human–human communication and relationships are just one model, and sometimes a misleading one, for the design of user interfaces.

- Humans relate to humans; humans operate computers. Improved interfaces will enable more people to carry out more tasks more rapidly and effectively.

- Voice is effective for human–human interaction, but visual designs of interfaces will be the dominant strategy because they enable users to operate computers rapidly. Voice user interfaces, such as Alexa and Siri, have an important role, especially when hands are busy and mobility is required, even though the ephemeral and slow nature of voice communication limits its utility.

- Furthermore, human generation of speech commands requires substantial cognitive effort and working memory resources, limiting the parallel effort possible when using hand gestures and controls.

- Interface designs that are consistent, predictable, and controllable are comprehensible, thus enabling mastery, satisfaction, and responsibility. They will be more widely used than ones that are adaptive, autonomous, and anthropomorphic.

- Take any examples

Will Automation, AI, and Robots Lead to Widespread Unemployment?

- Automation appears to increase concentration of wealth that produces evergrowing
  Inequalities,with poorly educated workers suffering the most.
  The focus on growth at the cost of poor employee treatment is a serious danger of increased automation.
- Example: Amazon has fought the introduction of labor unions, which have in the past enabled workers to get better wages, working conditions,health insurance, and benefits such as child care, sick pay, vacations, and pensions.
- **The central question is how to distribute the benefits of automation more equitably.**

The main effect of automation is typically to lower costs and increase quality, which increases demand and expands production, bringing benefits to customers while increasing employment. Automation is also disruptive by making certain skills obsolete, so there are substantial challenges to technology innovators and policy-makers, especially in helping those whose jobs are destroyed and in ensuring greater fairness in distributing the benefits of automation so that workers can make a living wage and receive better treatment.

- **Job Replacement or Job creation Justify**
- **Justify by taking any examples and compare human and computer**

MODULE-2
Human-Centered AI Framework

How human-centered AI ideas open up new possibilities for design of systems that offer high levels of human control and high levels of computer automation.

The human-centered artificial intelligence (HCAI) framework clarifies how to
(1) design for high levels of human control and high levels of computer automation so as to increase human performance,
(2) understand the situations in which full human control or full computer control are necessary, and
(3) avoid the dangers of excessive human control or excessive computer control.
Achieving these goals will also support human self-efficacy, creativity, responsibility, and social connections.

Here's the new possibilities by way of a two-dimensional framework of human-centered artificial intelligence (HCAI) that separates **levels of automation/autonomy from levels of human control**. The new guideline is to seek both *high levels of human control* and *high levels of automation*, which is more likely to produce computer applications that are reliable, safe, and trustworthy. Achieving these goals, especially for complex, poorly understood problems, will dramatically increase human performance, while supporting human self-efficacy, creativity, responsibility, and social connections.

Table 6.1 ***Summary of the widely cited, but one-dimensional levels of automation/autonomy***

Level	Description
	The Computer:
10 (High)	decides everything and acts autonomously, ignoring the human
9	informs the human only if the computer decides to
8	informs the human only if asked
7	executes automatically, then necessarily informs the human
6	allows the human a restricted time to veto before automatic execution
5	executes the suggestion, if the human approves
4	suggests one alternative
3	narrows the selection down to a few
2	offers a complete set of decision-and-action alternatives
1 (Low)	offers no assistance; the human must take all decisions and actions

there have been many refinements such as recognizing that there were at least **four *stages* of automation:**
(1) information acquisition,
(2) analysis of information,
(3) decision or choice of action,
and (4) execution of action

One dimensional level of autonomy self driving cars

Table 6.2 ***Persistent, but still misleading, one-dimensional thinking about levels of autonomy for self-driving cars***[6]

Level	Description
5 (High)	**Full autonomy**: equal to that of a human driver, in every driving scenario.
4	**High automation:** fully autonomous vehicles perform all safety-critical driving functions in certain areas and under defined weather conditions.
3	**Conditional automation:** driver shifts "safety critical functions" to the vehicle under certain traffic or environmental conditions.
2	**Partial automation:** at least one driver assistance system is automated. Driver is disengaged from physically operating the vehicle (hands off the steering wheel AND foot off the pedal at the same time).
1	**Driver assistance:** most functions are still controlled by the driver, but a specific function (like steering or accelerating) can be done automatically by the car.
0 (Low)	**No automation:** human driver controls all: steering, brakes, throttle, power.

