ACM Intelligent User Interfaces Conference, April 13, 2021

Human-Centered AI: Reliable, Safe & Trustworthy Part 1

Ben Shneiderman **@benbendc**

Founding Director (1983-2000), Human-Computer Interaction Lab
Professor, Department of Computer Science

Member, National Academy of Engineering

UNIVERSITY OF MARYLAND

PETER WALL INSTITUTE FOR ADVANCED STUDIES
THE UNIVERSITY OF BRITISH COLUMBIA VANCOUVER

Photo: BK Adams

HCIL
HUMAN COMPUTER
INTERACTION
LABORATORY

UNIVERSITY OF MARYLAND
18 56

Interdisciplinary research community

- Computer Science & Info Studies
- Psych, Socio, Educ, Jour & MITH

hcil.umd.edu
vimeo.com/72440805

Designing the User Interface

Design Theories

Direct manipulation

Menus, speech, search

Social Media

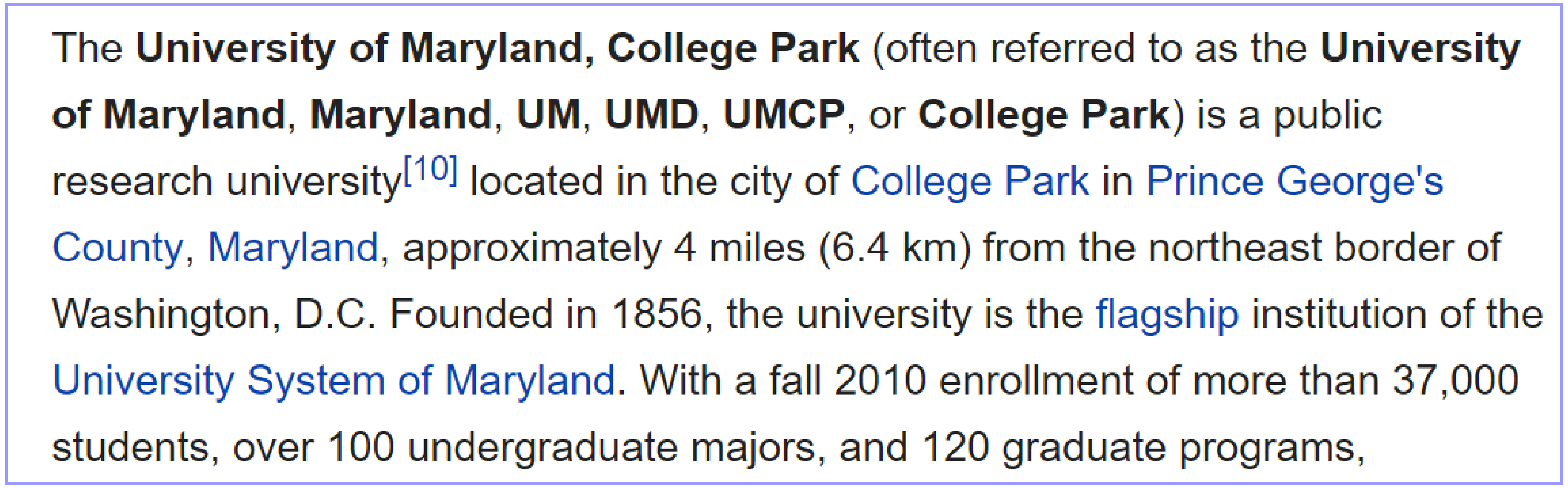
Information Visualization

SHNEIDERMAN • PLAISANT • COHEN • JACOBS • ELMQVIST
SIXTH EDITION
DESIGNING THE USER INTERFACE
STRATEGIES FOR EFFECTIVE
HUMAN-COMPUTER INTERACTION

www.cs.umd.edu/hcil/DTUI6

Sixth Edition: 2016

Web links



The **University of Maryland, College Park** (often referred to as the **University of Maryland**, **Maryland**, **UM**, **UMD**, **UMCP**, or **College Park**) is a public research university[10] located in the city of College Park in Prince George's County, Maryland, approximately 4 miles (6.4 km) from the northeast border of Washington, D.C. Founded in 1856, the university is the flagship institution of the University System of Maryland. With a fall 2010 enrollment of more than 37,000 students, over 100 undergraduate majors, and 120 graduate programs,

Tiny touchscreen keyboards

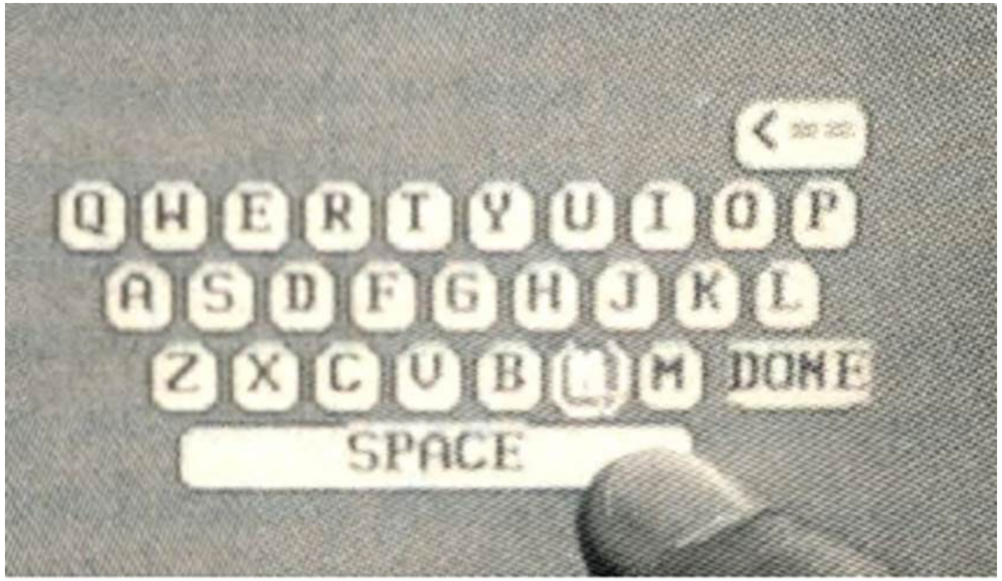
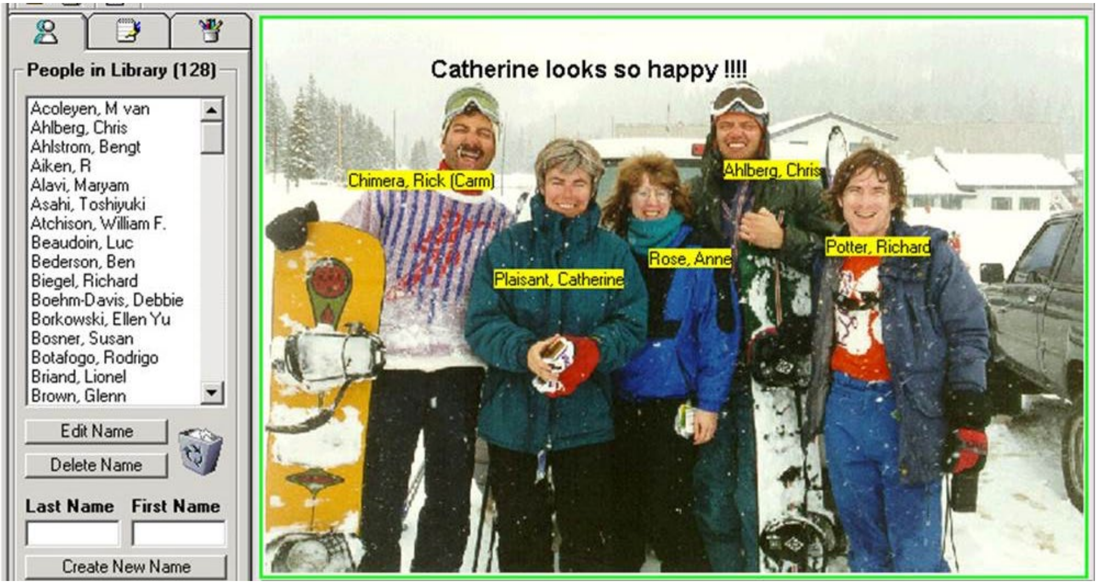


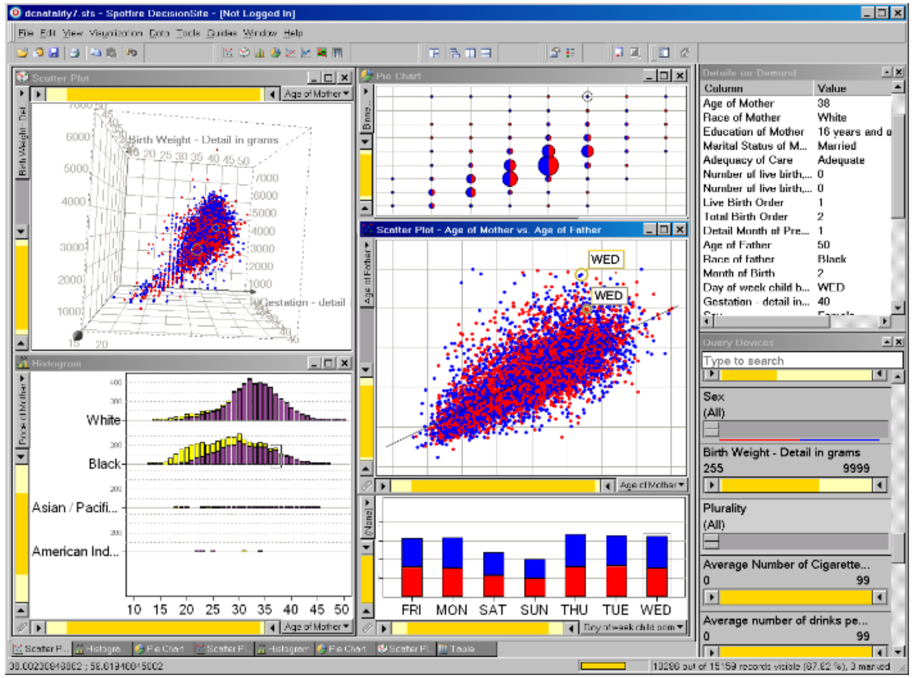
Photo tagging



People in Library (128)

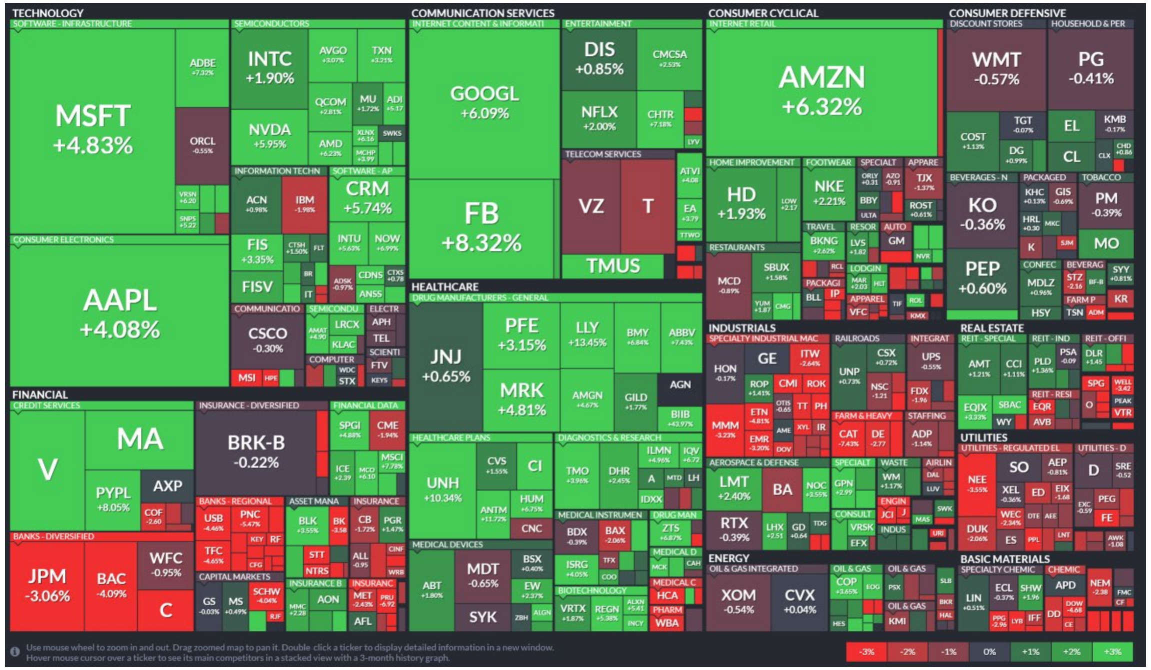
Acoleyen, M van
Ahlberg, Chris
Ahlstrom, Bengt
Aiken, R
Alavi, Maryam
Asahi, Toshiyuki
Atchison, William F.
Beaudoin, Luc
Bederson, Ben
Biegel, Richard
Boehm-Davis, Debbie
Borkowski, Ellen Yu
Bosner, Susan
Botafogo, Rodrigo
Briand, Lionel
Brown, Glenn

Catherine looks so happy !!!!

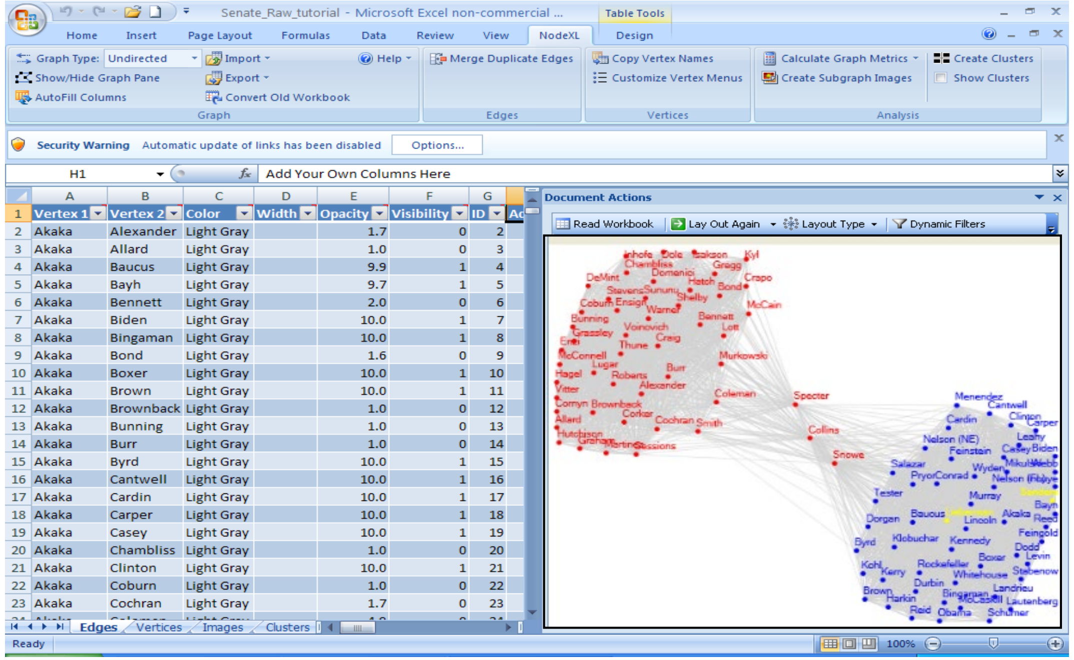

dcnatality7.sts - Spotfire DecisionSite - [Not Logged In]
Scatter Plot
Birth Weight - Detail in grams
Gestation - detail
Pie Chart
Scatter Plot - Age of Mother vs. Age of Father
WED
Histogram
White
Black
Asian / Pacifi...
American Ind...
FRI
MON
SAT
SUN
THU
TUE
WED
Details on Demand
Query Devices
Type to search
Sex
(All)
Birth Weight - Detail in grams
255
9999
Plurality
(All)
Average Number of Cigarette...
0
99
Average number of drinks pe...
0
99

Column	Value
Age of Mother	38
Race of Mother	White
Education of Mother	16 years and o
Marital Status of M...	Married
Adequacy of Care	Adequate
Number of live birth,...	0
Number of live birth,...	0
Live Birth Order	1
Total Birth Order	2
Detail Month of Pre...	1
Age of Father	50
Race of father	Black
Month of Birth	2
Day of week child b...	WED
Gestation - detail in...	40

Spotfire



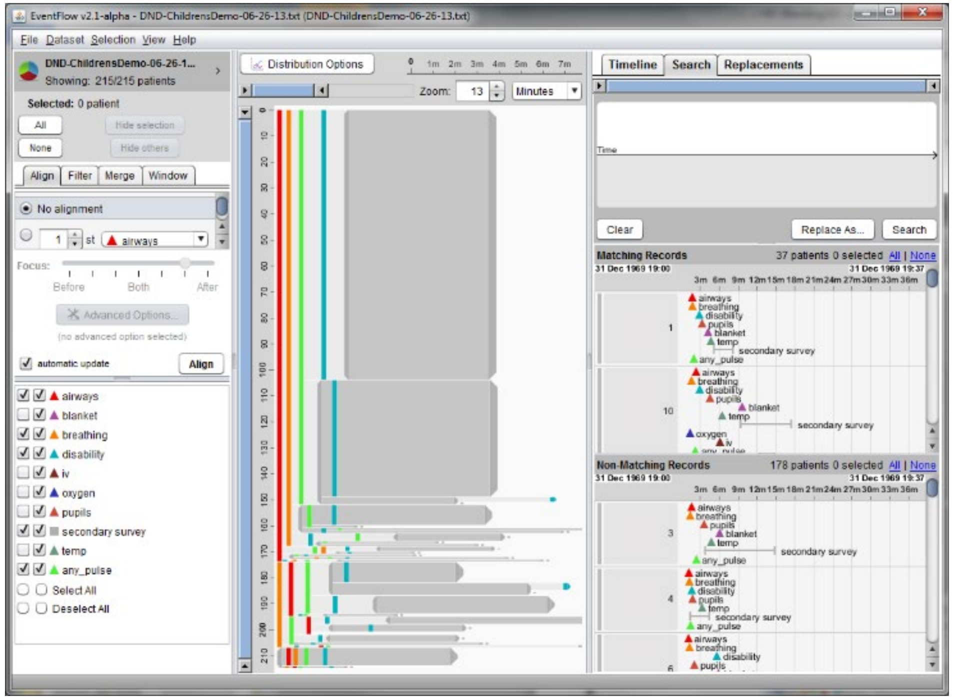
Treemaps
FinViz



H1 — Add Your Own Columns Here

	Vertex 1	Vertex 2	Color	Width	Opacity	Visibility	ID
2	Akaka	Alexander	Light Gray		1.7	0	2
3	Akaka	Allard	Light Gray		1.0	0	3
4	Akaka	Baucus	Light Gray		9.9	1	4
5	Akaka	Bayh	Light Gray		9.7	1	5
6	Akaka	Bennett	Light Gray		2.0	0	6
7	Akaka	Biden	Light Gray		10.0	1	7
8	Akaka	Bingaman	Light Gray		10.0	1	8
9	Akaka	Bond	Light Gray		1.6	0	9
10	Akaka	Boxer	Light Gray		10.0	1	10
11	Akaka	Brown	Light Gray		10.0	1	11
12	Akaka	Brownback	Light Gray		1.0	0	12
13	Akaka	Bunning	Light Gray		1.0	0	13
14	Akaka	Burr	Light Gray		1.0	0	14
15	Akaka	Byrd	Light Gray		10.0	1	15
16	Akaka	Cantwell	Light Gray		10.0	1	16
17	Akaka	Cardin	Light Gray		10.0	1	17
18	Akaka	Carper	Light Gray		10.0	1	18
19	Akaka	Casey	Light Gray		10.0	1	19
20	Akaka	Chambliss	Light Gray		1.0	0	20
21	Akaka	Clinton	Light Gray		10.0	1	21
22	Akaka	Coburn	Light Gray		1.0	0	22
23	Akaka	Cochran	Light Gray		1.7	0	23

NodeXL



EventFlow

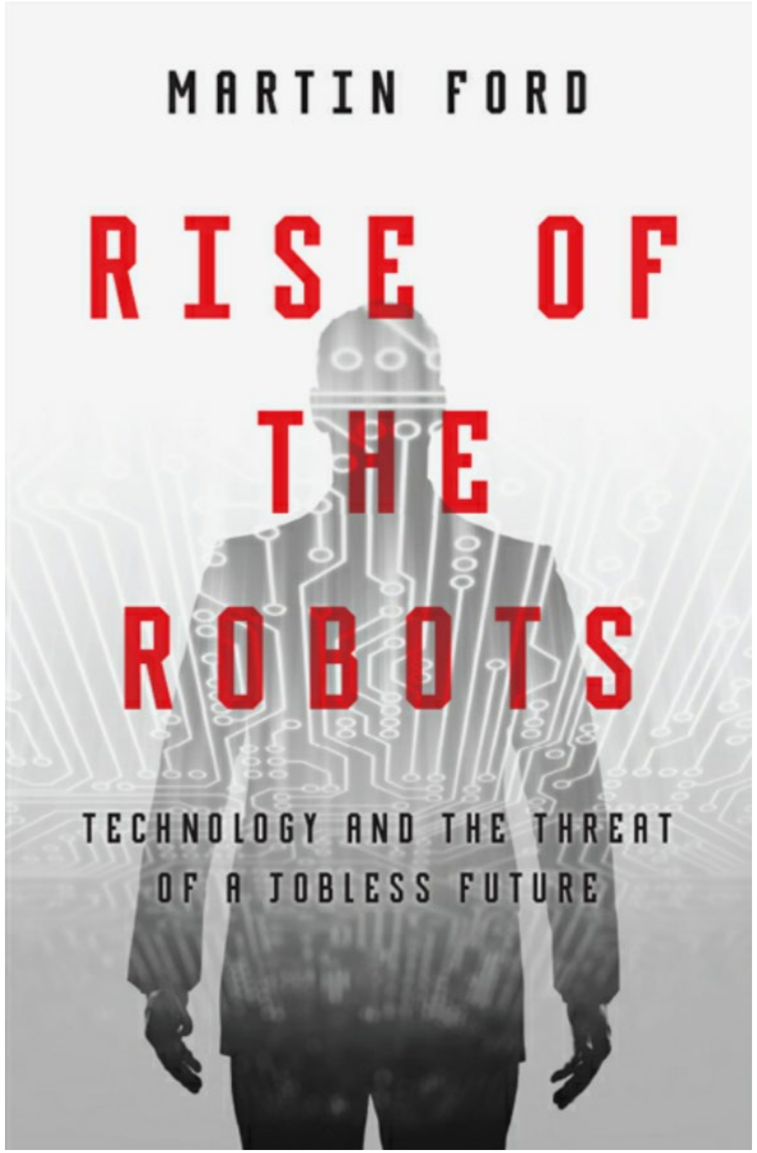
The Goal of Visualization is Insight, Not Pictures

The Goal of Visualization is Insight, Not Pictures

Information Visualization Mantra

- Overview first, zoom & filter, then details-on-demand
- Overview first, zoom & filter, then details-on-demand
- Overview first, zoom & filter, then details-on-demand
- Overview first, zoom & filter, then details-on-demand
- Overview first, zoom & filter, then details-on-demand
- Overview first, zoom & filter, then details-on-demand
- Overview first, zoom & filter, then details-on-demand

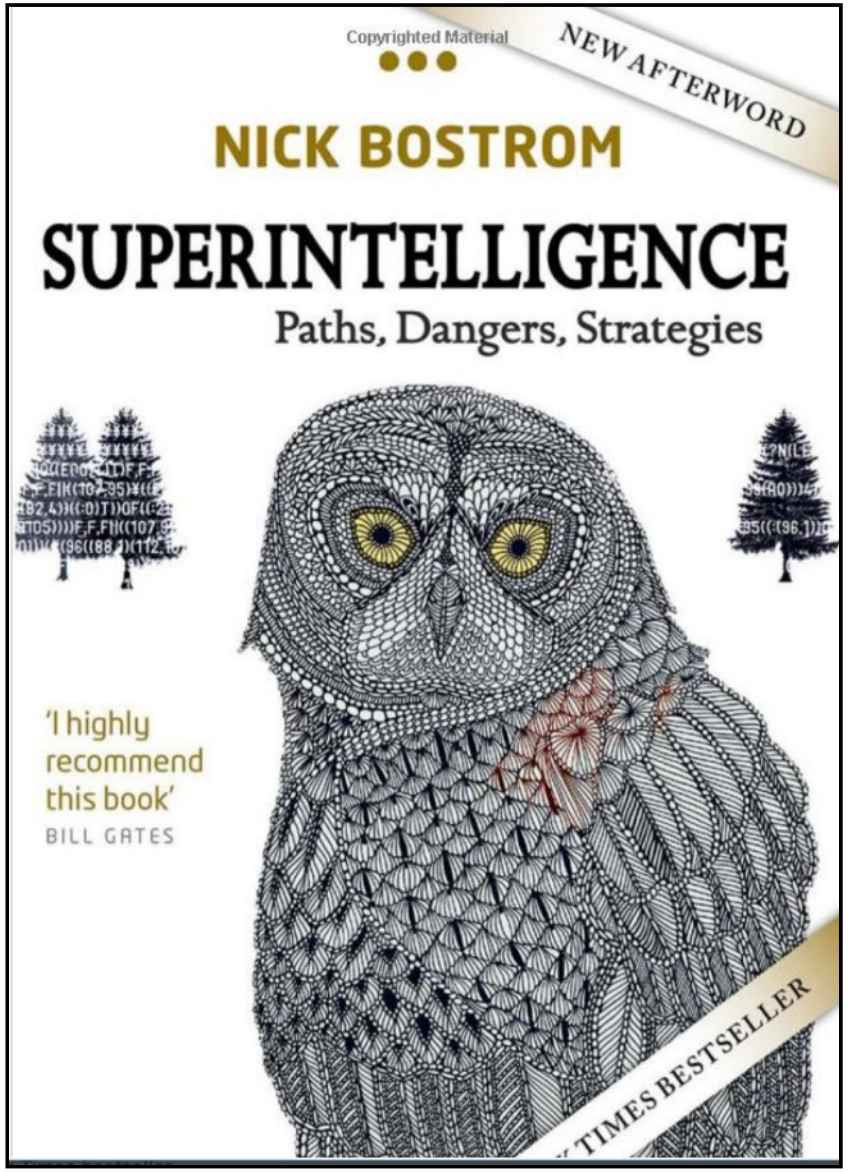
(IEEE, *VLC*,1996)



MARTIN FORD

RISE OF THE ROBOTS

TECHNOLOGY AND THE THREAT OF A JOBLESS FUTURE



Copyrighted Material

NEW AFTERWORD

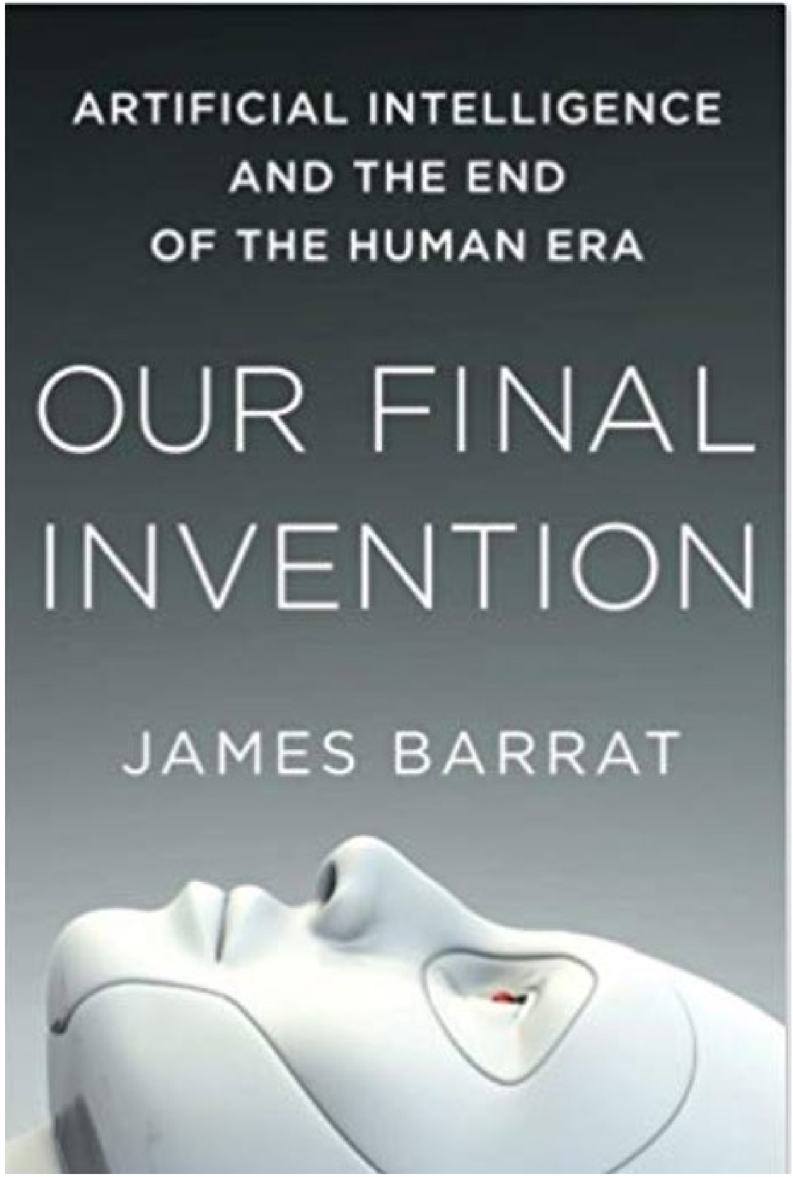
NICK BOSTROM

SUPERINTELLIGENCE

Paths, Dangers, Strategies

'I highly recommend this book'
BILL GATES

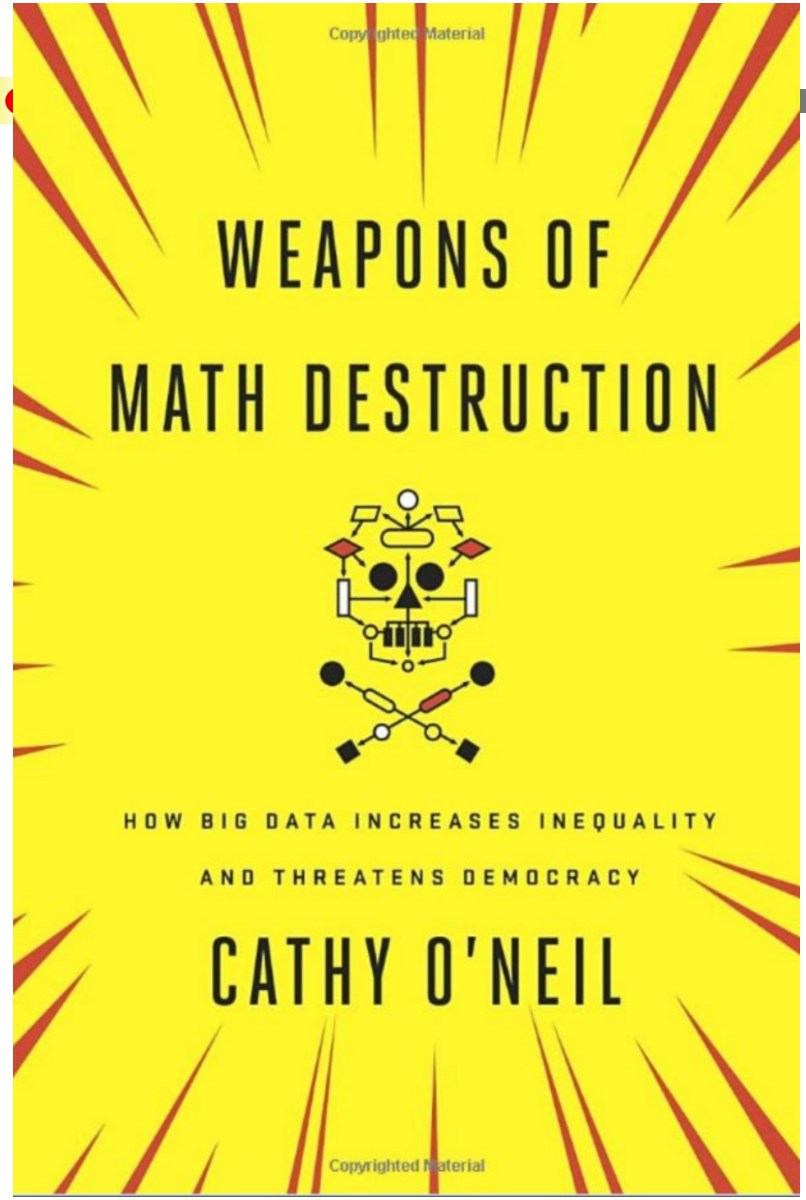
TIMES BESTSELLER

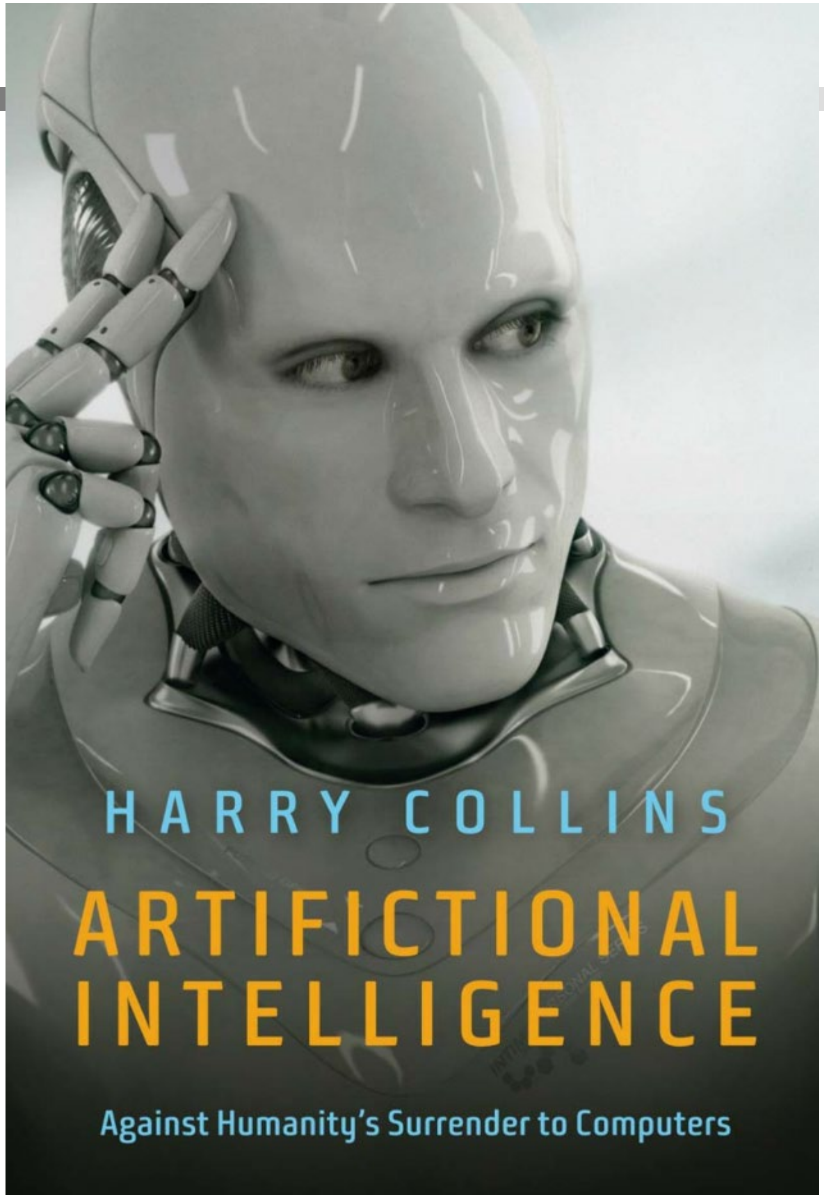


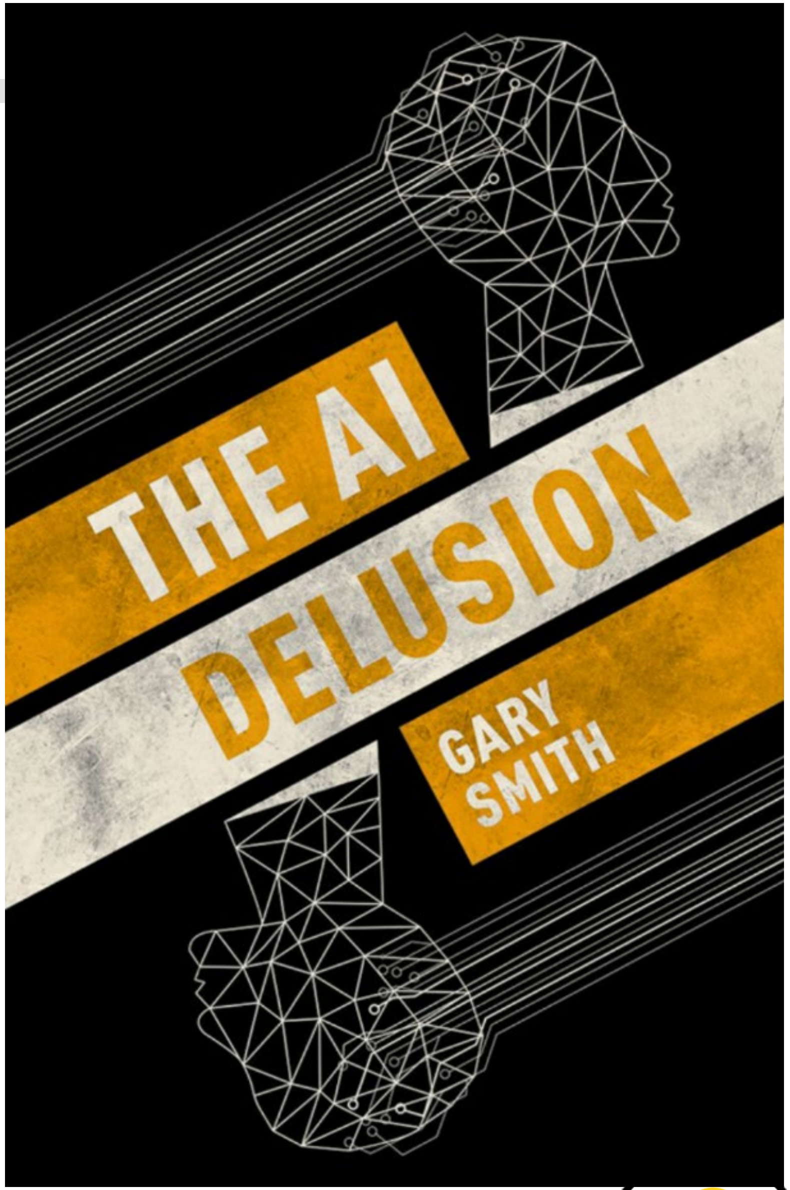
ARTIFICIAL INTELLIGENCE AND THE END OF THE HUMAN ERA

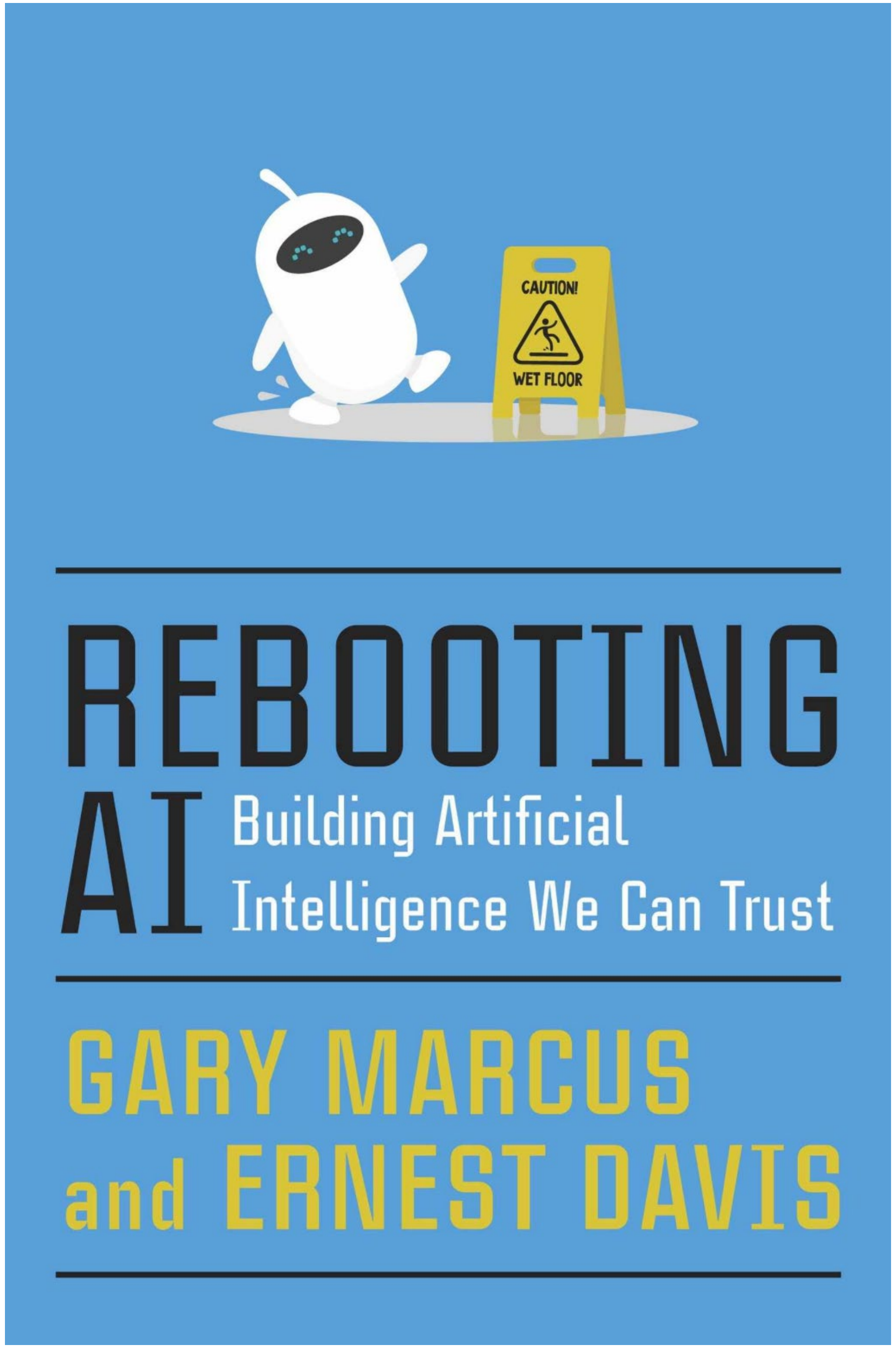
OUR FINAL INVENTION

JAMES BARRAT


Copyrighted Material
WEAPONS OF MATH DESTRUCTION
HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY
CATHY O'NEIL
Copyrighted Material


HARRY COLLINS
ARTIFICTIONAL INTELLIGENCE
Against Humanity's Surrender to Computers


THE AI DELUSION
GARY SMITH



“so much of what we read about AI strikes us as pure fantasy, predicated on a confidence in AI’s imagined strengths that bears no relation to current technological capabilities”

What is Human-Centered AI?

What is Human-Centered AI?

Amplify, Augment, Empower & Enhance People

Human-Centered AI

Human Values
Rights, Justice & Dignity

Human-Centered AI

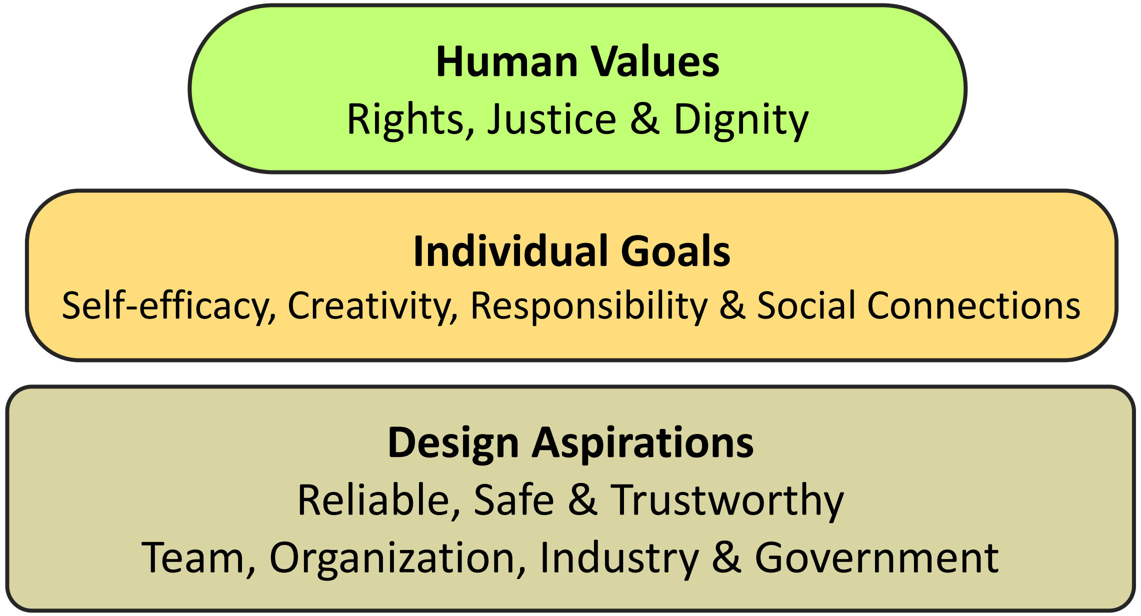
Human Values

Rights, Justice & Dignity

Individual Goals

Self-efficacy, Creativity, Responsibility & Social Connections

Human-Centered AI



Human Values

Rights, Justice & Dignity

Individual Goals

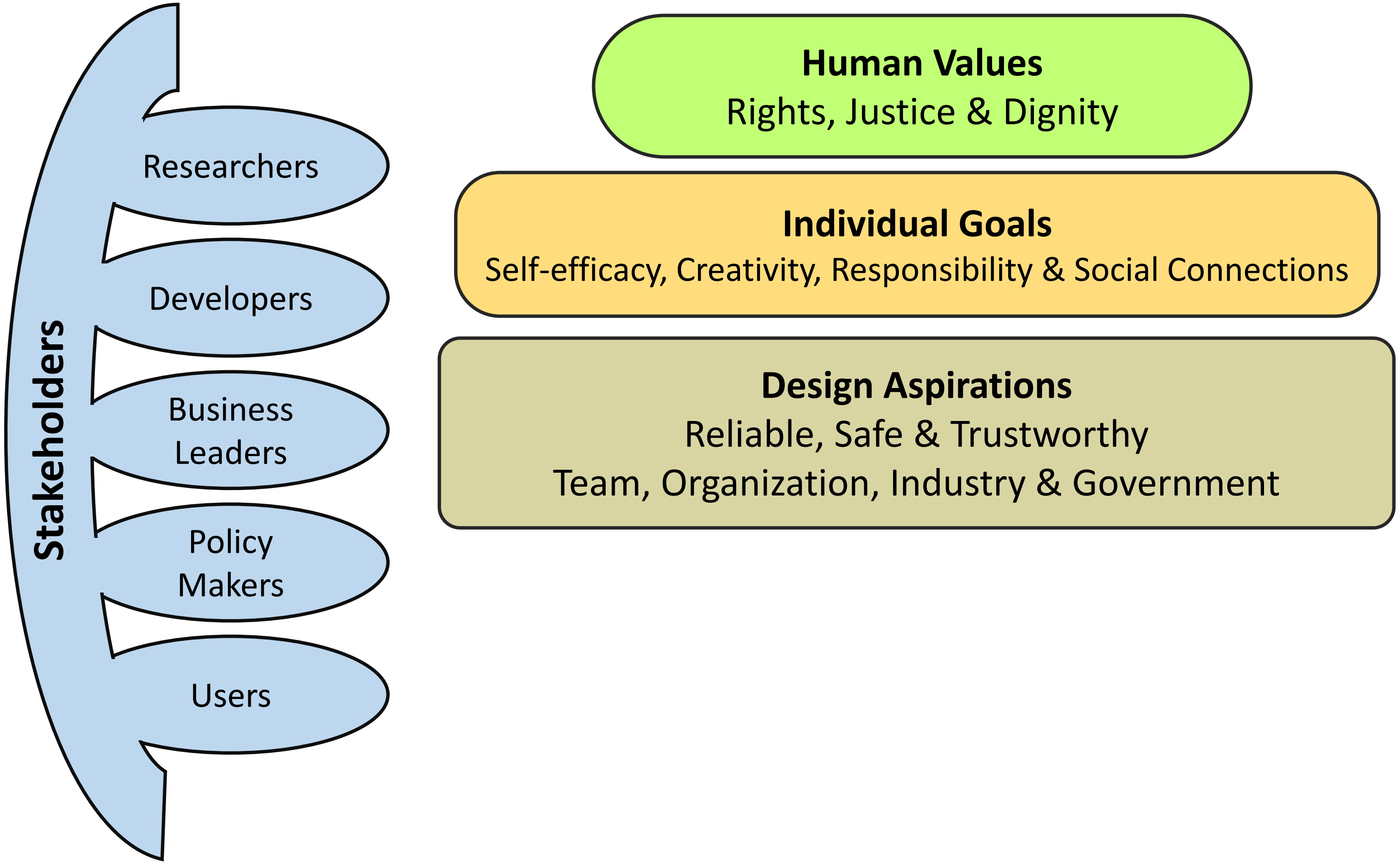
Self-efficacy, Creativity, Responsibility & Social Connections

Design Aspirations

Reliable, Safe & Trustworthy

Team, Organization, Industry & Government

Human-Centered AI



Stakeholders

Researchers

Developers

Business Leaders

Policy Makers

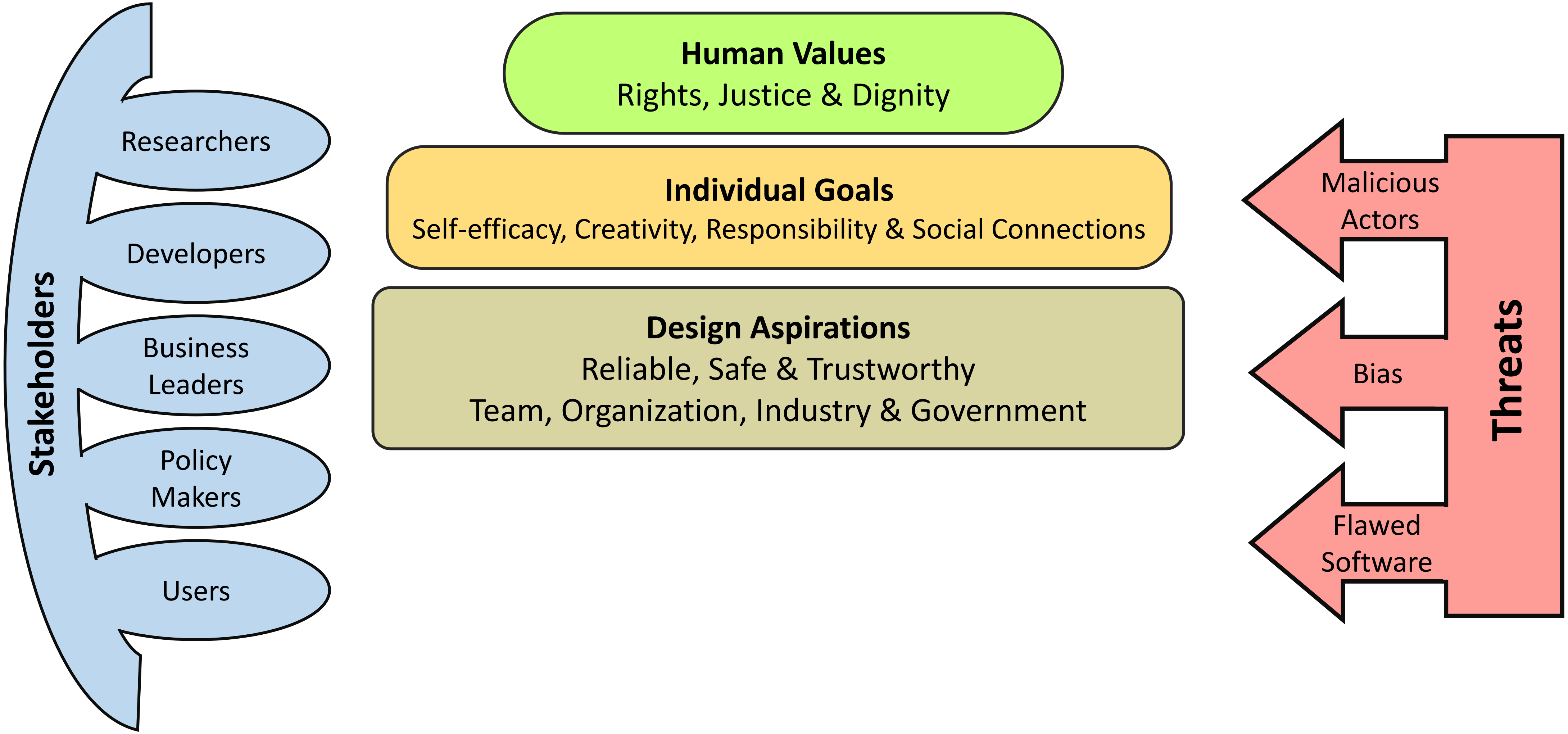
Users

Human Values
Rights, Justice & Dignity

Individual Goals
Self-efficacy, Creativity, Responsibility & Social Connections

Design Aspirations
Reliable, Safe & Trustworthy
Team, Organization, Industry & Government