Defining Reliable, Safe, and Trustworthy Systems

Machine and human autonomy are both valuable in certain contexts, but a combined strategy uses automation when it is reliable and human control when it is necessary. To guide design improvements it will be helpful to focus on the attributes that make HCAI systems ***reliable, safe, and trustworthy*.**

These terms are complex, but I choose to define them with **four levels of recommendations**
(1) **reliable** systems based on sound software engineering practices,
(2) **safety** culture through business management strategies,
(3) **trustworthy** certification by independent oversight,
and (4) regulation by government bodies.

1. ***Reliable*** systems produce expected responses when needed. Reliability comes from appropriate technical practices for software engineering teams

   When failures occur, investigators can review detailed audit trails, much like the logs of flight data recorders, which have been so effective in civil aviation. The technical practices that support human responsibility, fairness, and explainability include: Audit trails and analysis tools

• Software engineering workflows
• Verification and validation testing
• Bias testing to enhance fairness
• Explainable user interfaces

2. **Safety**

Ample testing of software and analyses of training data used in machine learning promotes fairness. Explainability comes from many design features, but I focus on visual user interfaces that prevent or reduce the need for explainability
Cultures of *safety* are created by managers who focus on these strategies
Leadership commitment to safety

• Hiring and training oriented to safety
• Extensive reporting of failures and near misses
• Internal review boards for problems and future plans
• Alignment with industry standard practices

3. **Trustworthy**

However, public expectations go beyond trust or trusted systems; users want *trustworthy* systems. A system could be mistakenly trusted, but a trustworthy system is one that deserves trust, even though stakeholders struggle to measure Accounting firms conducting external audits

• Insurance companies compensating for failures
• Non-governmental and civil society organizations
• Professional organizations and research institutes

GOVERNMENT AND POLICY MAKERS

- **Government** intervention and regulation also play an important role as protectors of the public interest, especially when large corporations elevate their agendas above the needs of residents and citizens.
- Governments can encourage innovation as much as they can limit it, so learning from successes and failures will help policy-makers to make better decisions.
- I chose to focus on reliable, safe, and trustworthy to simplify the discussion, but the rich literature on these topics advocates other attributes of systems, their performance, and user perceptions
- Users of mature technologies such as elevators, cameras, home appliances, or medical devices know when these devices are reliable, safe, and trustworthy.
- They appreciate the high levels of automation but think of themselves as operating these devices in ways that give them control so as to accomplish their goals.
- Designers who adopt the HCAI mindset will emphasize strategies for enabling diverse users to steer, operate, and control their highly automated devices, while inviting users to exercise their creativity to refine designs.
- Well-designed automation can ensure finer human control, such as in surgical robots that enable surgeons to make precise incisions in difficult to reach organs.

One dimensional levels of automation

Designers of recommender systems, consequential applications, and life critical systems were guided by the one-dimensional levels of autonomy, shared the belief that more automation was better and that to increase automation, designers had to reduce human control.
In short, designers had to choose a point on the one-dimensional line from human control to computer automation

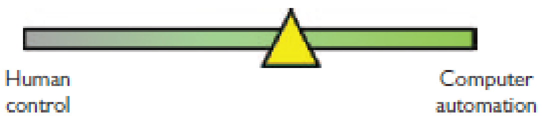

Fig 8.1 Misleading one-dimensional thinking suggests that designers must choose between human control and computer automation.

Two Dimensional levels of automation

The decoupling of these concepts leads to a two-dimensional HCAI framework, which suggests that achieving high levels of human control and high levels of computer automation is possible (yellow triangle in Figure 8.2).

The two axes go from low to high human control and from low to high computer automation. This simple expansion to two dimensions is already helping designers imagine fresh possibilities

1. **Automation Level (X-axis)** – Ranges from **low automation** (AI provides only recommendations) to **high automation** (AI acts autonomously with minimal human intervention).
2. **Human Control & Oversight (Y-axis)** – Ranges from **low human control** (AI acts independently) to **high human control** (humans make all final decisions).

This creates a **four-quadrant model**, categorizing AI systems based on how much control humans retain versus how much automation is involved.