Human-Centered AI



Stakeholders

Researchers

Developers

Business Leaders

Policy Makers

Users

Human Values
Rights, Justice & Dignity

Individual Goals
Self-efficacy, Creativity, Responsibility & Social Connections

Design Aspirations
Reliable, Safe & Trustworthy
Team, Organization, Industry & Government

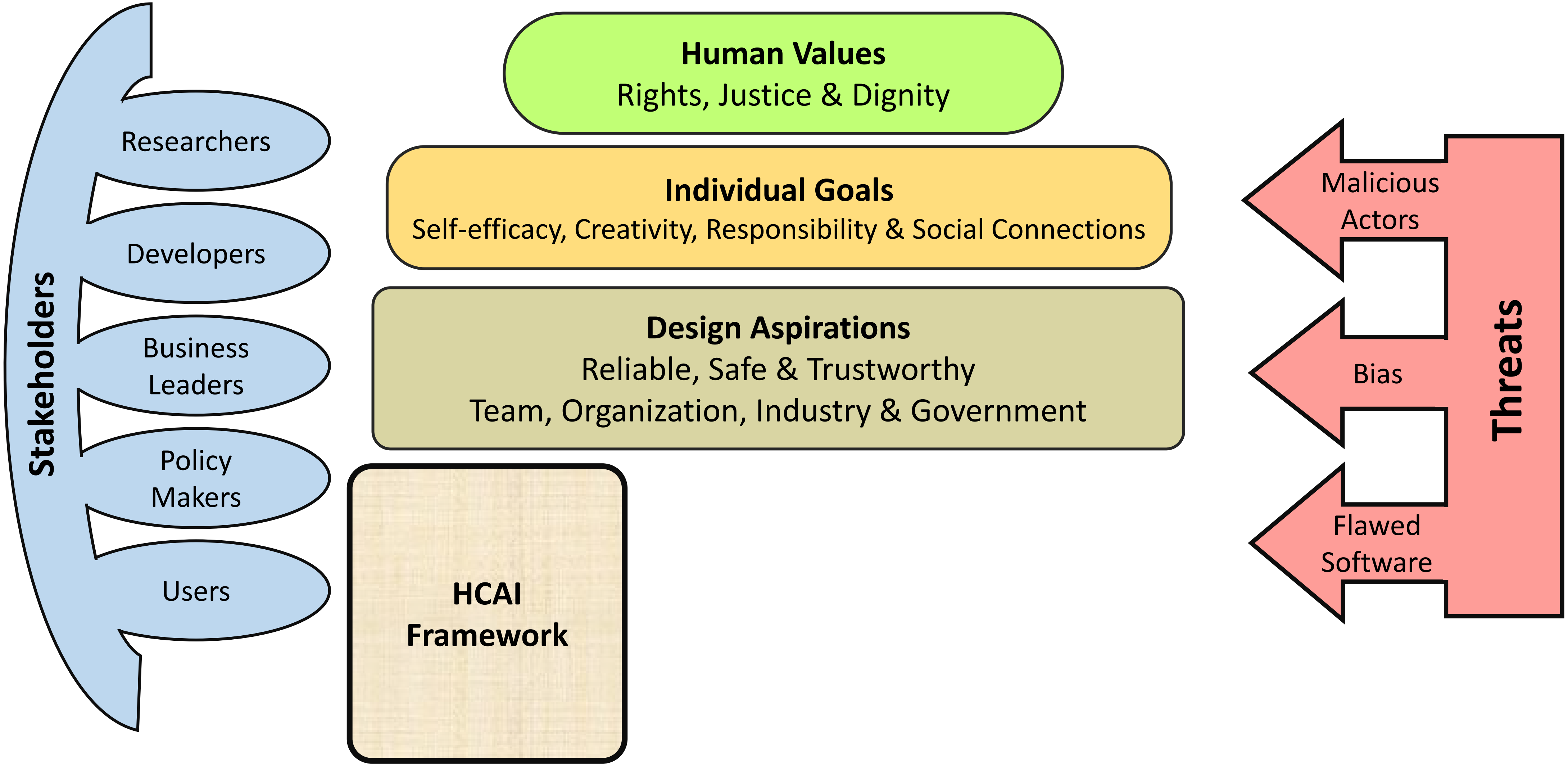
Threats

Malicious Actors

Bias

Flawed Software

Human-Centered AI



Human Values
Rights, Justice & Dignity

Individual Goals
Self-efficacy, Creativity, Responsibility & Social Connections

Design Aspirations
Reliable, Safe & Trustworthy
Team, Organization, Industry & Government

HCAI Framework

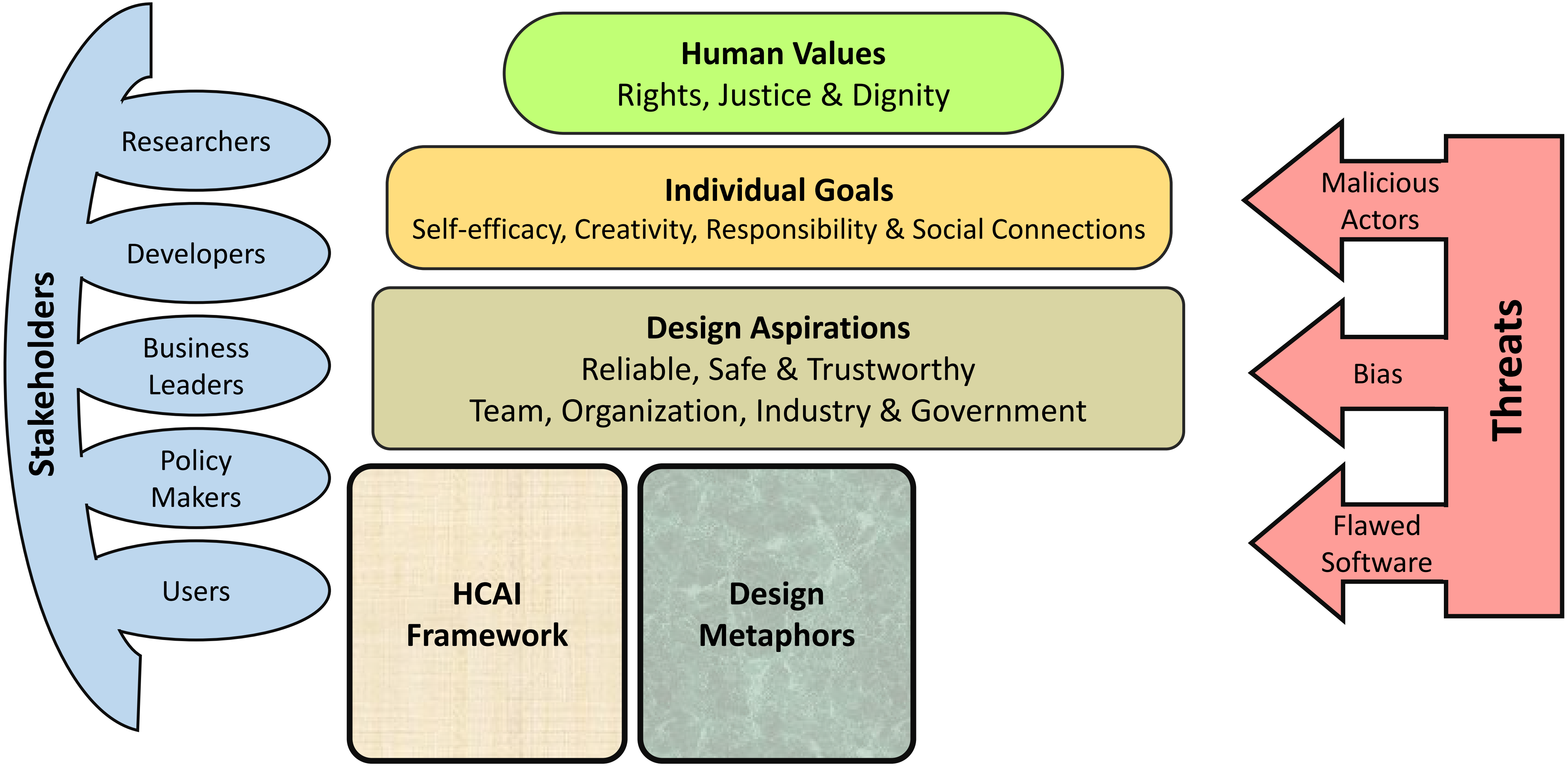
Stakeholders

- Researchers
- Developers
- Business Leaders
- Policy Makers
- Users

Threats

- Malicious Actors
- Bias
- Flawed Software

Human-Centered AI



Stakeholders

Researchers

Developers

Business Leaders

Policy Makers

Users

Human Values
Rights, Justice & Dignity

Individual Goals
Self-efficacy, Creativity, Responsibility & Social Connections

Design Aspirations
Reliable, Safe & Trustworthy
Team, Organization, Industry & Government

HCAI Framework

Design Metaphors

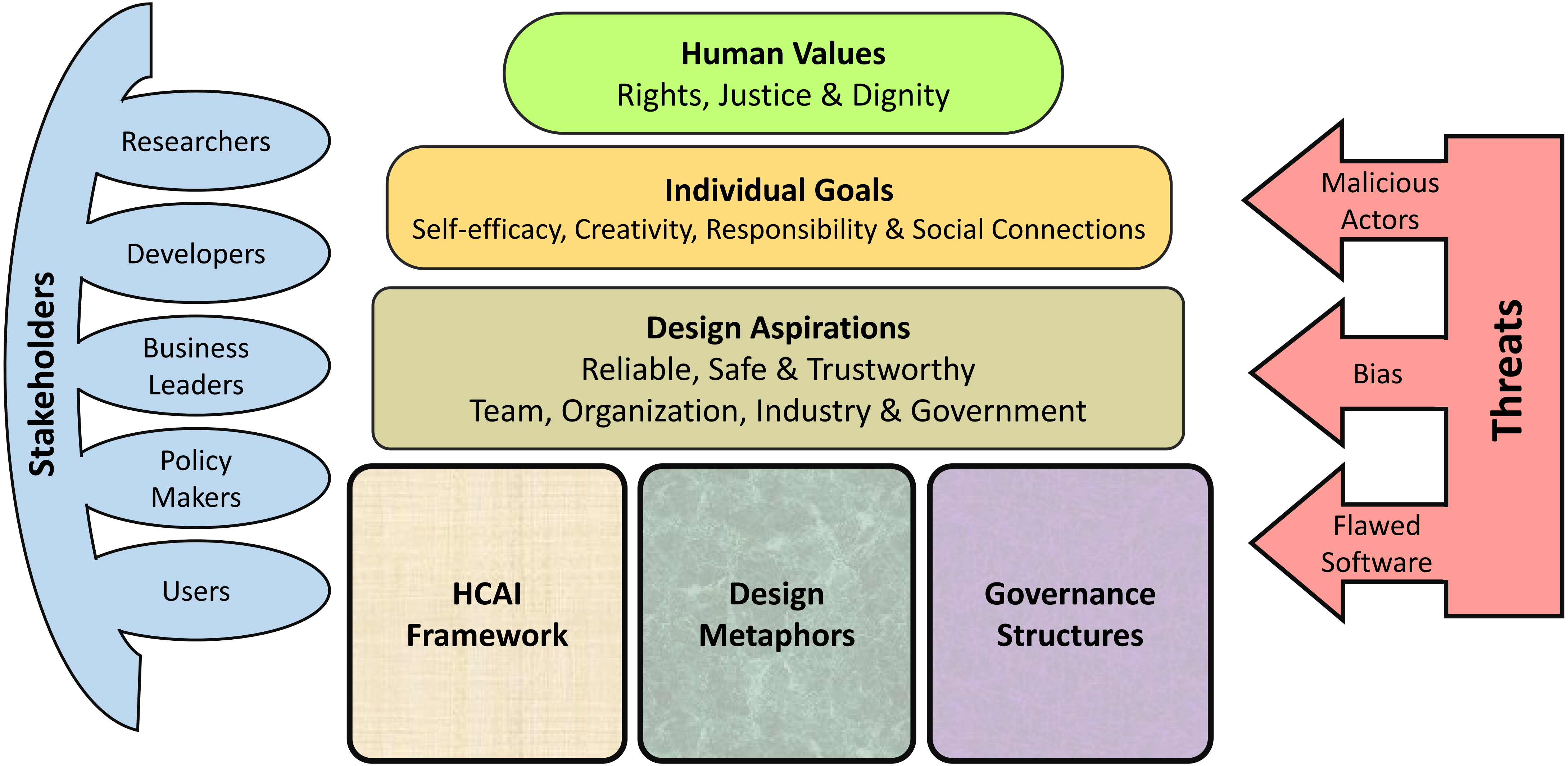
Threats

Malicious Actors

Bias

Flawed Software

Human-Centered AI



Stakeholders

Researchers

Developers

Business Leaders

Policy Makers

Users

Human Values
Rights, Justice & Dignity

Individual Goals
Self-efficacy, Creativity, Responsibility & Social Connections

Design Aspirations
Reliable, Safe & Trustworthy
Team, Organization, Industry & Government

HCAI Framework

Design Metaphors

Governance Structures

Threats

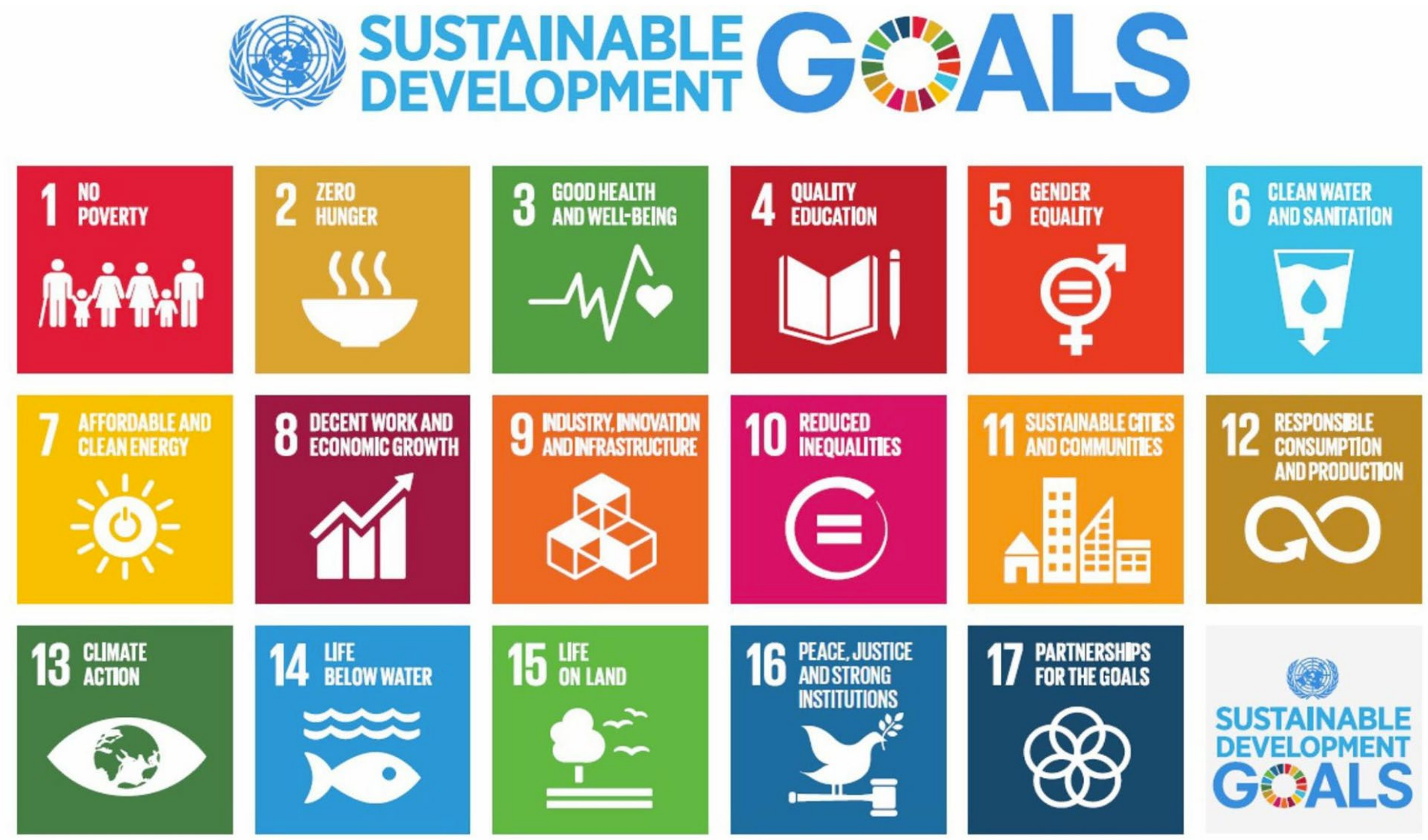
Malicious Actors

Bias

Flawed Software

Oxford University Press (Early 2022) `https://hcil.umd.edu/human-centered-ai/`

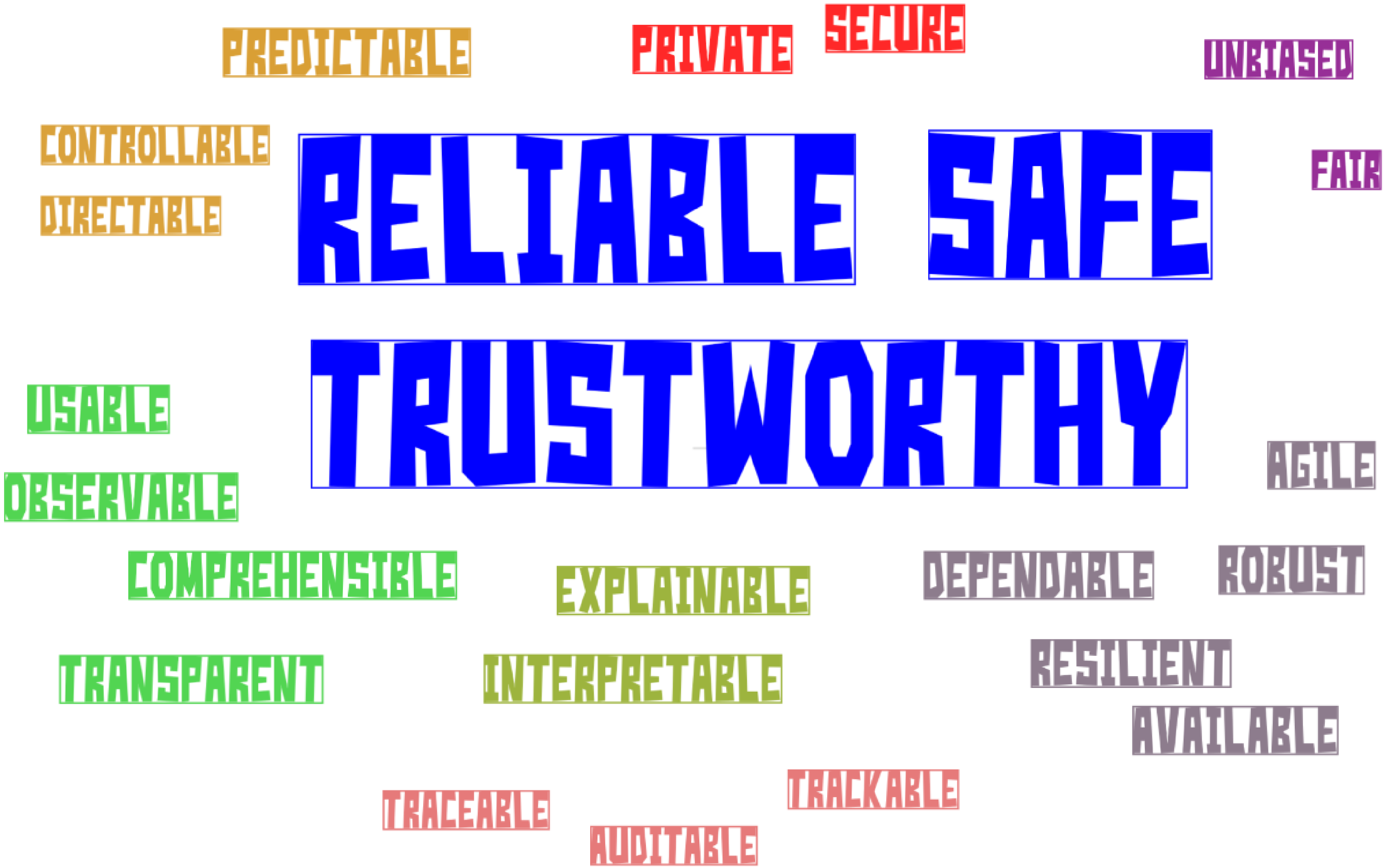
UN Sustainable Development Goals


SUSTAINABLE DEVELOPMENT GOALS
1 NO POVERTY
2 ZERO HUNGER
3 GOOD HEALTH AND WELL-BEING
4 QUALITY EDUCATION
5 GENDER EQUALITY
6 CLEAN WATER AND SANITATION
7 AFFORDABLE AND CLEAN ENERGY
8 DECENT WORK AND ECONOMIC GROWTH
9 INDUSTRY, INNOVATION AND INFRASTRUCTURE
10 REDUCED INEQUALITIES
11 SUSTAINABLE CITIES AND COMMUNITIES
12 RESPONSIBLE CONSUMPTION AND PRODUCTION
13 CLIMATE ACTION
14 LIFE BELOW WATER
15 LIFE ON LAND
16 PEACE, JUSTICE AND STRONG INSTITUTIONS
17 PARTNERSHIPS FOR THE GOALS
SUSTAINABLE DEVELOPMENT GOALS

(https://sdgs.un.org/goals)

People are not computers; Computers are not people

HCAI Attributes



PREDICTABLE
PRIVATE
SECURE
UNBIASED
CONTROLLABLE
DIRECTABLE
RELIABLE
SAFE
FAIR
TRUSTWORTHY
USABLE
OBSERVABLE
AGILE
COMPREHENSIBLE
EXPLAINABLE
DEPENDABLE
ROBUST
TRANSPARENT
INTERPRETABLE
RESILIENT
AVAILABLE
TRACKABLE
TRACEABLE
AUDITABLE

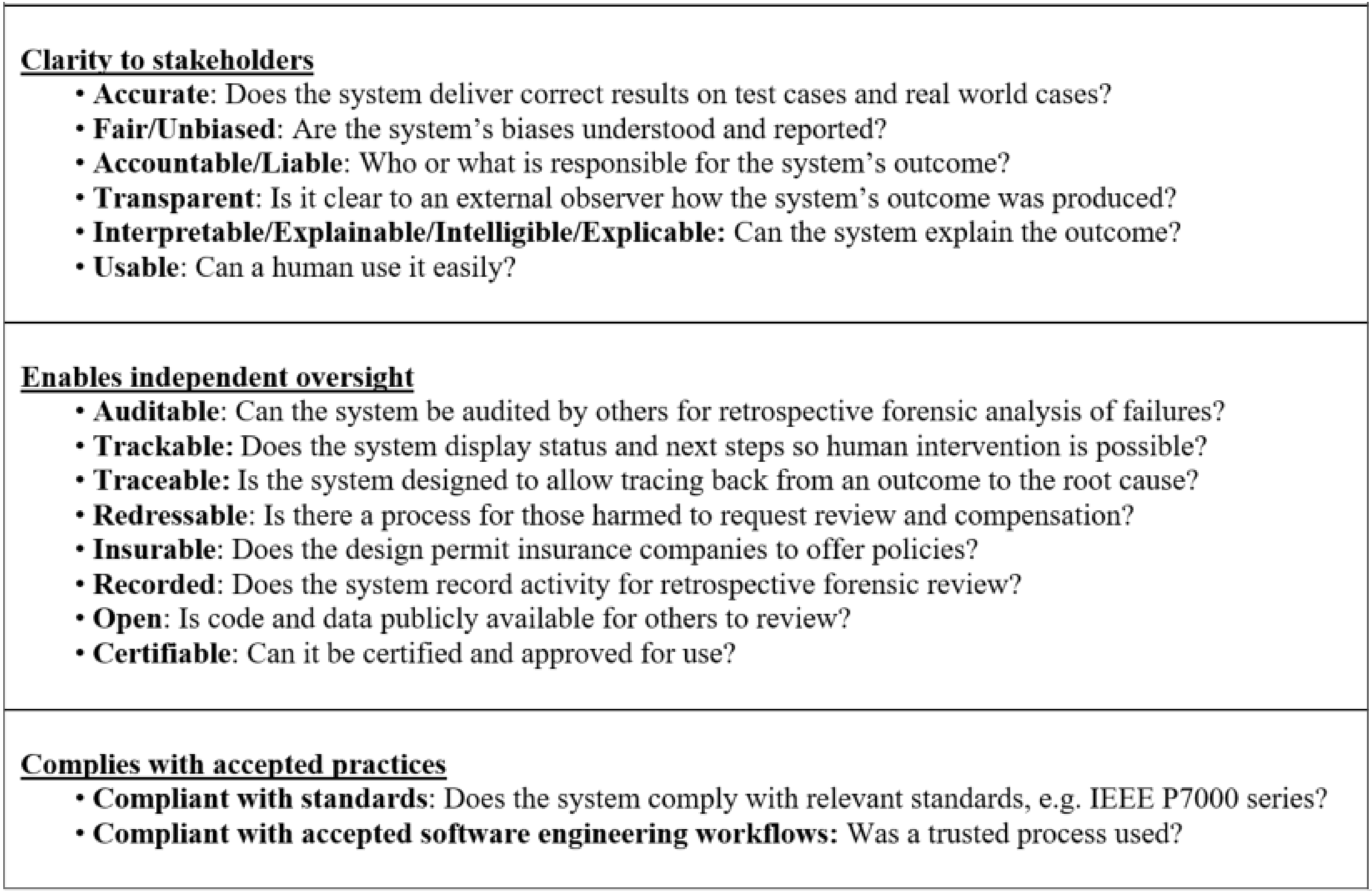
HCAI Attributes that Are Candidates for Assessment

General virtues of the system itself

- **Trustworthy**: Can users trust the system to perform correctly?
- **Responsible/Humane**: Has the system been designed, developed, and tested in a responsible way?
- **Ethical Design**: Were stakeholders involved in the design?
- **Ethical Data**: Was the data collected in an ethical manner?
- **Ethical Use**: Will the system's outcome be used in an ethical manner?
- **Well-being/Benevolence**: Does the system support human health, comfort, and values?
- **Secure**: How vulnerable is the system to attack?
- **Private**: Does the system protect a person's identity and data?

Performs well in practice

- **Robust/Agile**: Does the system perform well when inputs change?
- **Reliable/Dependable**: Does the system do the right thing?
- **Available**: Is the system running when needed?
- **Resilient/Adaptive**: Can the system recover from disruptions?
- **Testable/Verifiable/Validatable/Certifiable:** Can be tested to verify adherence to requirements?
- **Safe**: Does the system have a history of safe use?



Clarity to stakeholders

- **Accurate**: Does the system deliver correct results on test cases and real world cases?
- **Fair/Unbiased**: Are the system's biases understood and reported?
- **Accountable/Liable**: Who or what is responsible for the system's outcome?
- **Transparent**: Is it clear to an external observer how the system's outcome was produced?
- **Interpretable/Explainable/Intelligible/Explicable:** Can the system explain the outcome?
- **Usable**: Can a human use it easily?

Enables independent oversight

- **Auditable**: Can the system be audited by others for retrospective forensic analysis of failures?
- **Trackable:** Does the system display status and next steps so human intervention is possible?
- **Traceable:** Is the system designed to allow tracing back from an outcome to the root cause?
- **Redressable**: Is there a process for those harmed to request review and compensation?
- **Insurable**: Does the design permit insurance companies to offer policies?
- **Recorded**: Does the system record activity for retrospective forensic review?
- **Open**: Is code and data publicly available for others to review?
- **Certifiable**: Can it be certified and approved for use?

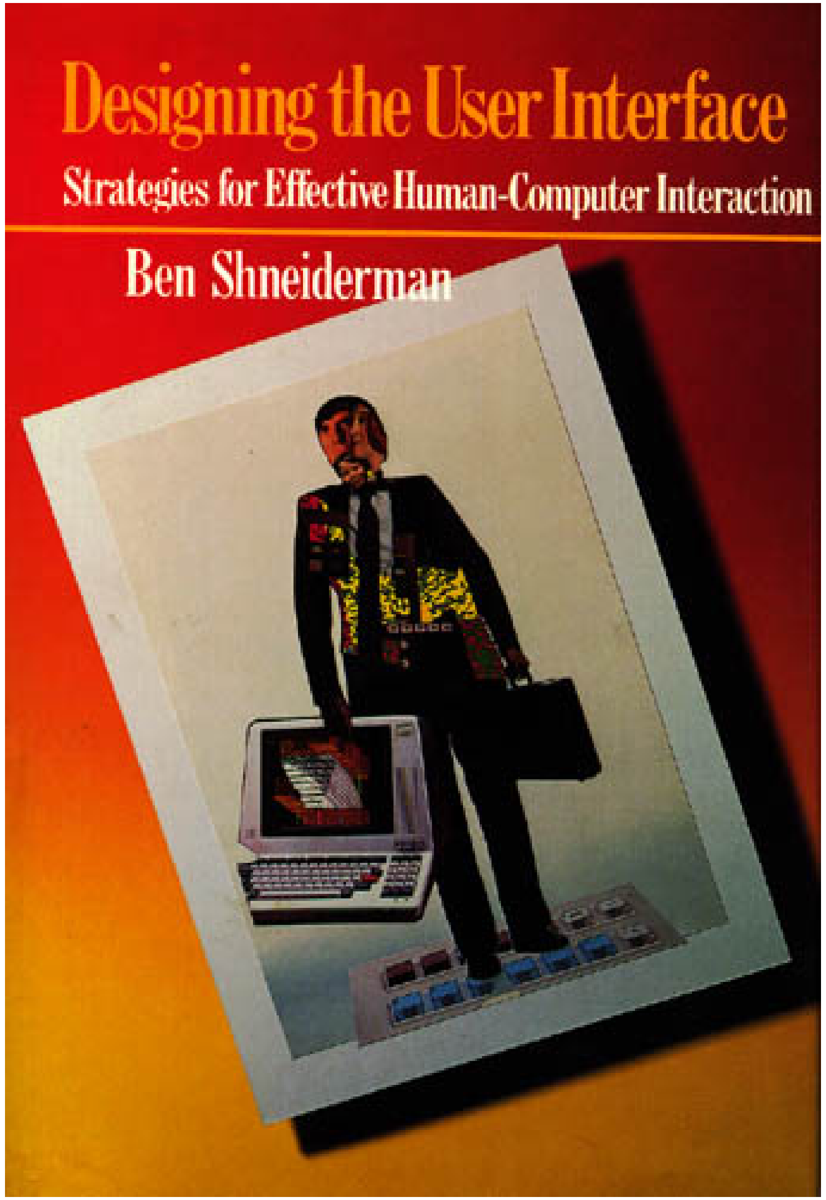
Complies with accepted practices

- **Compliant with standards**: Does the system comply with relevant standards, e.g. IEEE P7000 series?
- **Compliant with accepted software engineering workflows:** Was a trusted process used?

HCAI Framework

Designing the User Interface

Balancing automation & human control



Designing the User Interface

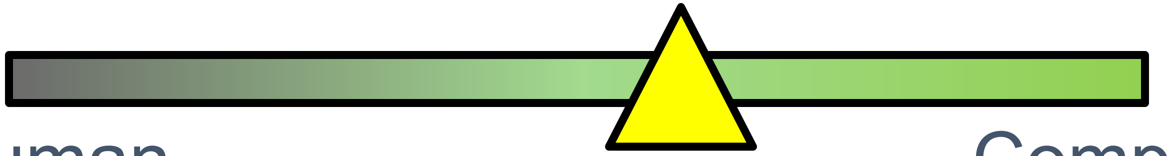
Strategies for Effective Human-Computer Interaction

Ben Shneiderman

First Edition: 1986

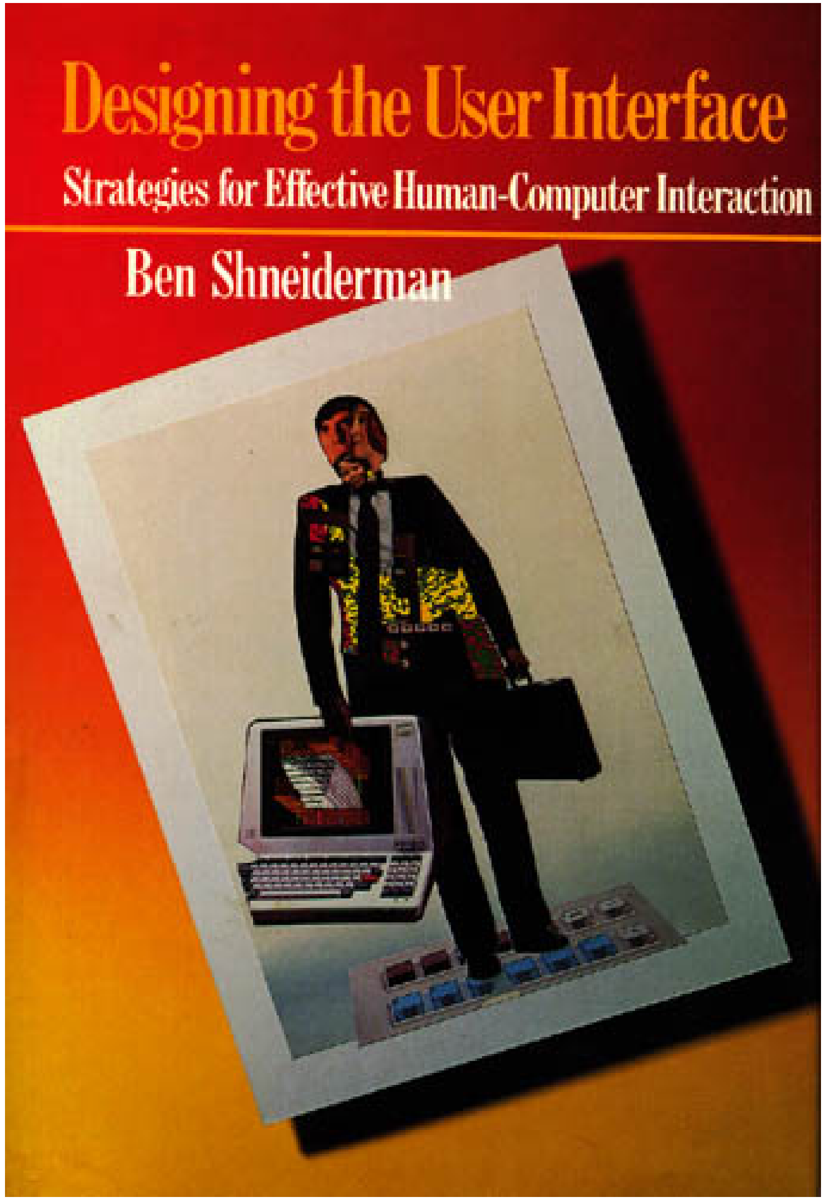
Designing the User Interface

Balancing automation & human control



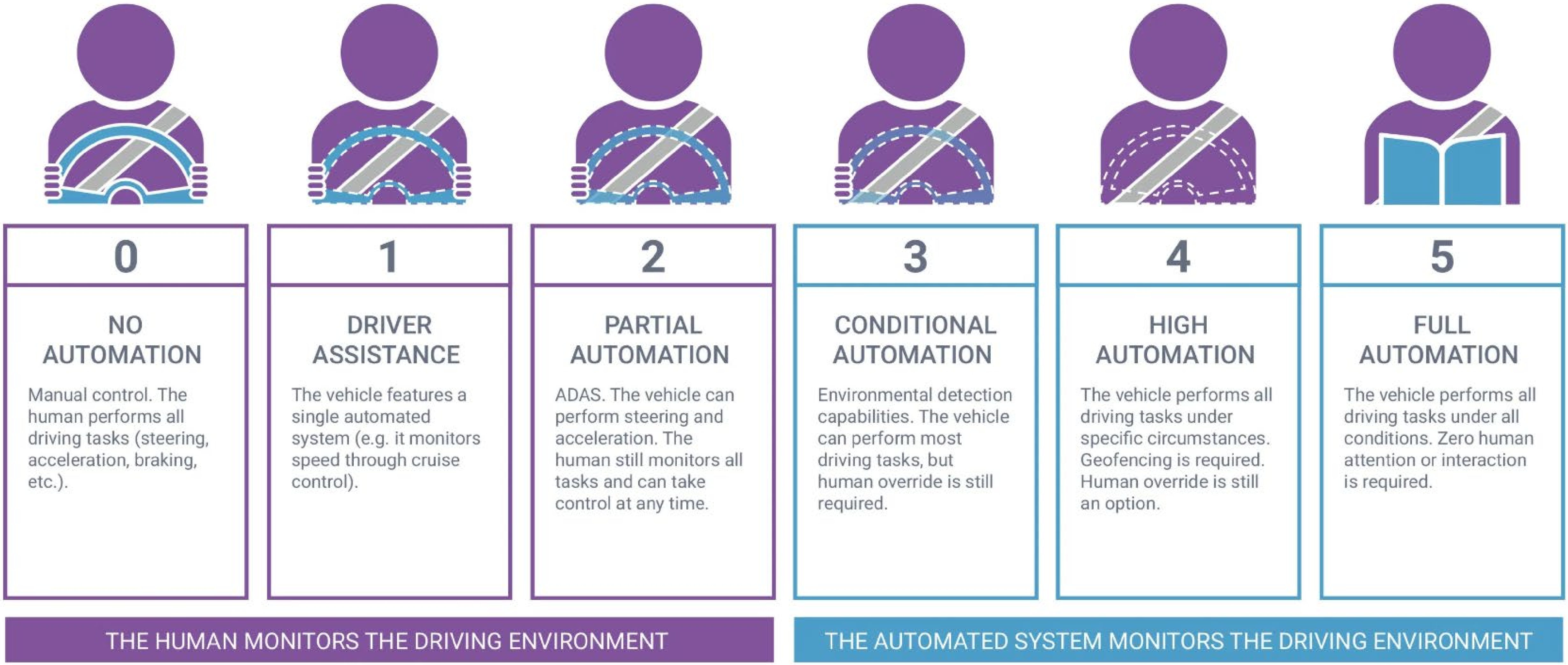
Human control

Computer automation


Designing the User Interface
Strategies for Effective Human-Computer Interaction
Ben Shneiderman

First Edition: 1986

LEVELS OF DRIVING AUTOMATION



0

NO AUTOMATION

Manual control. The human performs all driving tasks (steering, acceleration, braking, etc.).

1

DRIVER ASSISTANCE

The vehicle features a single automated system (e.g. it monitors speed through cruise control).

2

PARTIAL AUTOMATION

ADAS. The vehicle can perform steering and acceleration. The human still monitors all tasks and can take control at any time.

3

CONDITIONAL AUTOMATION

Environmental detection capabilities. The vehicle can perform most driving tasks, but human override is still required.

4

HIGH AUTOMATION

The vehicle performs all driving tasks under specific circumstances. Geofencing is required. Human override is still an option.

5

FULL AUTOMATION

The vehicle performs all driving tasks under all conditions. Zero human attention or interaction is required.

THE HUMAN MONITORS THE DRIVING ENVIRONMENT

THE AUTOMATED SYSTEM MONITORS THE DRIVING ENVIRONMENT

(Society of Automotive Engineers, 2016)

Designing the User Interface

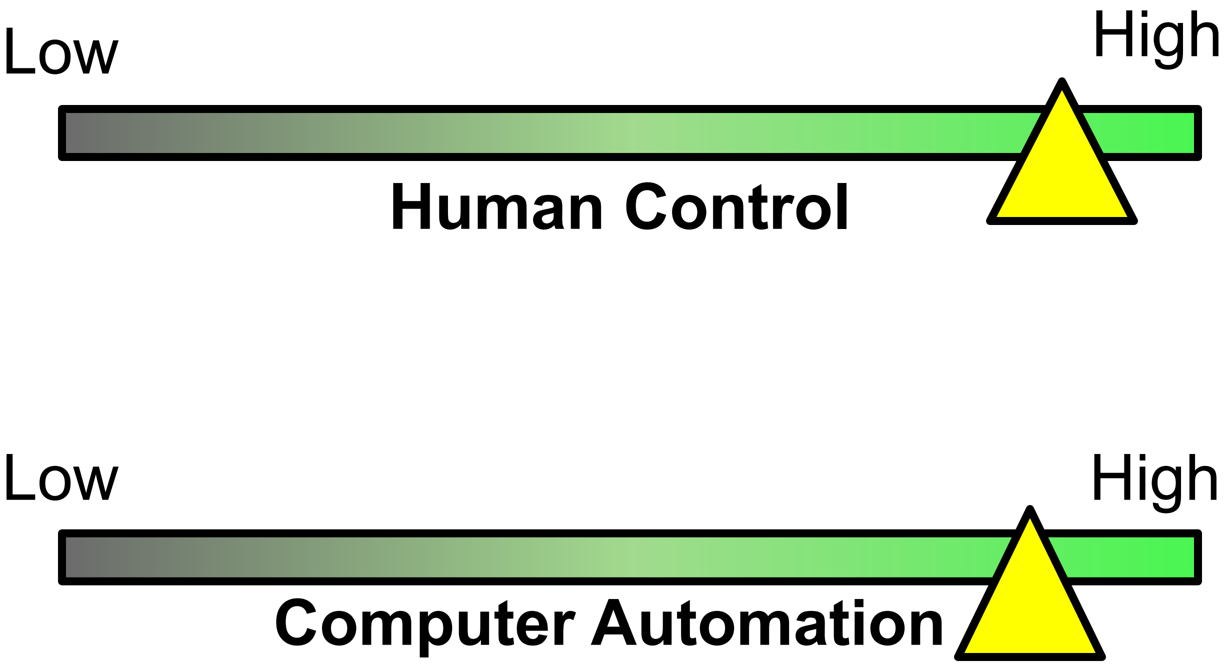
Ensuring human control
while increasing automation

SHNEIDERMAN • PLAISANT • COHEN • JACOBS • ELMQVIST
SIXTH EDITION
DESIGNING THE USER INTERFACE
STRATEGIES FOR EFFECTIVE
HUMAN-COMPUTER INTERACTION