1. Low Automation + High Human Control (Assistive AI)
AI helps humans by providing **recommendations, alerts, or insights**.
Humans make **all final decisions** and remain in full control.
Example: **Lane departure warning systems** in cars (alerts driver but doesn't take control).

2. High Automation + High Human Control (Supervised AI)
AI performs tasks but **requires active human oversight** to intervene when needed.
Example: **Tesla's Autopilot (Level 2)** – the car drives but requires hands on the wheel.

3. Low Automation + Low Human Control (Passive AI)
AI provides **limited automation and minimal human oversight**.
These systems are mostly **static** and don't actively intervene.

Example: **Basic traffic cameras or speed limit detectors** – they provide information but don't take action.

4. High Automation + Low Human Control (Fully Autonomous AI)
AI operates **independently with little or no human oversight**.
Example: **Waymo's self-driving taxis (Level 4/5 autonomy)** – cars navigate without human intervention.

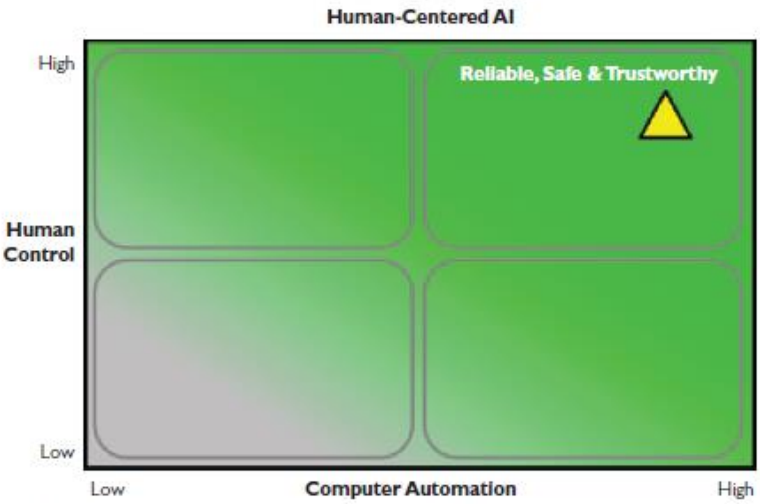

Fig 8.2 Two-dimensional framework with the goal of reliable, safe, and trustworthy designs, which are achieved by a high level of human control and a high level of computer automation (yellow triangle).

The desired goal is often, but not always, to create designs that are in the upper right quadrant. Most reliable, safe, and trustworthy systems are on the right side. The lower right quadrant is home to relatively mature, well understood systems for predictable tasks, for example, automobile automatic transmission or skid control on normal highways. For poorly understood and complex tasks with varying contexts of use, the upper right quadrant is needed. These tasks involve creative decisions, making them currently at the research frontier. As contexts of use are standardized (e.g. elevator shafts) these tasks can come under greater computer control with high levels of automation.

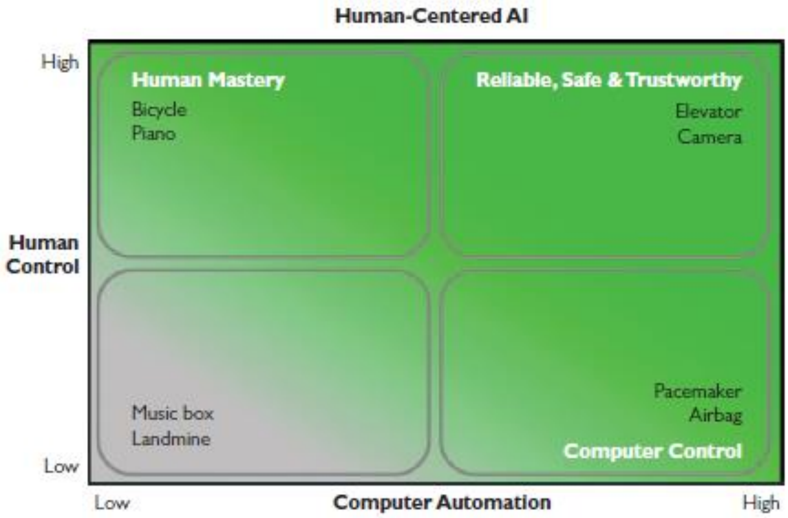

Fig 8.3 Regions requiring rapid action (high automation, low human control) and human mastery (high human control, low automation).