Sixth Edition: 2016

Designing the User Interface

Ensuring human control
while increasing automation



Low

High

Human Control

Low

High

Computer Automation

SHNEIDERMAN • PLAISANT • COHEN • JACOBS • ELMQVIST

SIXTH EDITION

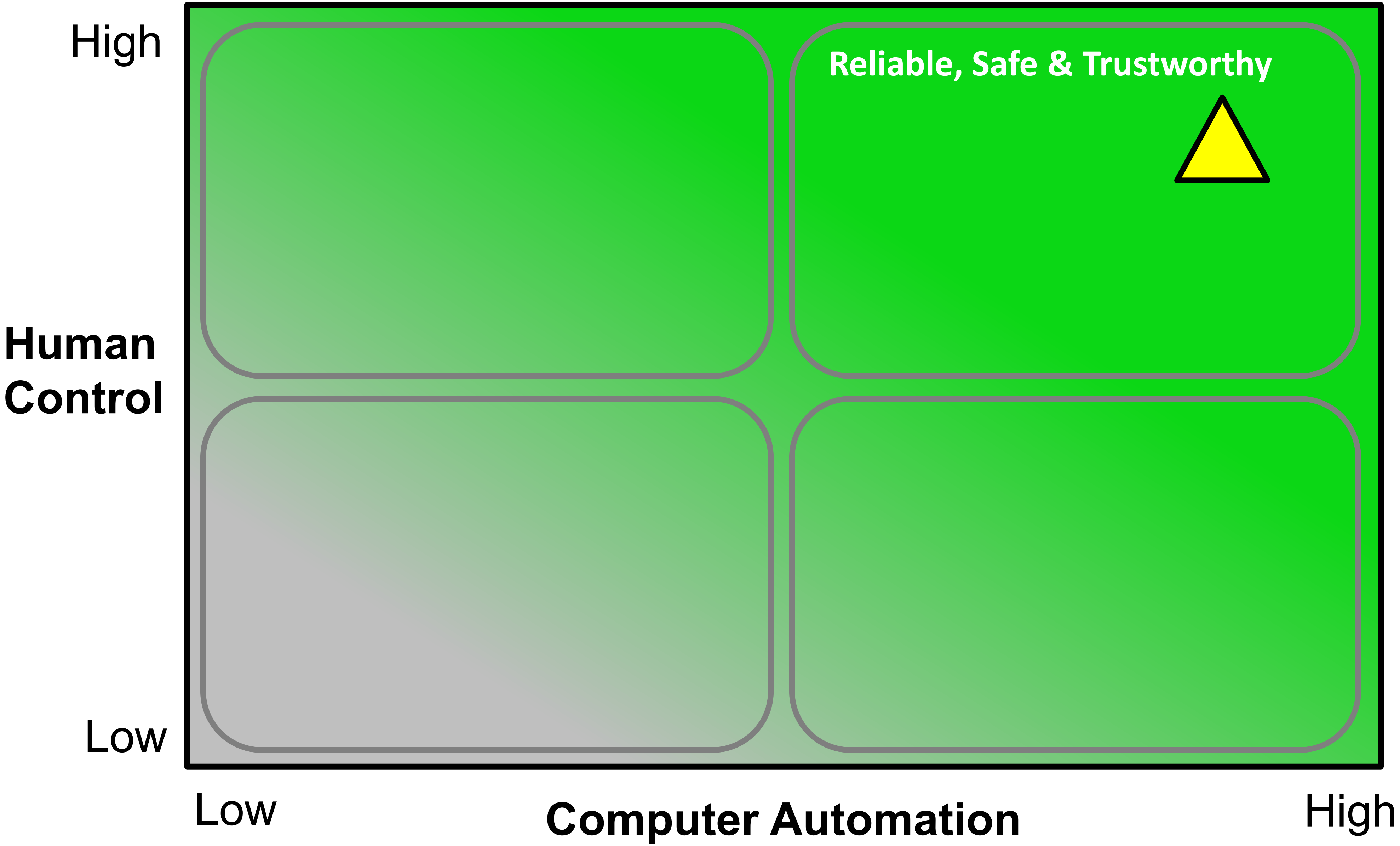
DESIGNING THE USER INTERFACE

STRATEGIES FOR EFFECTIVE
HUMAN-COMPUTER INTERACTION

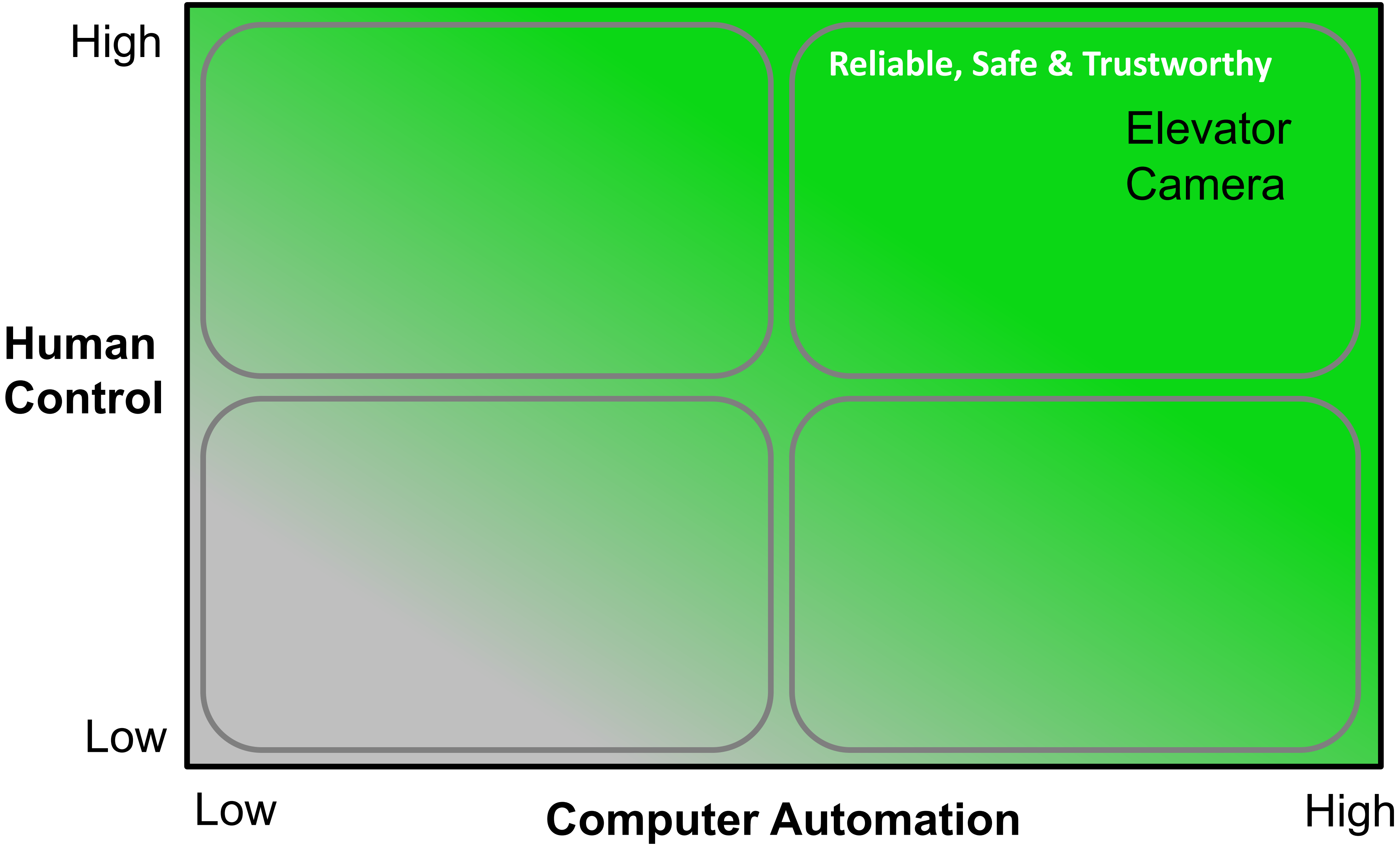
Sixth Edition: 2016

IJHCI (2020) `https://doi.org/10.1080/10447318.2020.1741118`

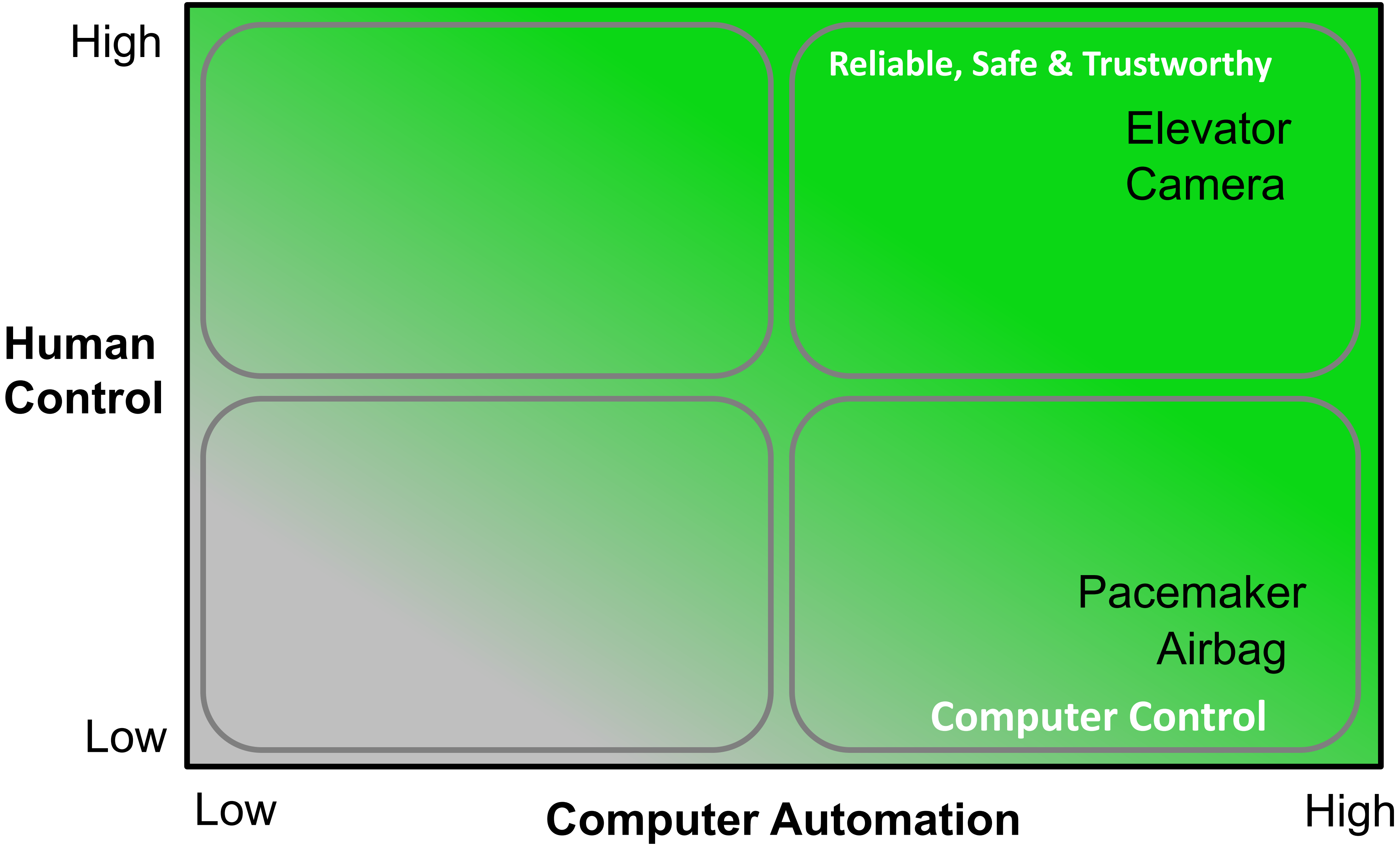
Human-Centered AI



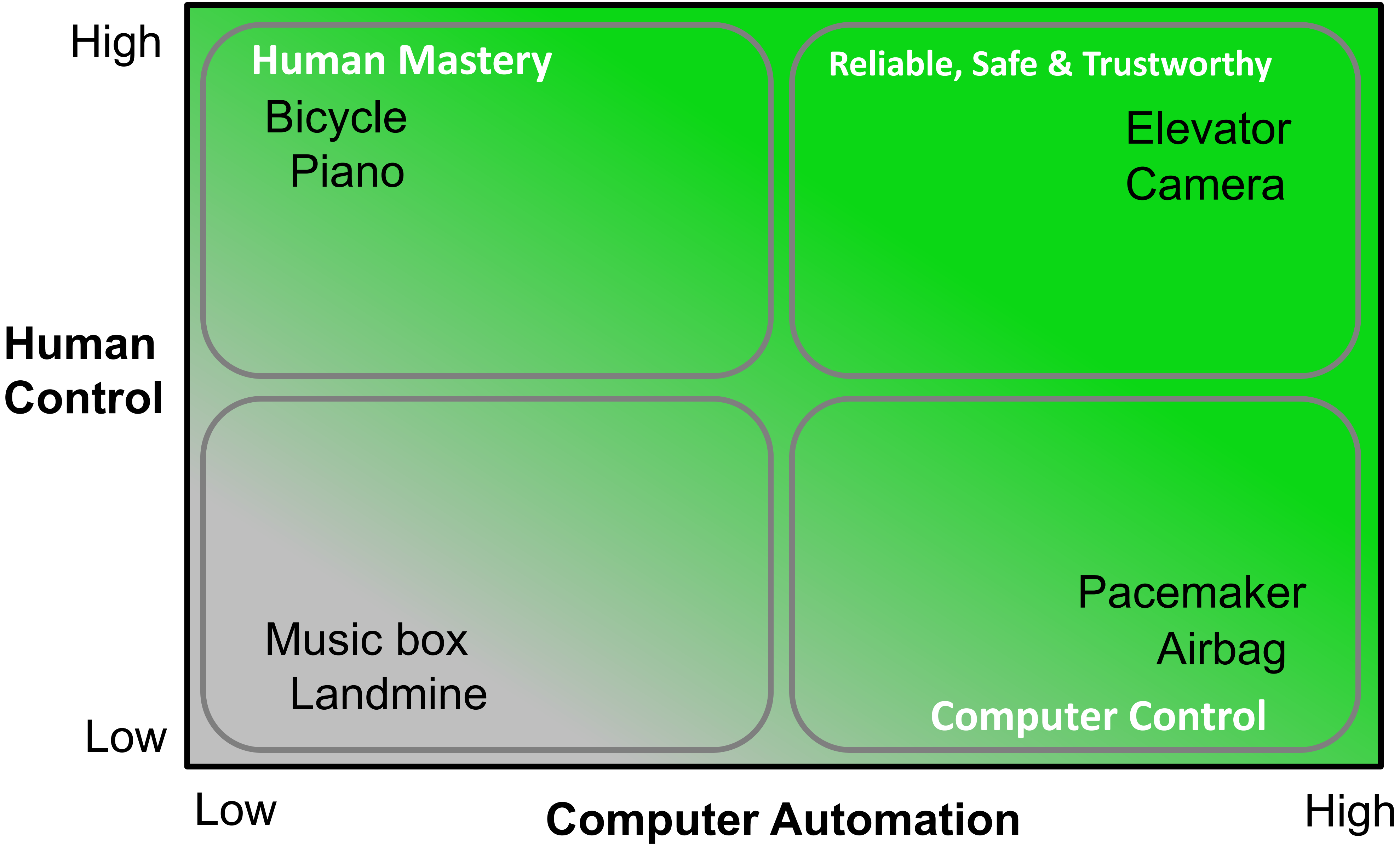
Human-Centered AI



Human-Centered AI



Human-Centered AI



Human Mastery

Bicycle
Piano

Reliable, Safe & Trustworthy

Elevator
Camera

Music box
Landmine

Pacemaker
Airbag

Computer Control

High

Human Control

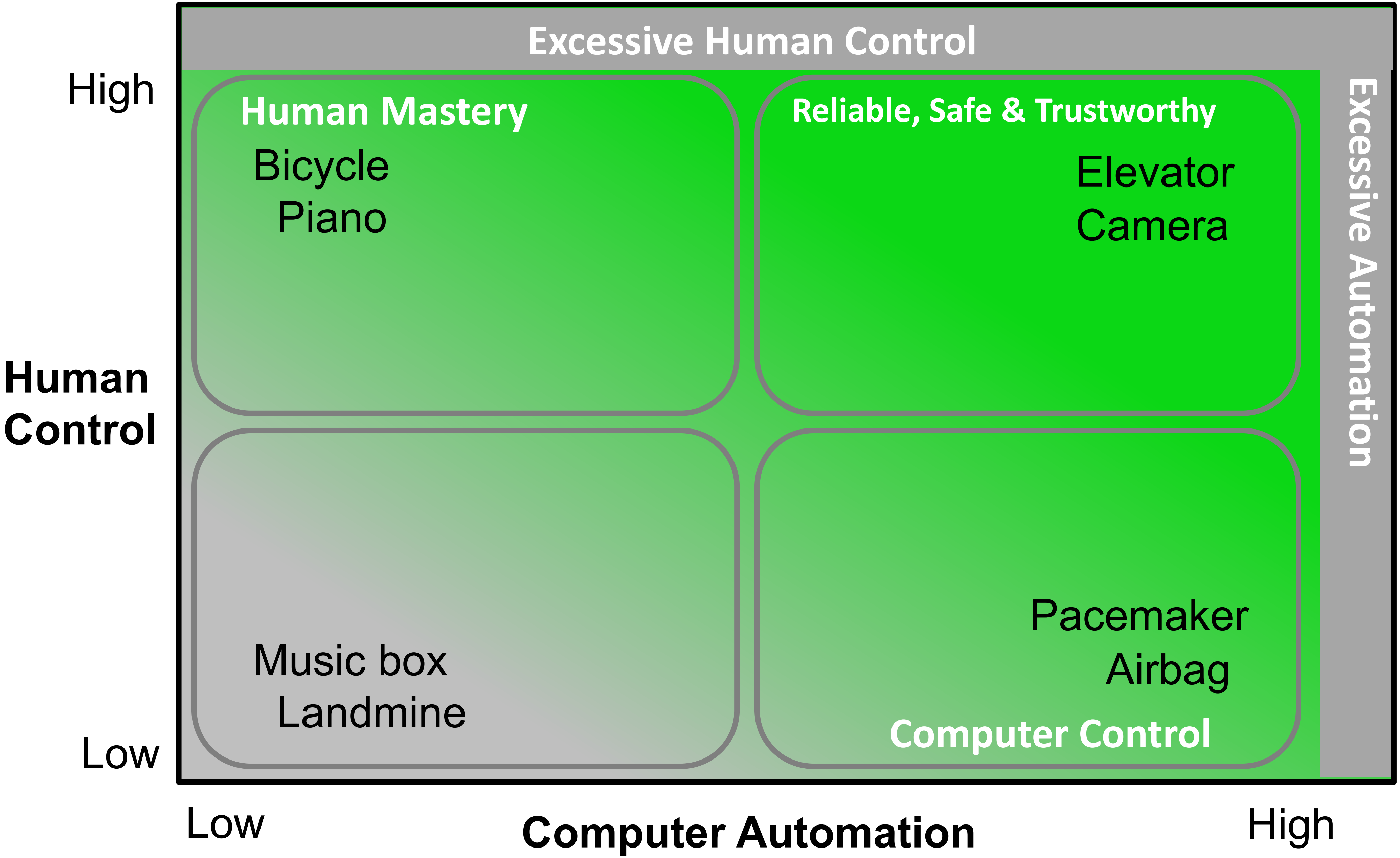
Low

Low

Computer Automation

High

Human-Centered AI



Excessive Human Control

High

Human Control

Low

Human Mastery

Bicycle

Piano

Reliable, Safe & Trustworthy

Elevator

Camera

Music box

Landmine

Pacemaker

Airbag

Computer Control

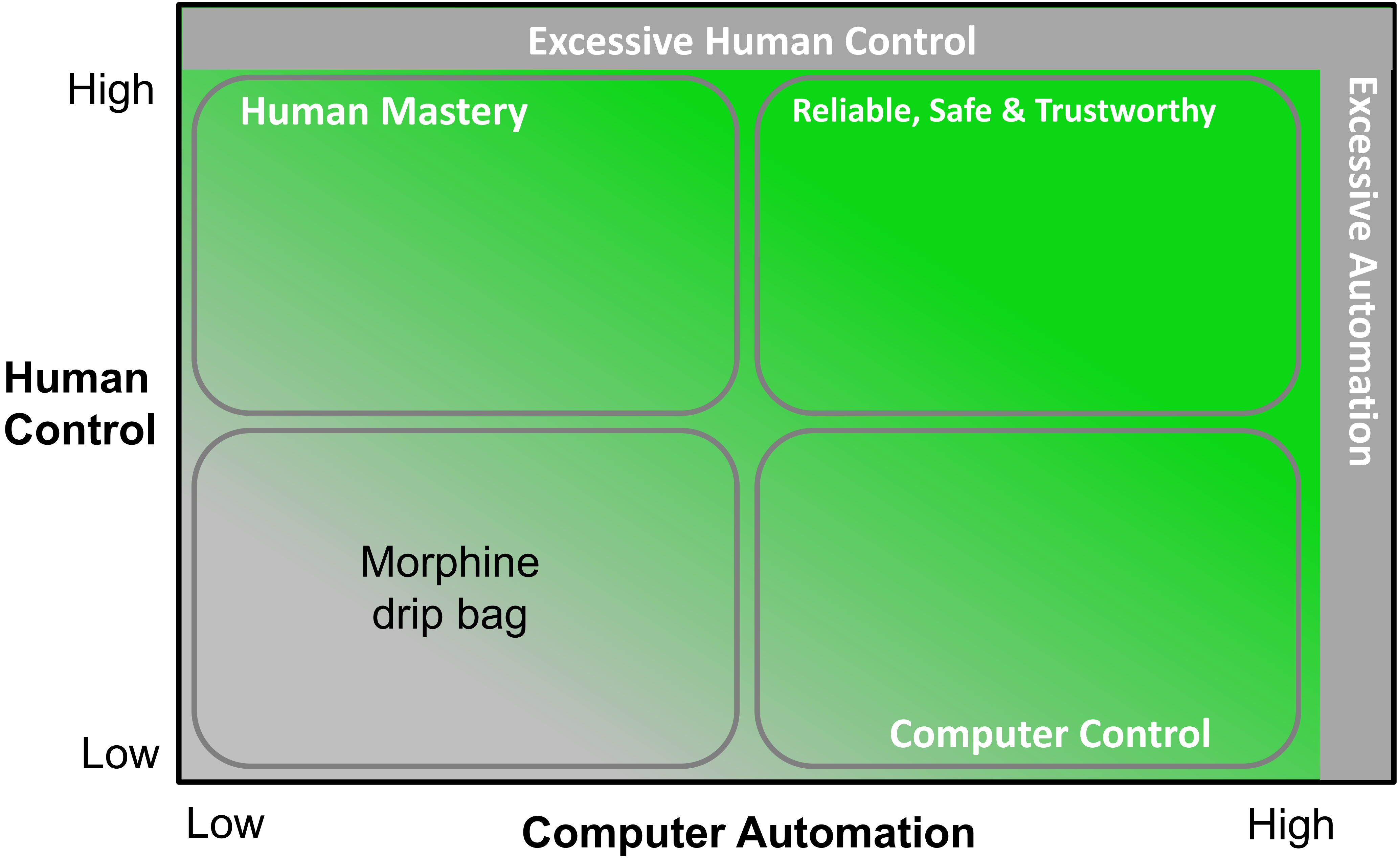
Excessive Automation

Low

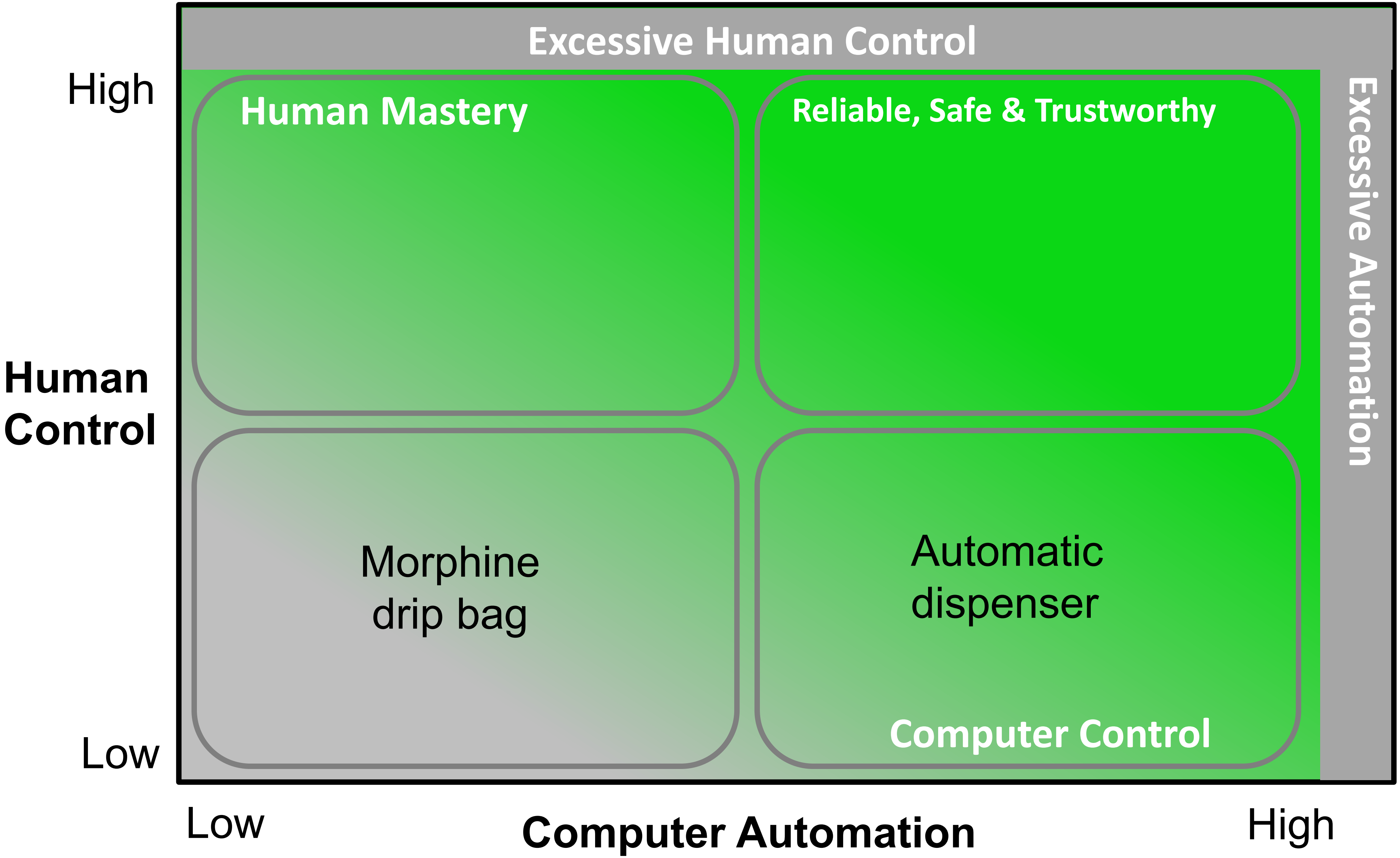
Computer Automation

High

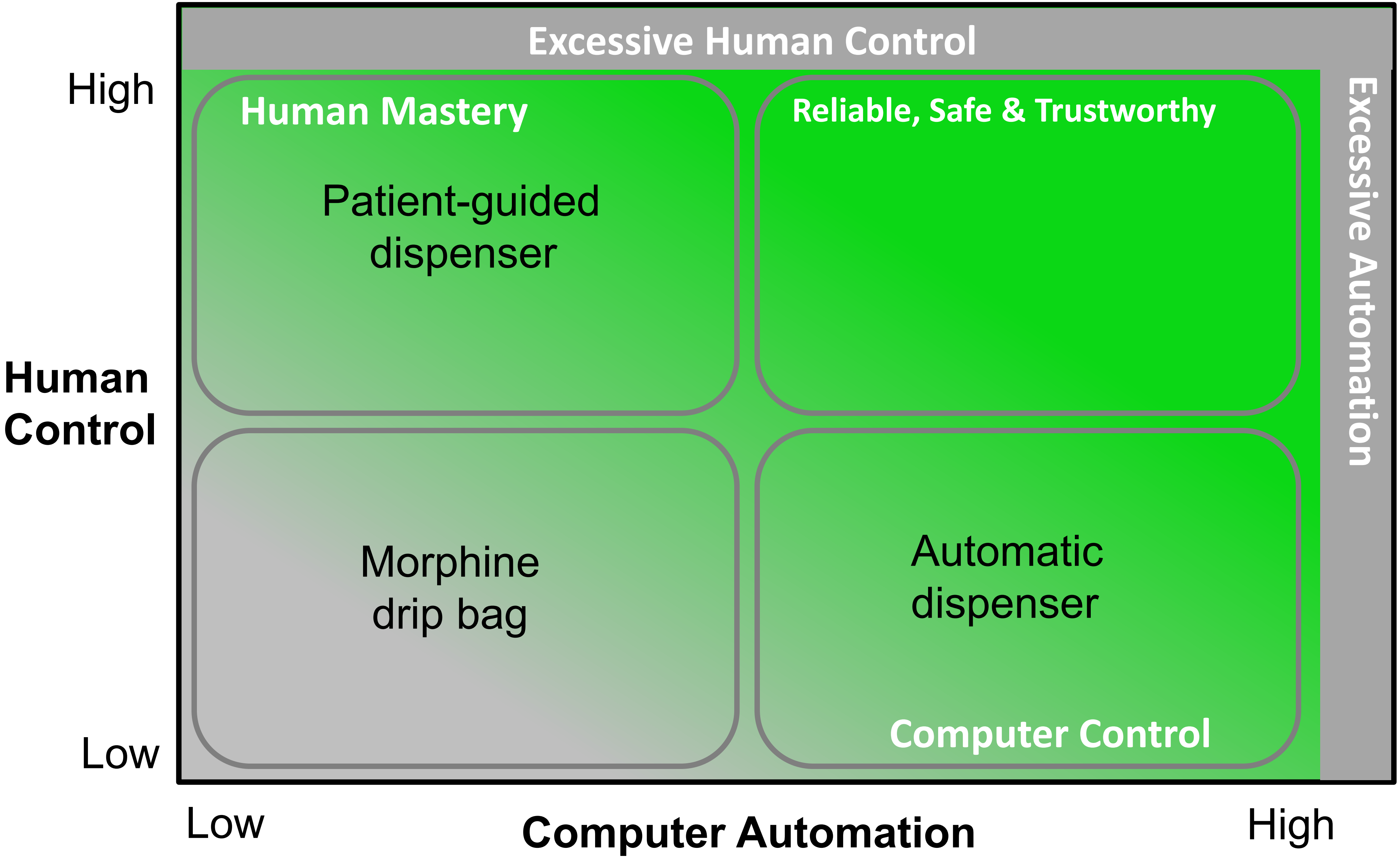
Pain Control Designs



Pain Control Designs



Pain Control Designs



Excessive Human Control

Human Mastery

Patient-guided dispenser

Reliable, Safe & Trustworthy

Morphine drip bag

Automatic dispenser

Computer Control

Excessive Automation

High

Human Control

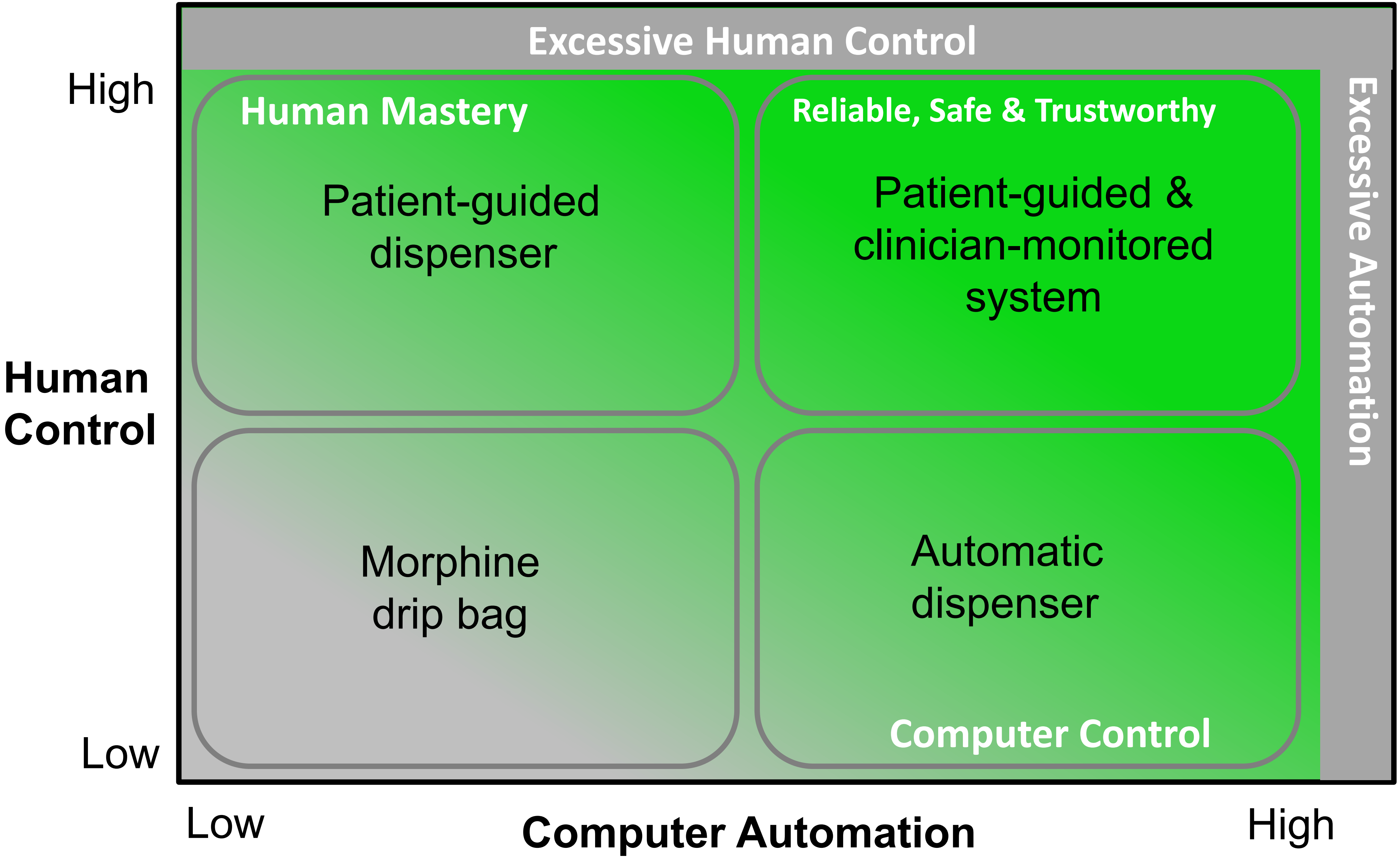
Low

Low

Computer Automation

High

Pain Control Designs



Excessive Human Control

Human Mastery

Patient-guided dispenser

Reliable, Safe & Trustworthy

Patient-guided & clinician-monitored system

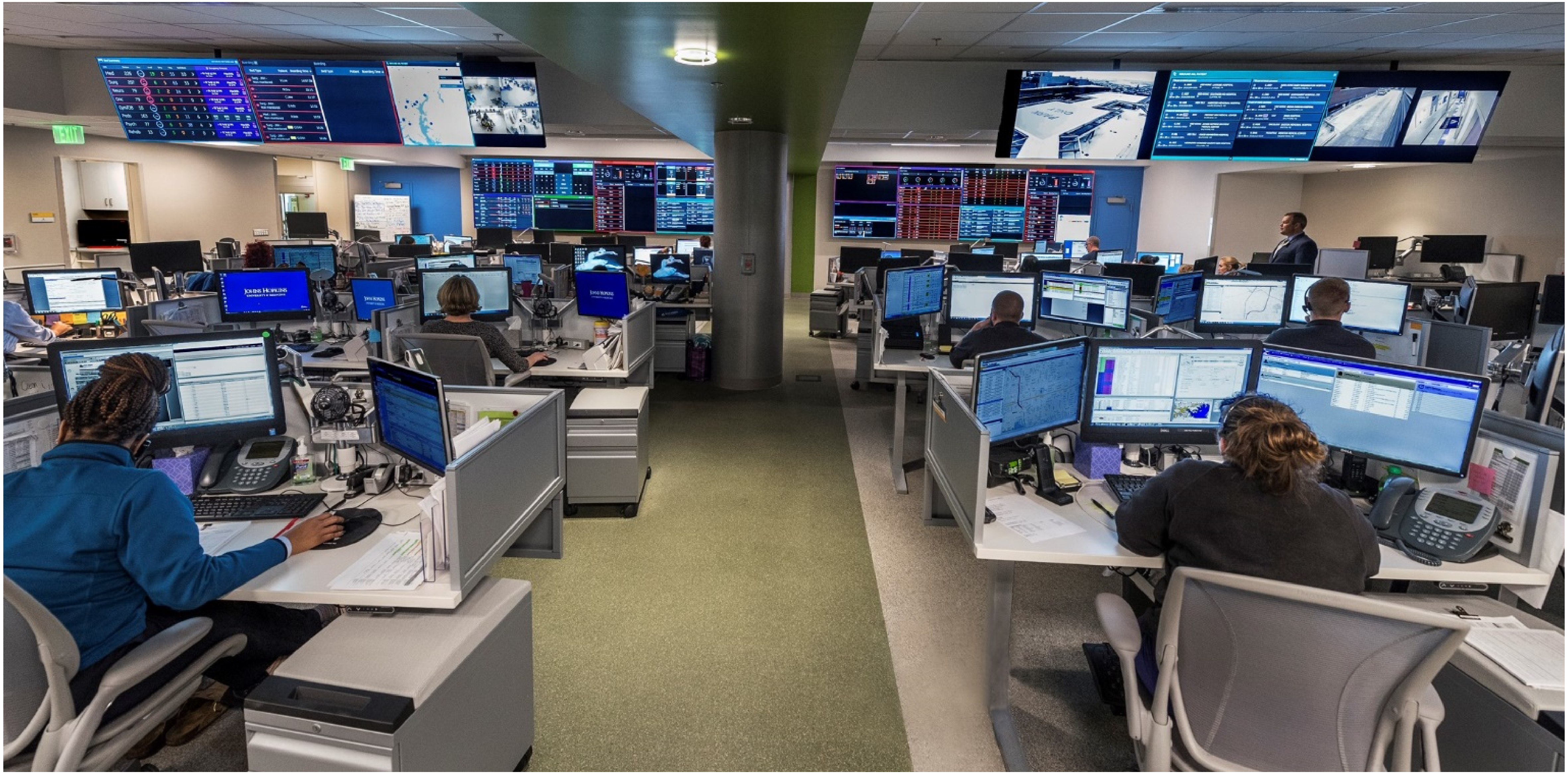
Morphine drip bag

Automatic dispenser

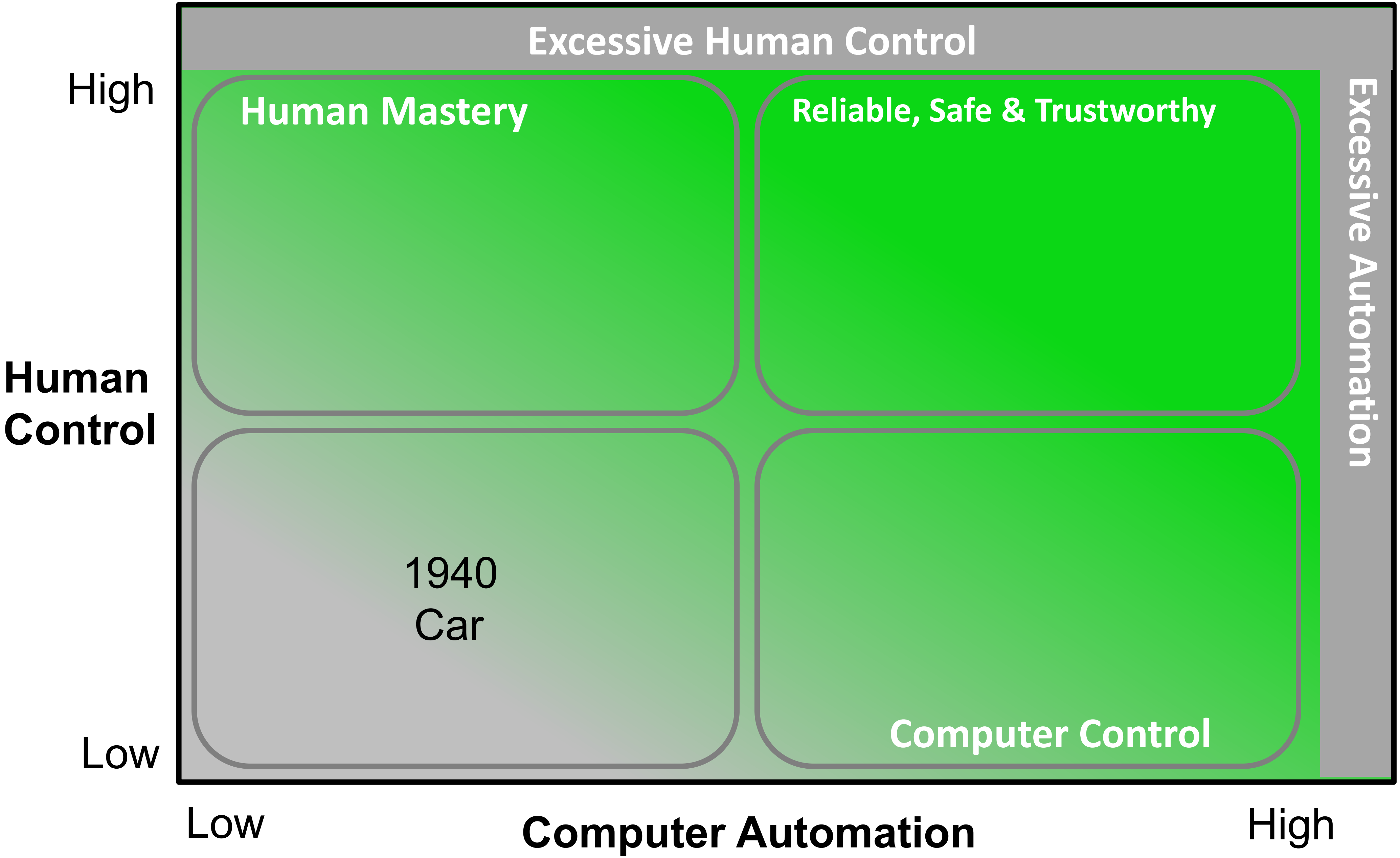
Computer Control

Excessive Automation

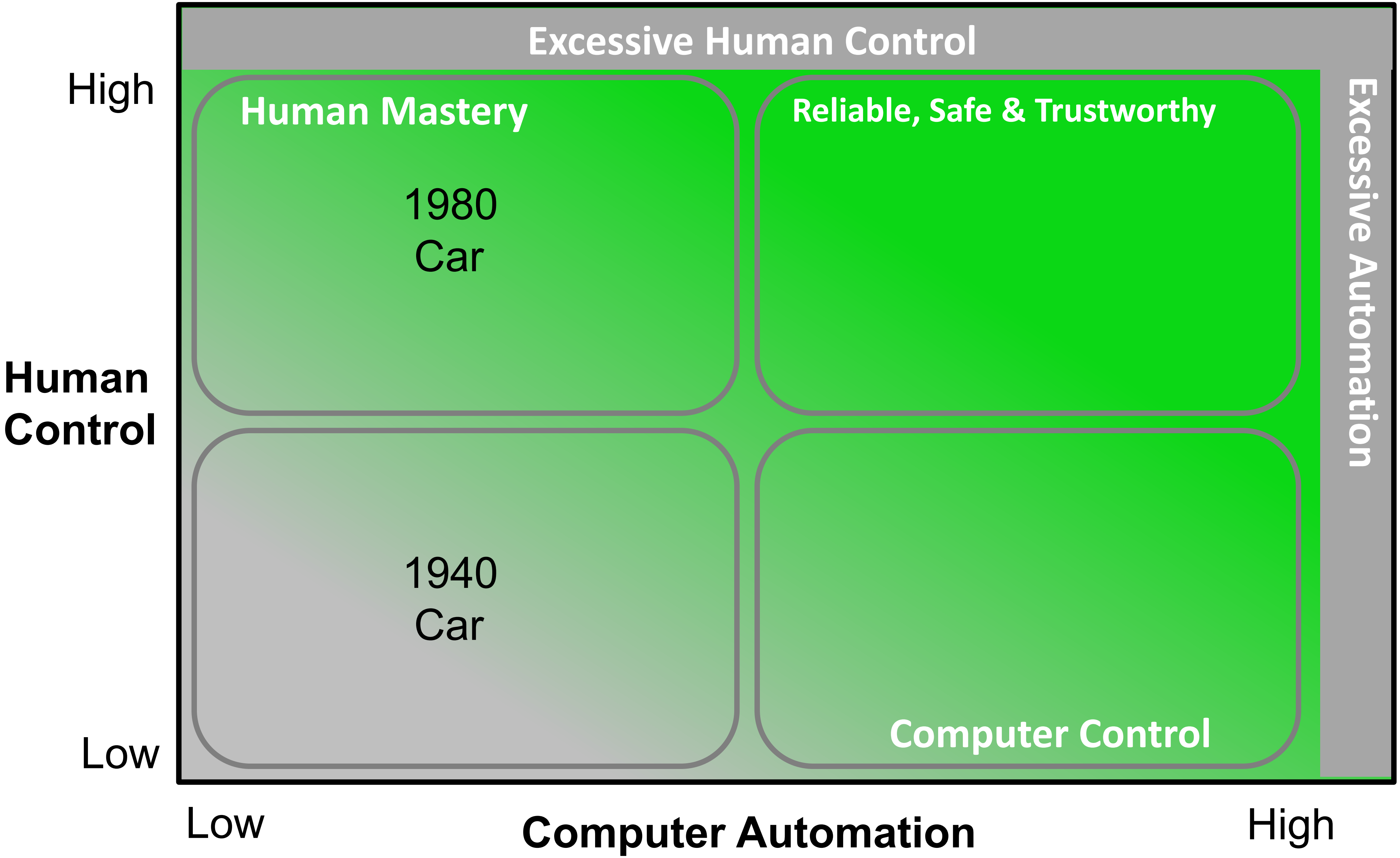
Johns Hopkins University Hospital Control Center



Car Control Designs



Car Control Designs



Excessive Human Control

Excessive Automation

Human Mastery

1980
Car

Reliable, Safe & Trustworthy

1940
Car

Computer Control

High

Human Control

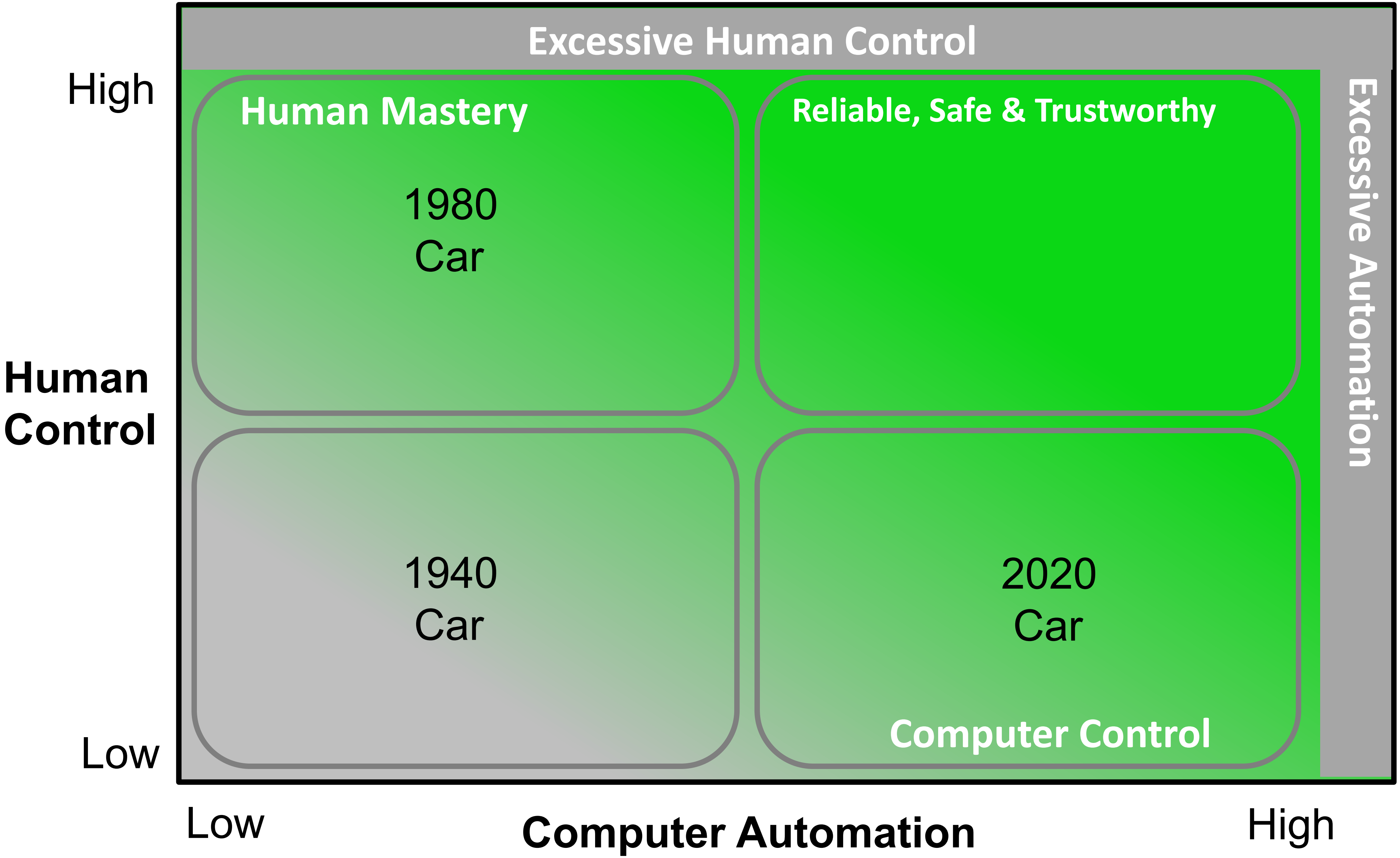
Low

Low

Computer Automation

High

Car Control Designs



Excessive Human Control

Excessive Automation

Human Mastery

1980 Car

Reliable, Safe & Trustworthy

1940 Car

2020 Car

Computer Control

High

Human Control

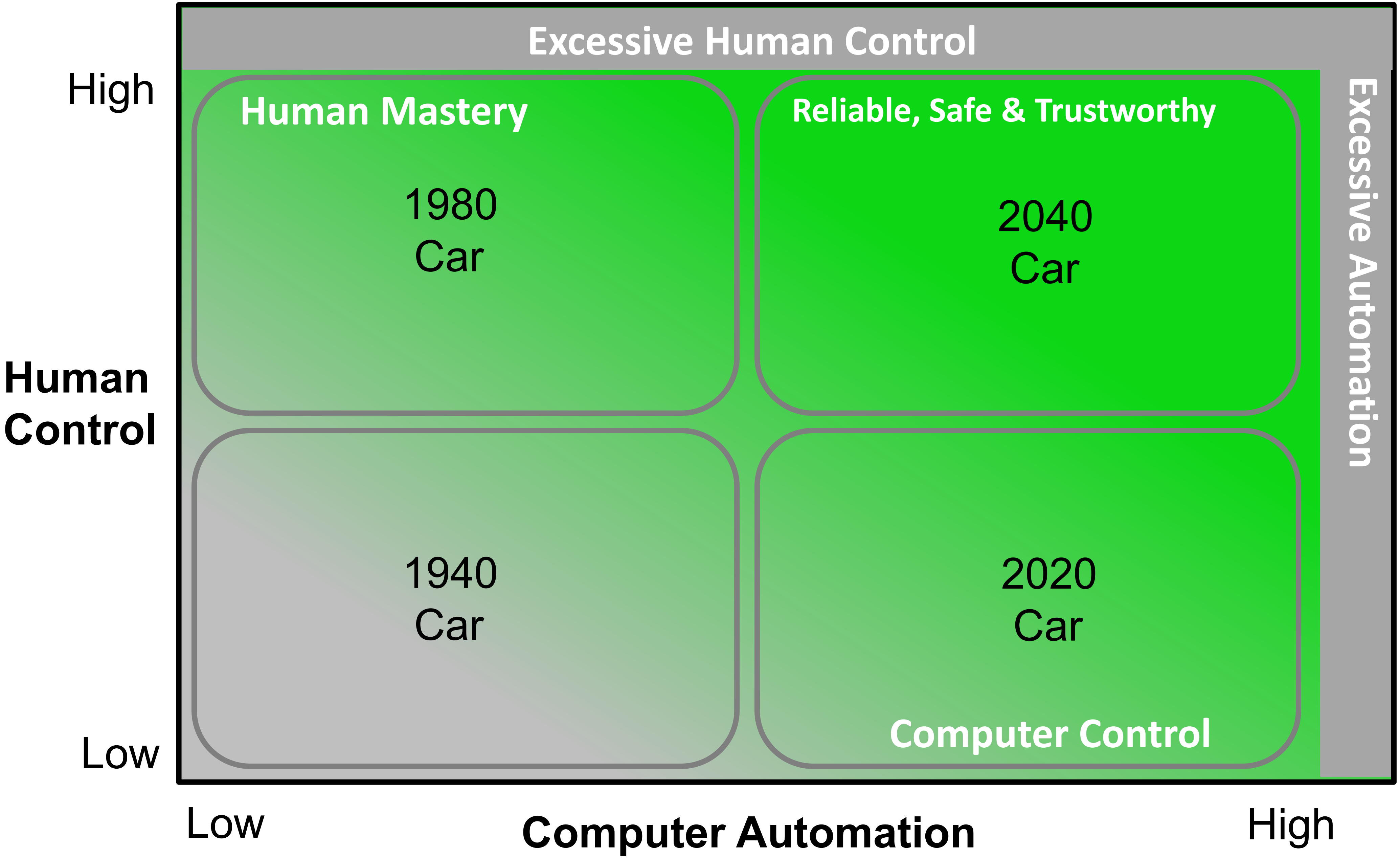
Low

Low

Computer Automation

High

Car Control Designs



Excessive Human Control

Excessive Automation

Human Mastery

1980 Car

Reliable, Safe & Trustworthy

2040 Car

1940 Car

2020 Car

Computer Control

Human Control

High

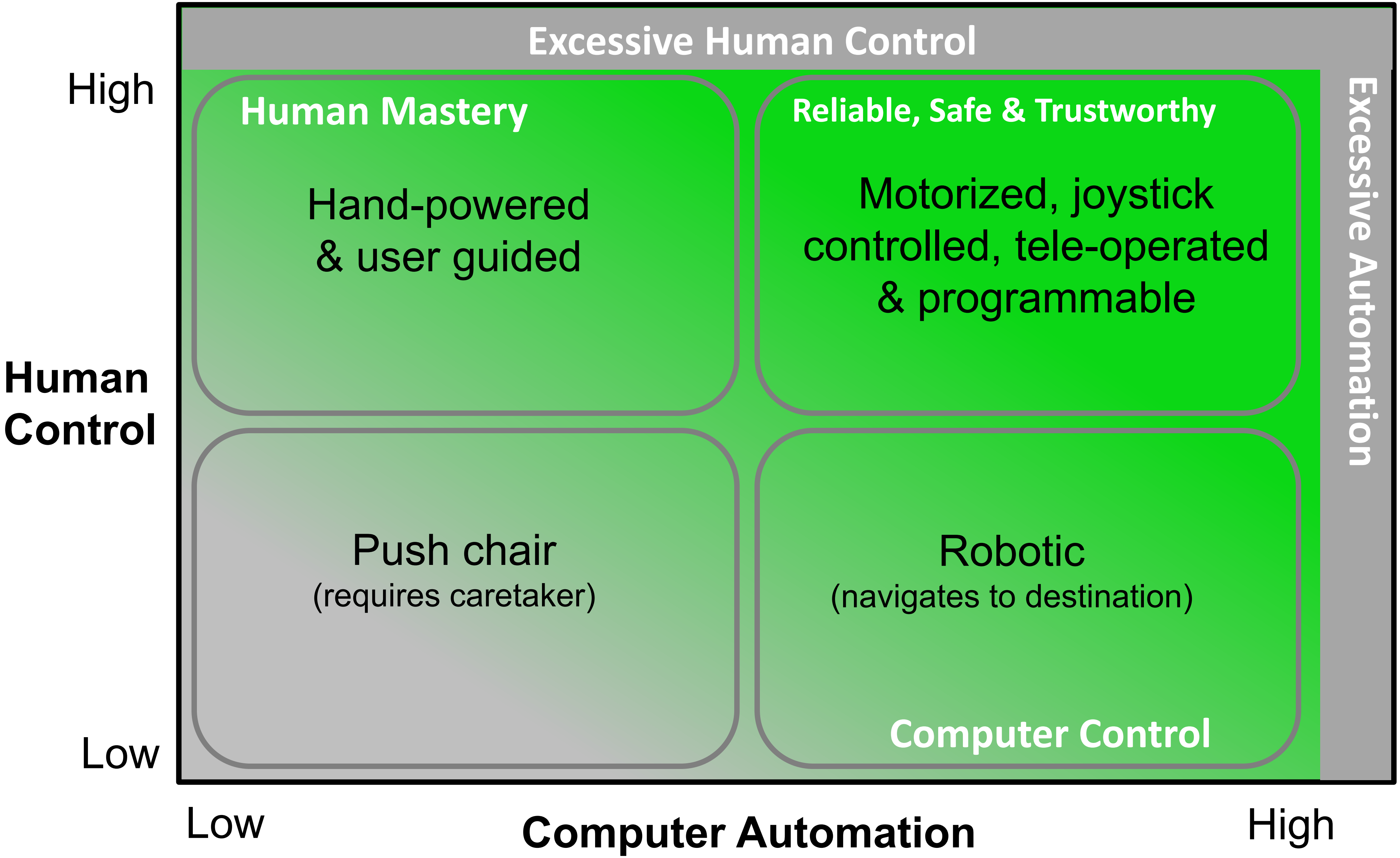
Low

Computer Automation

Low

High

Wheelchair Designs



Excessive Human Control

Excessive Automation

High

Human Control

Low

Human Mastery

Hand-powered & user guided

Reliable, Safe & Trustworthy

Motorized, joystick controlled, tele-operated & programmable

Push chair (requires caretaker)

Robotic (navigates to destination)

Computer Control

Low

Computer Automation

High

Application Categories

- **Recommenders**
- **Consequential**
- **Life-critical**

Application Categories

- Recommenders
- Consequential
- Life-critical

- Rapid performance
- Long duration
- Remote locations

Micro-Structure of Design

- **Automate where**
  - **+ Reliable performance is possible**
  - **- But allow overrides**

Micro-Structure of Design

- **Automate where**
  - **+ Reliable performance is possible**
  - **- But allow overrides**
- **Give human control where**
  - **+ Desired for creative flexibility**
  - **+ Automation is uncertain**
  - - **But prevent human errors**

Micro-Structure of Design

- **Automate where**

  + Reliable performance is possible

  - But allow overrides

- **Give human control where**

  + Desired for creative flexibility

  + Automation is uncertain

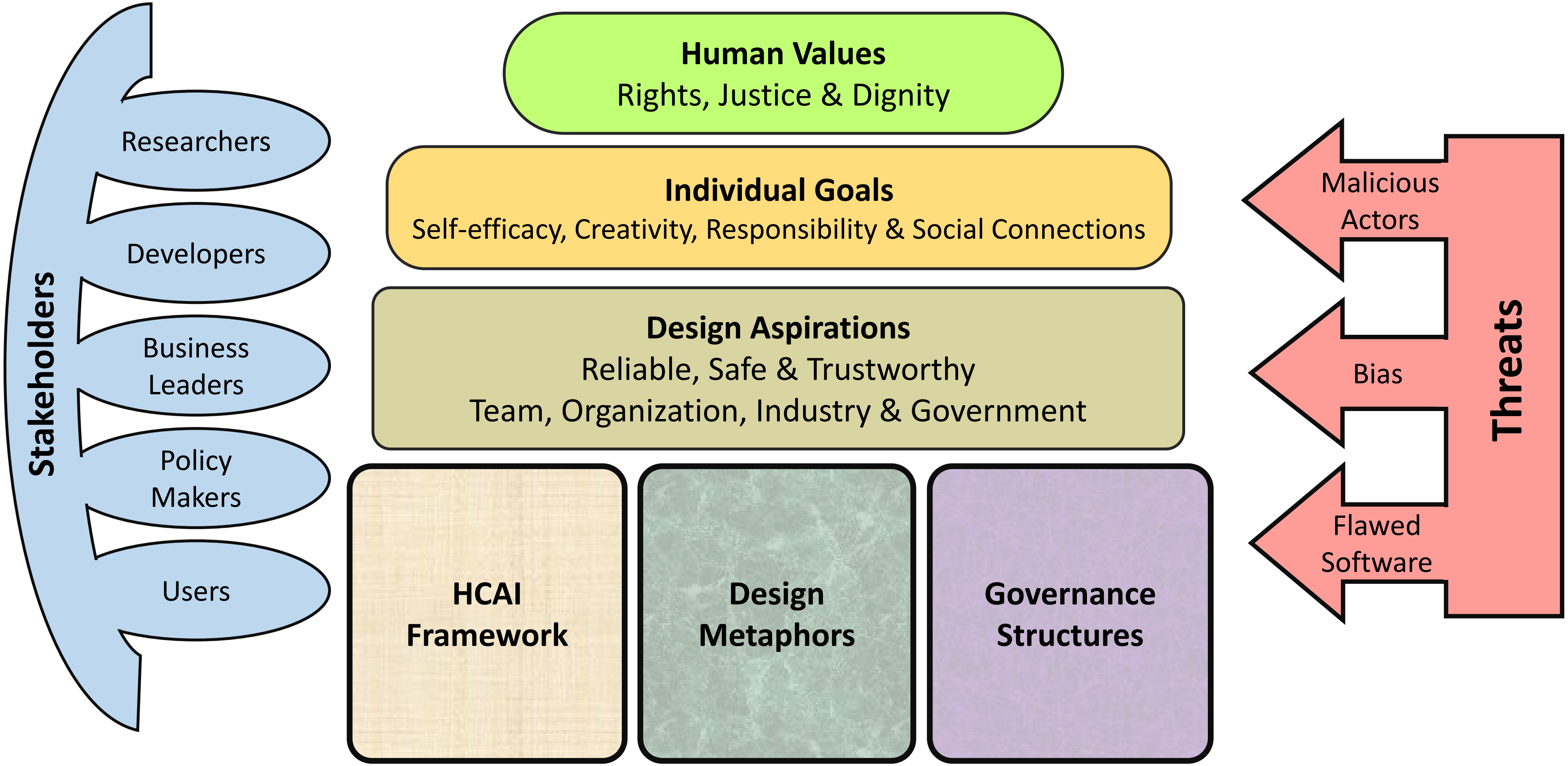
  - **But prevent human errors**

- **Design supervisory control**

  + Teleoperate remotely

  + Collect aggregate data

Human-Centered AI



Stakeholders

Researchers

Developers

Business Leaders

Policy Makers

Users

Human Values
Rights, Justice & Dignity

Individual Goals
Self-efficacy, Creativity, Responsibility & Social Connections

Design Aspirations
Reliable, Safe & Trustworthy
Team, Organization, Industry & Government