The lower right quadrant (Figure 8.3), with high computer automation and low human control, is the home of computer autonomy requiring rapid action, **for example, airbag deployment**, anti-lock brakes, pacemakers, implantable defibrillators,
or defensive weapons systems. In these applications, there is no time for human intervention or control. Because the price of failure is so high, these applications require extremely careful design, extensive testing, and monitoring during usage to refine designs for different use contexts. As systems mature, users appreciate the effective and proven designs, paving the way for higher
levels of automation and less human supervision.

The upper left quadrant, with high human control and low automation, is the home of human autonomy where human mastery is desired to enable competence building, free exploration, and creativity.
Examples include bicycle riding, piano playing, baking, or playing with children where errors are part of the experience. During these activities, humans generally want to derive pleasure from seeking mastery, improving their skills, and feeling fully engaged. A safe bike ride or a flawless violin performance is events to celebrate. They may elect to use computer-based systems for training, review, or guidance, but many humans desire independent action to achieve mastery that builds self-efficacy. In these actions, the goal is in the doing and the personal satisfaction that it
provides, plus the potential for creative exploration.

The lower left quadrant is the home of simple devices such as clocks, music boxes, or mousetraps, as well as deadly devices such as land mines.

Two other implementation aspects greatly influence reliability, safety, and trustworthiness: the accuracy of the sensors and the fairness of the data. When sensors are unstable or data sources incomplete, error-laden, or biased, then human controls become more important. On the other hand, stable sensors and complete, accurate, and unbiased data favor higher levels of automation.
The take-away message for designers is that, for certain tasks, there is value in full computer control or full human mastery. However, the challenge is to develop effective and proven designs, supported by reliable practices, safety cultures, and trustworthy oversight.

In addition to the extreme cases of full computer and human autonomy, there are two other extreme cases that signal danger—excessive automation and excessive human control.
Ex: Aircraft

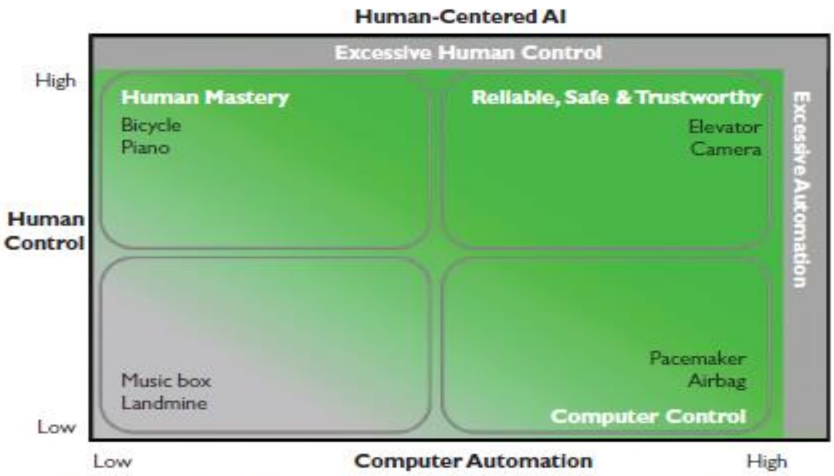

Fig 8.4 Designers need to prevent the failures from excessive automation and excessive human control (gray areas).

On the far right of Figure 8.4 is the region of excessive automation, where there are dangers from designs such as the Boeing which led to two crashes with so many deaths.There are many aspects to this failure, but some basic design principles were violated.

The automated control system took the readings from only one of the two angle of attack sensors which indicate whether the plane is ascending or descending. When this sensor failed, the control system forced the plane's nose downwards, but the pilots did not know why, so they tried to pull the nose up more than twenty times in the minutes before the crash, which killed everyone on board. The aircraft designers made the terrible mistake in believing that their autonomous system for controlling the plane could not fail. Therefore, its existence was not described in the user manual and the pilots were not trained in how to switch to manual override.

Design Guidelines and Examples

Designers can produce HCAI by integrating artificial intelligence algorithms with user interface designs in ways that amplify, augment, empower, and enhance people.