HCAI Framework

Design Metaphors

Governance Structures

Threats

Malicious Actors

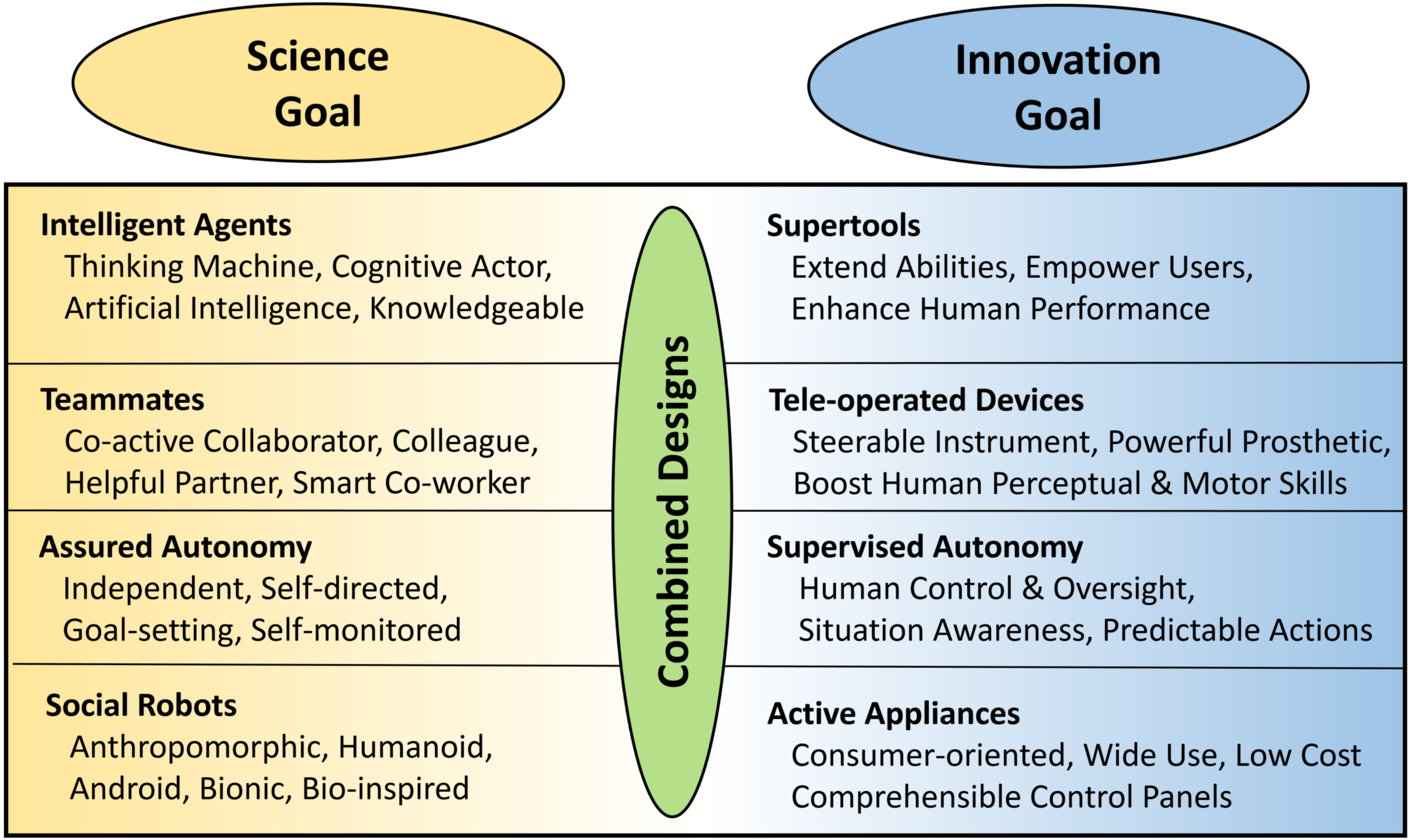
Bias

Flawed Software

Oxford University Press (Early 2022) `https://hcil.umd.edu/human-centered-ai/`

Design Metaphors

Design Metaphors



Science Goal

Innovation Goal

Combined Designs

Science Goal	Innovation Goal
Intelligent Agents Thinking Machine, Cognitive Actor, Artificial Intelligence, Knowledgeable	**Supertools** Extend Abilities, Empower Users, Enhance Human Performance
Teammates Co-active Collaborator, Colleague, Helpful Partner, Smart Co-worker	**Tele-operated Devices** Steerable Instrument, Powerful Prosthetic, Boost Human Perceptual & Motor Skills
Assured Autonomy Independent, Self-directed, Goal-setting, Self-monitored	**Supervised Autonomy** Human Control & Oversight, Situation Awareness, Predictable Actions
Social Robots Anthropomorphic, Humanoid, Android, Bionic, Bio-inspired	**Active Appliances** Consumer-oriented, Wide Use, Low Cost Comprehensible Control Panels

Teammate & Supertool

- **Social Teammate:**
  Since many people respond socially to robots,
  -> design robots to be human-like social teammates.

Teammate & Supertool

- **Social Teammate:**
  Since many people respond socially to robots,
  -> design robots to be human-like social teammates.

- **Human-Centered Conjecture:**
  Since only humans can be responsible &
  computers have distinct capabilities (speed, storage, display...)
  -> design computers to be supertools

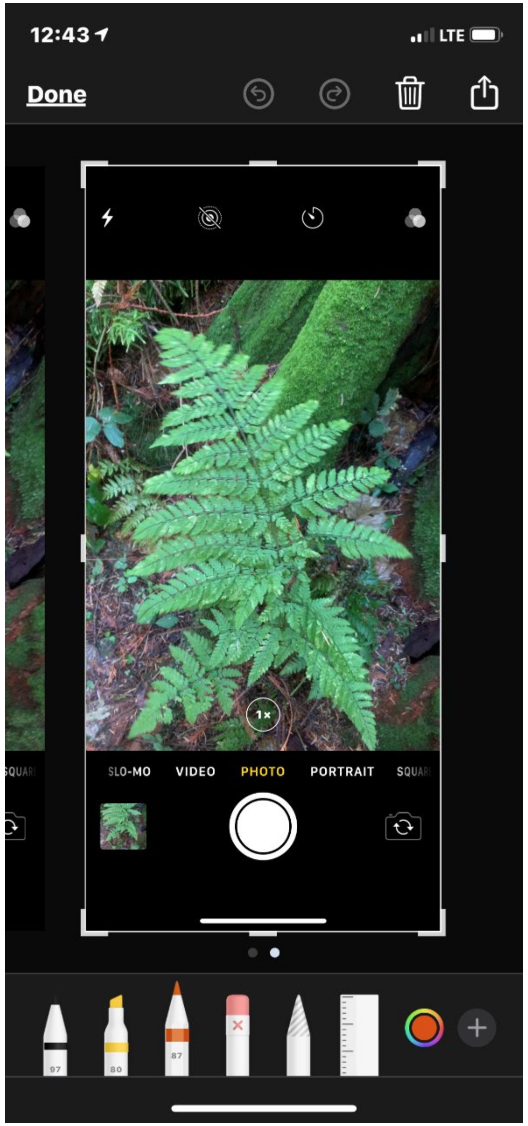
Teammate & Supertool

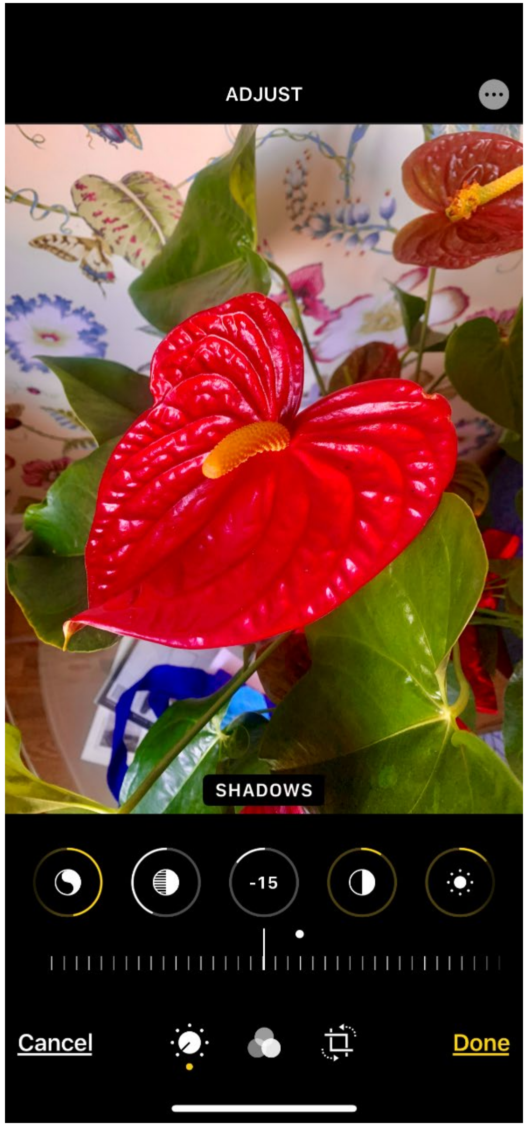
- **Social Teammate:**
  Since many people respond socially to robots,
  -> design robots to be human-like social teammates.

- **Human-Centered Conjecture:**
  Since only humans can be responsible &
  computers have distinct capabilities (speed, storage, display...)
  -> design computers to be supertools
  -> invite users to fix, personalize & extend the design

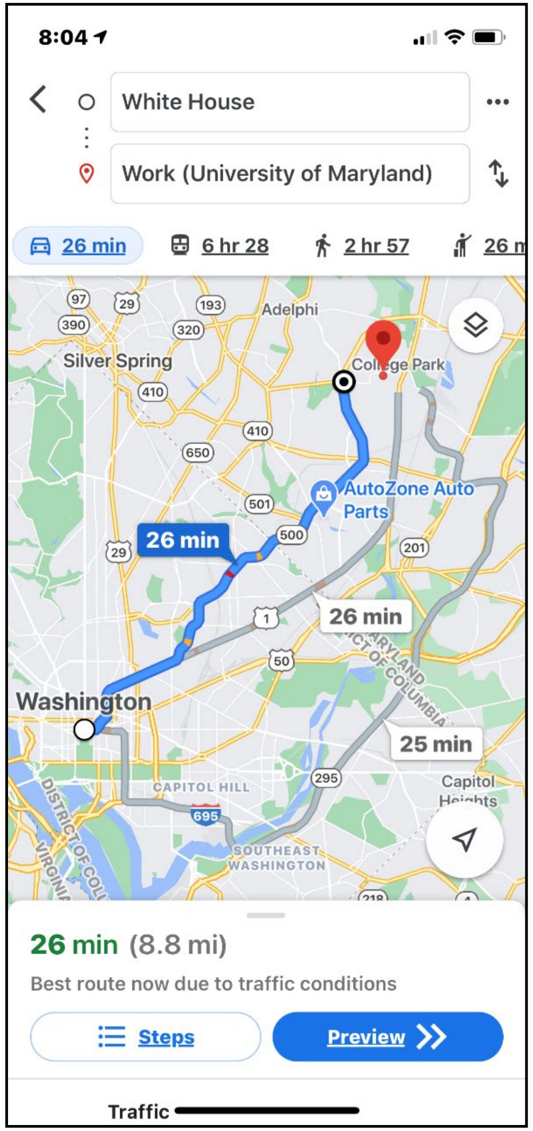
Supertools

Digital Camera Controls





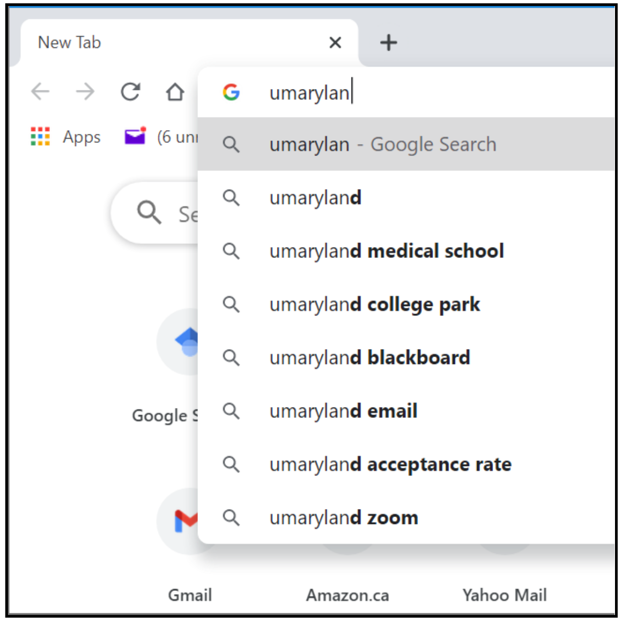
Navigation Choices



26 min (8.8 mi)

Best route now due to traffic conditions

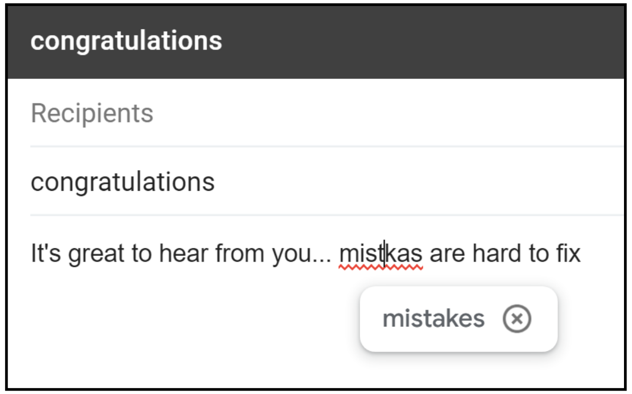
Texting Autocompletion



umarylan

- umarylan - Google Search
- umaryland
- umaryland medical school
- umaryland college park
- umaryland blackboard
- umaryland email
- umaryland acceptance rate
- umaryland zoom