Google's complementary website gives guidelines for responsible AI that are well-aligned with principles.
• Use a human-centered design approach
• Identify multiple metrics to assess training and monitoring
• When possible, directly examine your raw data
• Understand the limitations of your data set and model
• Test, Test, Test
• Continue to monitor and update the system after deployment

"Guidelines for the Design of Human-Autonomy Systems"

For everything we have have aguidelines through that only we will get to know how to use what to you and when to use

Eight Golden Rules

The design decisions to craft user interfaces based on the Eight Golden Rules typically involve trade-offs, so careful study, creative design, and rigorous testing will help designers to produce high-quality user interfaces.

Table 9.1 ***Eight Golden Rules for design***[14]

1. Strive for consistency
2. Seek universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Keep users in control
8. Reduce short-term memory load

Example Digital cameras in most cell phones display an image of what the users would get if they clicked on the large button (Figure 9.3). The image is updated smoothly as users adjust their composition or zoom in. At the same time, the camera makes automatic adjustments to the aperture and focus, while compensating for shaking hands, a wide range of lighting conditions (high dynamic range), and many other factors. Flash can be set to be on or off, or automatically set by the camera. Users can choose portrait modes, panorama, or video, including slow motion and time lapse. Users also can set various filters and once the image is taken they can make further adjustments such as brightness, contrast, saturation, vibrance, shadows, cropping, and red-eye elimination. These designs give users a high degree of control while also providing a high level of automation. Of course, there are mistakes, such as when the automatic focus feature puts the focus on a nearby bush, rather than the person standing just behind it. However, knowledgeable users can touch the desired focus point to override this mistake.

Fig 9.3 Digital cameras give users great flexibility such as taking photos and editing tools to mark up photos, adjust color, change brightness, crop photos, and much more.

There are several examples in text book check it

Eight Golden Rules with an *HCAI pattern language*

There is room to build on these Eight Golden Rules with an *HCAI pattern language* (Table 9.2). Pattern languages have been developed for many design challenges from architecture to social media systems.

Table 9.2 *An HCAI pattern language*

1.	Overview first, zoom and filter, then details-on-demand
2.	Preview first, select and initiate, then manage execution
3.	Steer by way of interactive control panels
4.	Capture history and audit trails from powerful sensors
5.	People thrive on human-to-human communication
6.	Be cautious when outcomes are consequential
7.	Prevent adversarial attacks
8.	Incident reporting websites accelerate refinement

They are brief expressions of important ideas that suggest solutions to common design problems.
Patterns remind designers of vital ideas in compact phrases meant to provoke more careful thinking.

1) Overview first, zoom and filter, then details-on-demand: The first one will be a familiar information visualization mantra that suggests users should be able to get an overview of all the data, possibly as a scattergram, geographic map, or a network diagram. This overview shows the scope and context for individual items, and allows users to zoom in on what they want, filter out what they don't want, and then click for details-on demand.

2) Preview first, select and initiate, then manage execution: For temporal sequences or robot operations the second pattern is to show a preview of the entire process; it allows users to select their plan, initiate the activity, and then manage the execution. This is what navigation tools and digital cameras do so successfully.

3) Steer by way of interactive control panels: Enable users to steer the process or device by way of interactive control panels. This is what users do when driving cars, flying drones, or playing video games. The control panel can include joysticks, buttons, or onscreen buttons, sliders, and other controls, often placed on maps, rooms, or imaginary spaces. Augmented and virtual reality extend the possibilities.

4) Capture history and audit trails from powerful sensors: Aircraft sensors record engine temperature, fuel flow, and dozens of other values, which are saved on the flight data recorder, but are also useful for pilots who want to review what happened 10 minutes ago. Cars and trucks record many items for maintenance reviews; so should applications, websites, data exploration tools, and machine learning models, so users can review their history easily.

5) People thrive on human-to-human communication: Applications are improved when users can easily share content, ask for help, and jointly edit documents. Remember the bumper sticker: Humans in the Group; Computers in the Loop.

6) Be cautious when outcomes are consequential: When applications can affect people's lives, violate privacy, create physical damage, or cause injury, thorough evaluations and continuous monitoring become vital. Independent oversight helps limit damage. Humility is a good attribute for designers to have.

7) Prevent adversarial attacks: Failures can come not only from biased data and flawed software but also from attacks by malicious actors or vandals who would put technology to harmful purposes or simply disrupt normal use.

8) Incident reporting websites accelerate refinement: Openness to feedback from users and stakeholders will bring designers information about failures and near misses, which will enable them to continuously improve their technology products and services

NOTE: Refer Textbook for more content and examples