Spelling correction



congratulations

It's great to hear from you... mistkas are hard to fix

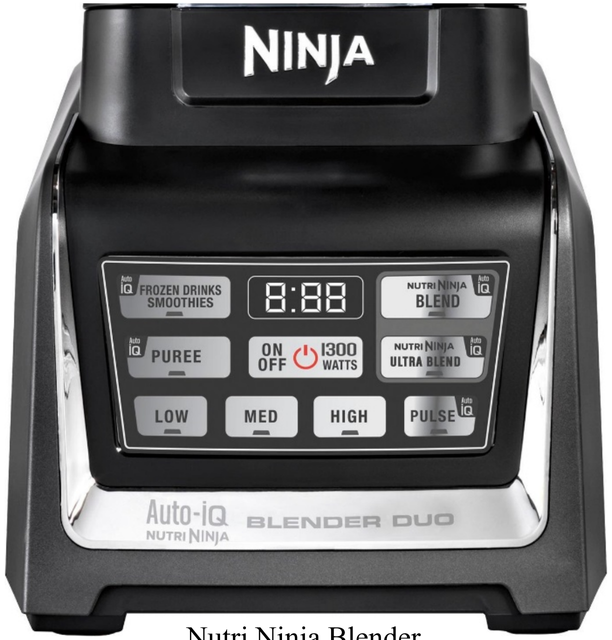
Active Appliances

Coffee maker, Rice cooker, Blender

Dishwasher, Clothes Washer/Dryer

Cuisinart Grind & Brew Coffee Maker

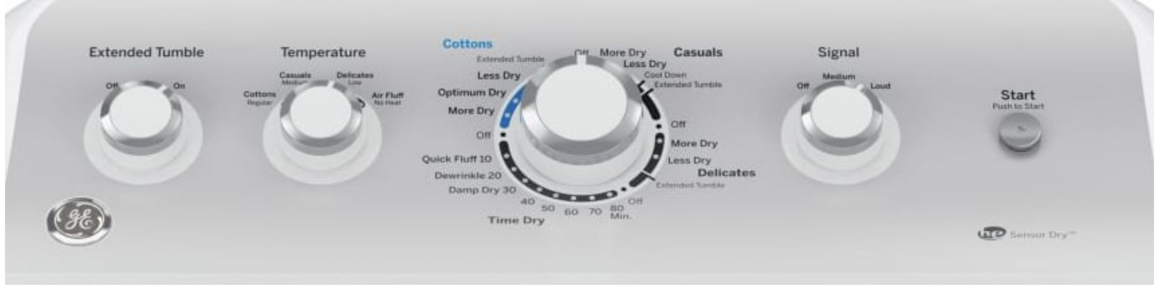
Panasonic Rice Cooker



Nutri Ninja Blender

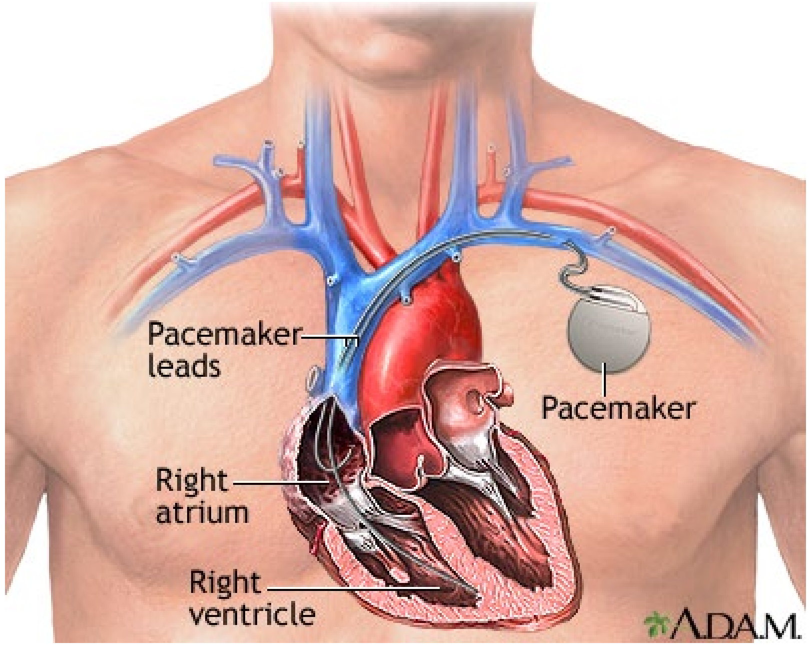
Miele Dishwasher

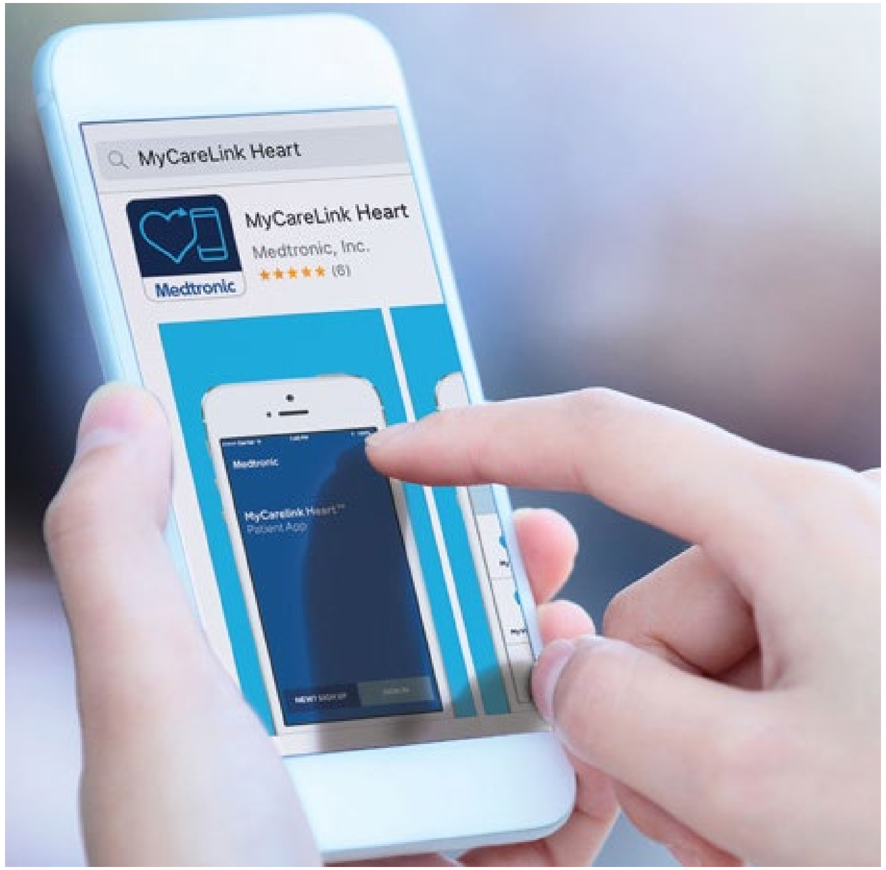
General Electric Washer

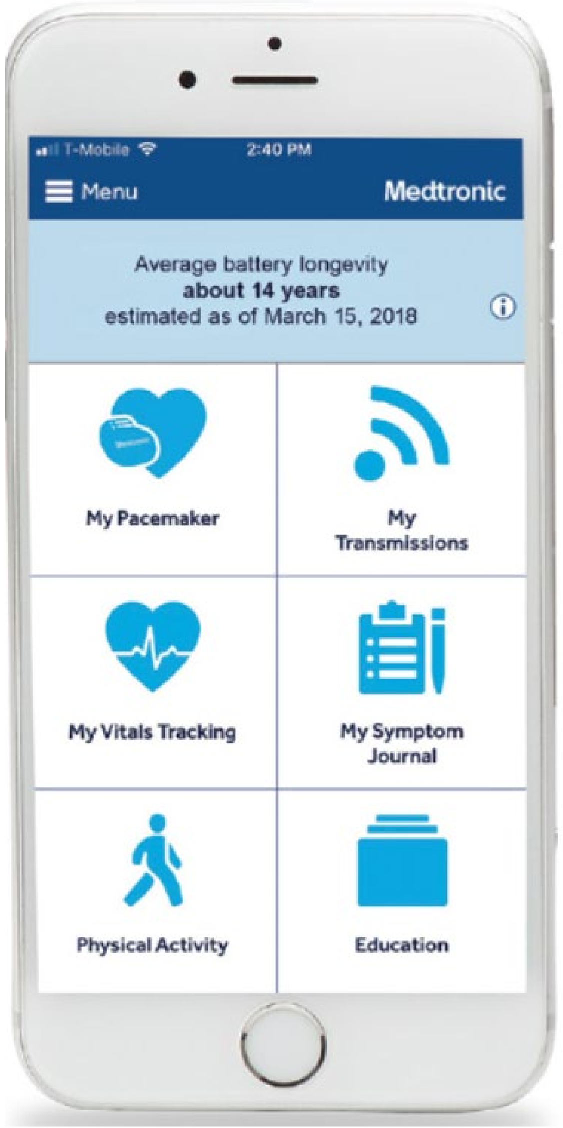


General Electric Dryer

Implanted Cardiac Pacemakers



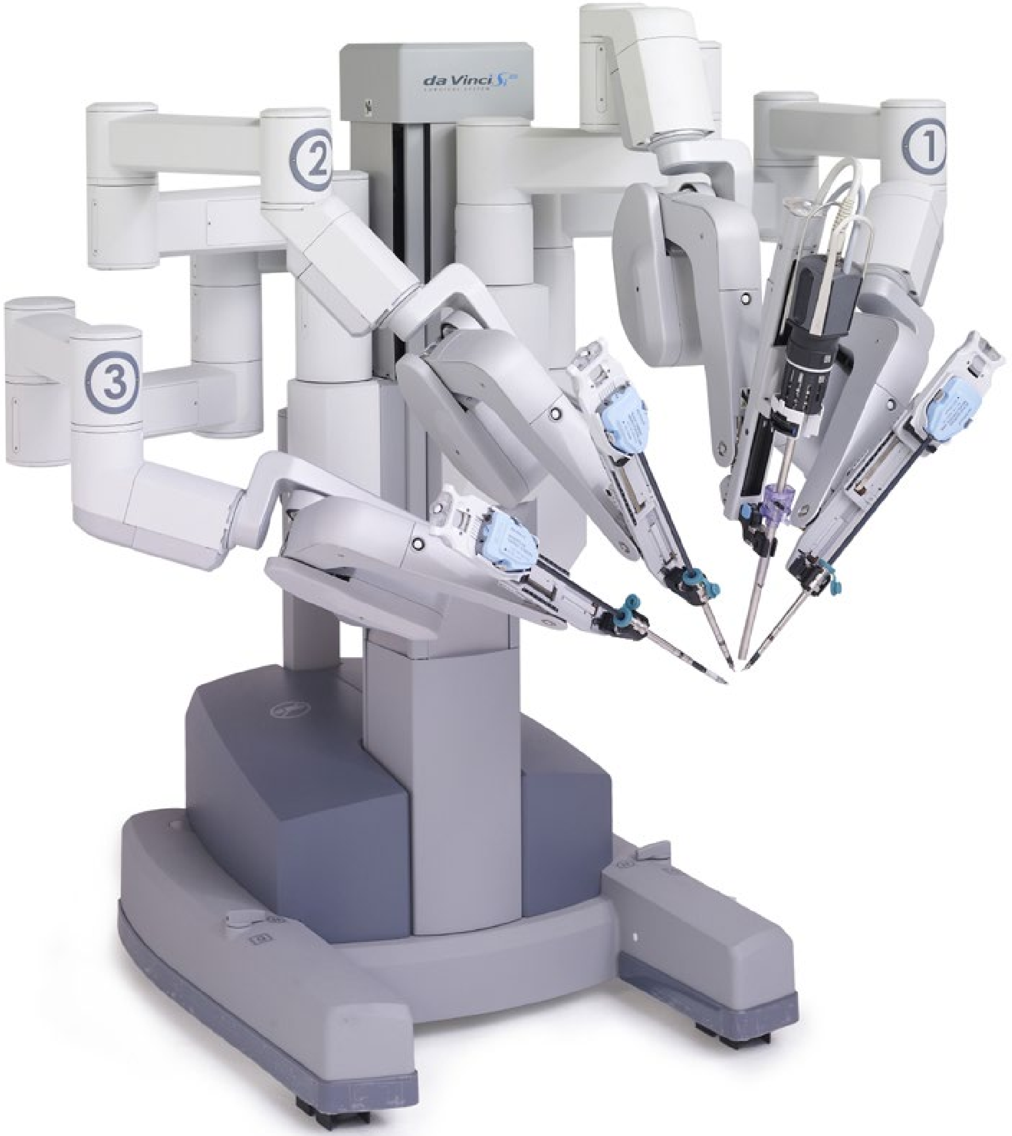




Average battery longevity
about 14 years
estimated as of March 15, 2018

NASA Mars Rovers are Tele-Operated

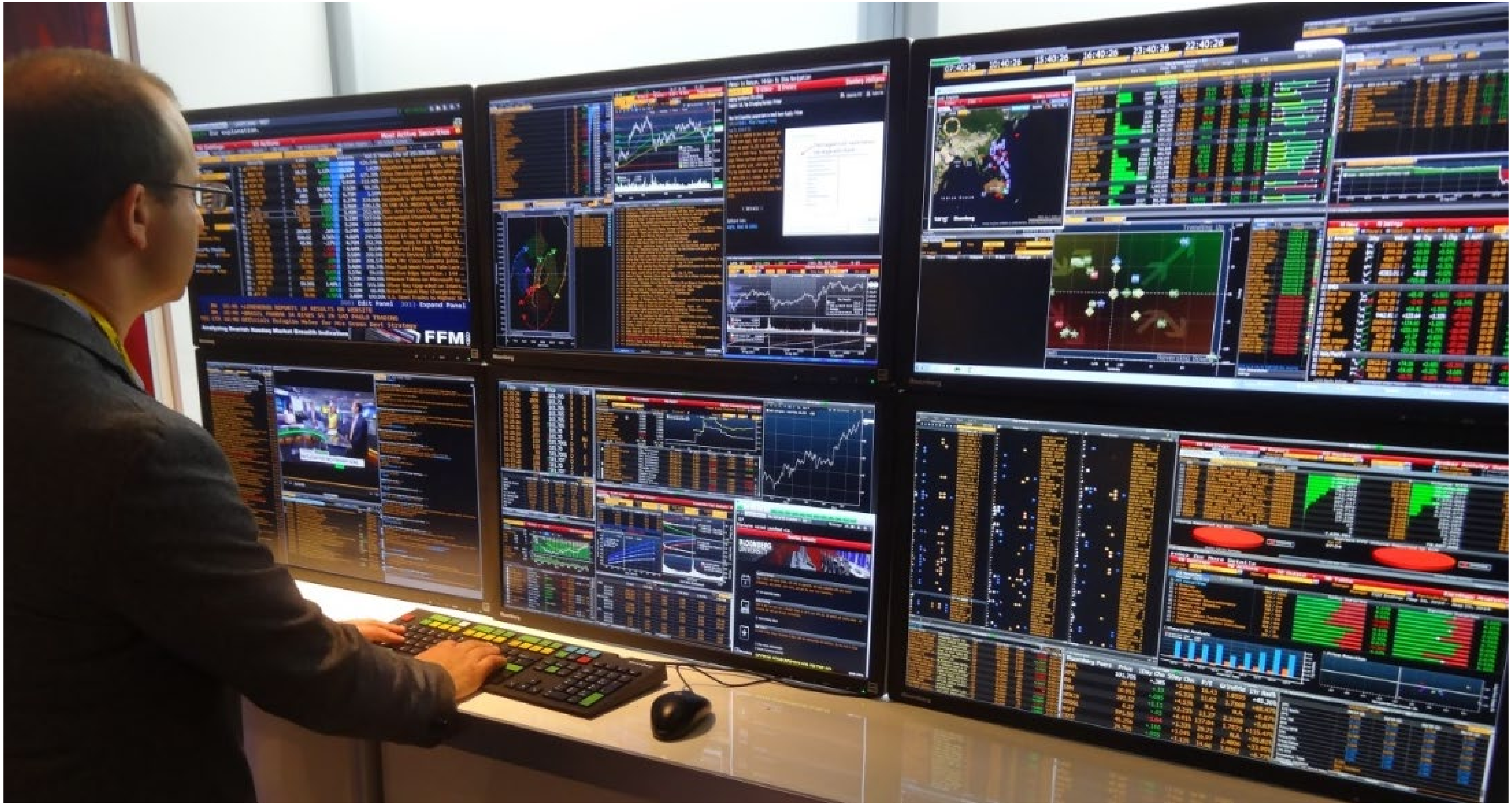
Da Vinci Tele-Operated Surgery



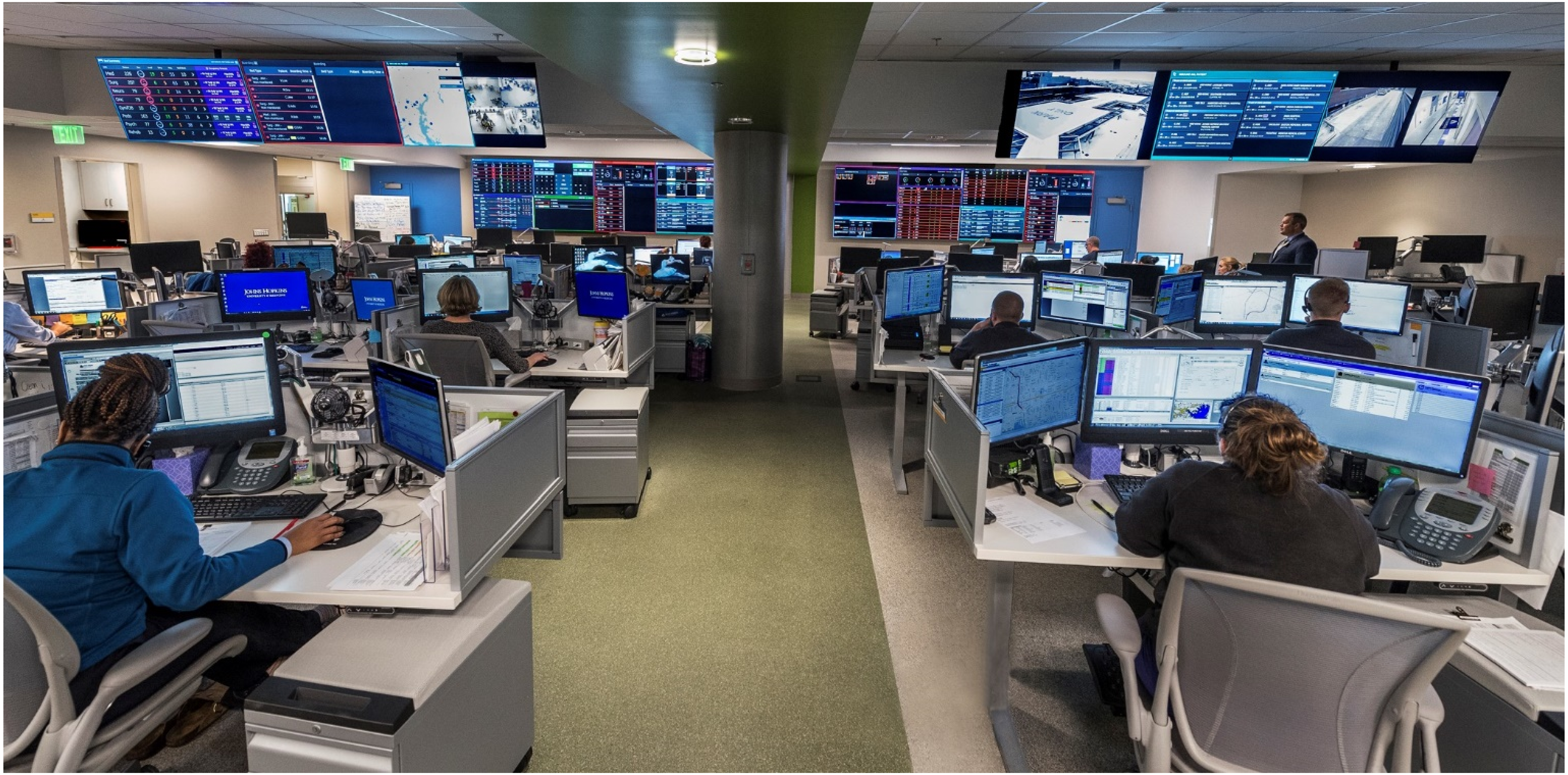
“Robots don’t perform surgery. Your surgeon performs surgery with da Vinci by using instruments that he or she guides via a console.”

https://www.davincisurgery.com/

Bloomberg Terminal

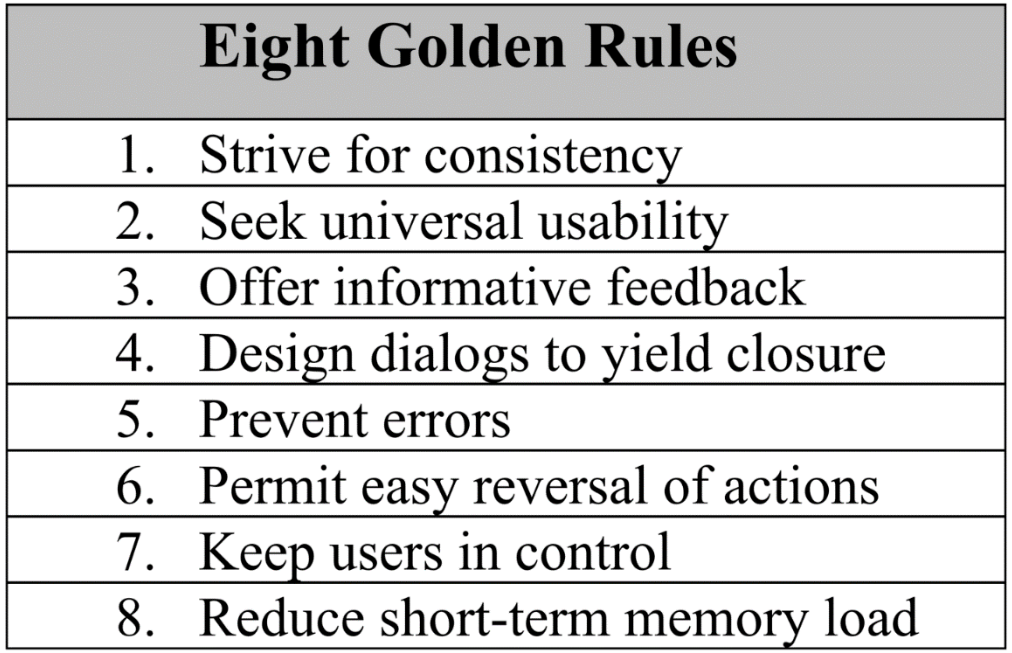


Hospital Control Center



Counter Terrorism Center

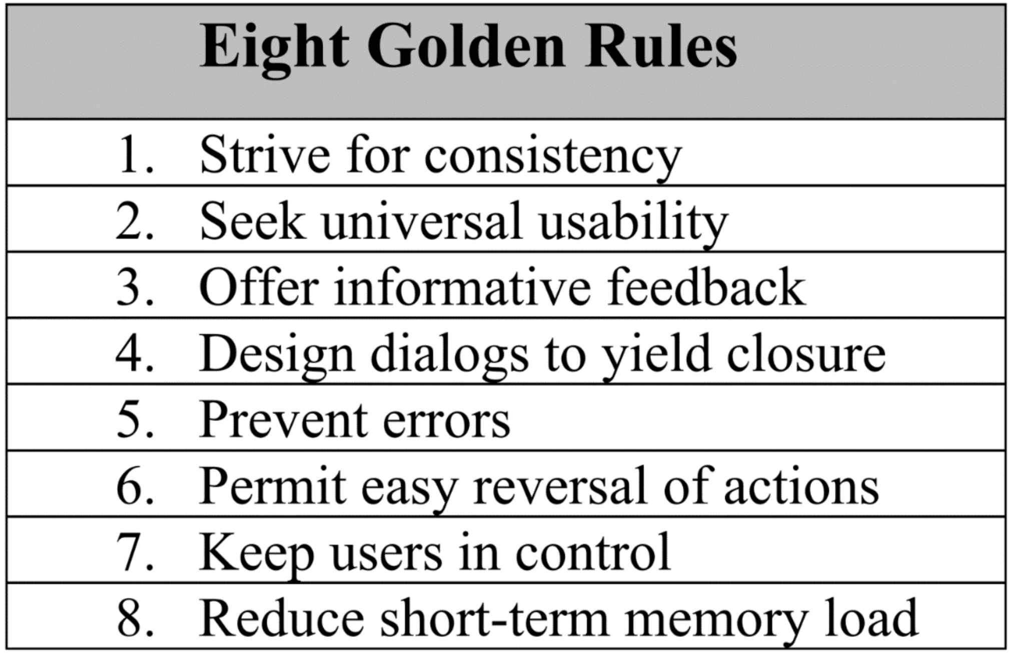
Design Guidelines



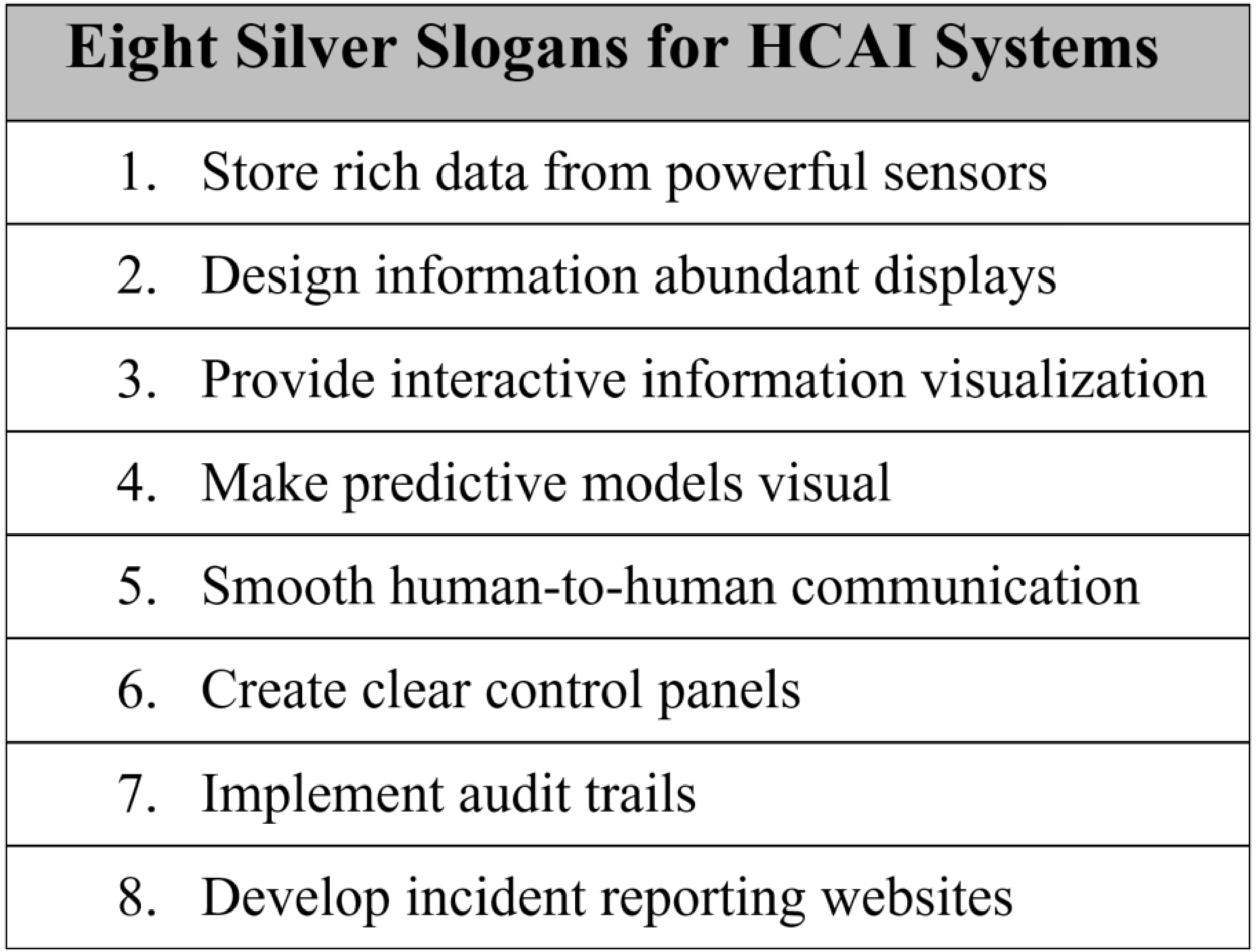
Eight Golden Rules
1. Strive for consistency
2. Seek universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Keep users in control
8. Reduce short-term memory load

https://www.cs.umd.edu/~ben/goldenrules.html

Design Guidelines



Eight Golden Rules
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

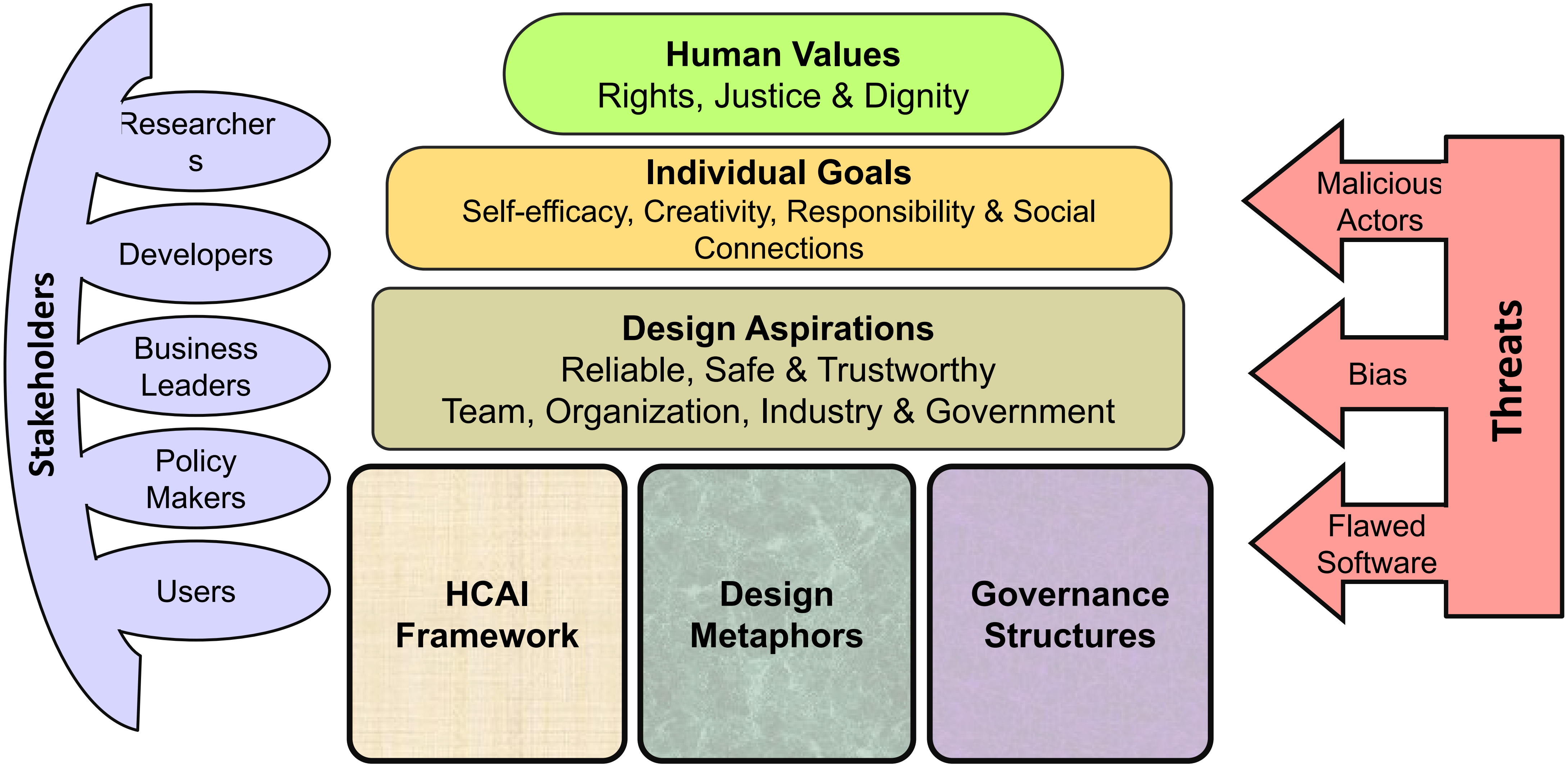


Eight Silver Slogans for HCAI Systems
1. Store rich data from powerful sensors
2. Design information abundant displays
3. Provide interactive information visualization
4. Make predictive models visual
5. Smooth human-to-human communication
6. Create clear control panels
7. Implement audit trails
8. Develop incident reporting websites

`https://www.cs.umd.edu/~ben/goldenrules.html`

Summary

Human-Centered AI



Stakeholders

Researchers

Developers

Business Leaders

Policy Makers

Users

Human Values
Rights, Justice & Dignity

Individual Goals
Self-efficacy, Creativity, Responsibility & Social Connections

Design Aspirations
Reliable, Safe & Trustworthy
Team, Organization, Industry & Government

HCAI Framework

Design Metaphors

Governance Structures

Threats

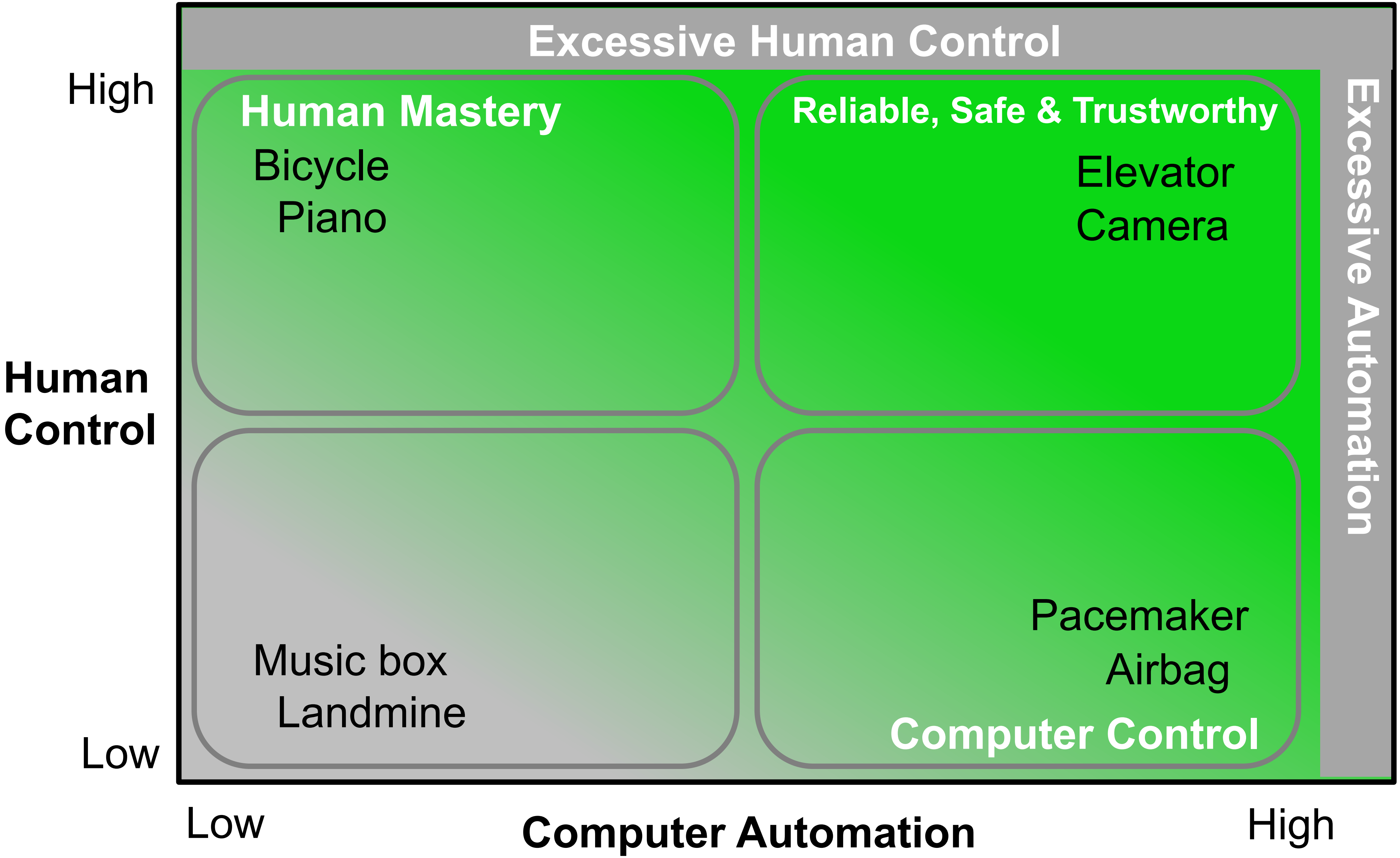
Malicious Actors

Bias

Flawed Software

Oxford University Press (Early 2022) **`https://hcil.umd.edu/human-centered-ai/`**

Human-Centered AI Framework



Excessive Human Control

Human Mastery

Bicycle

Piano

Reliable, Safe & Trustworthy

Elevator

Camera

Music box

Landmine

Pacemaker

Airbag

Computer Control

Excessive Automation

High

Human Control

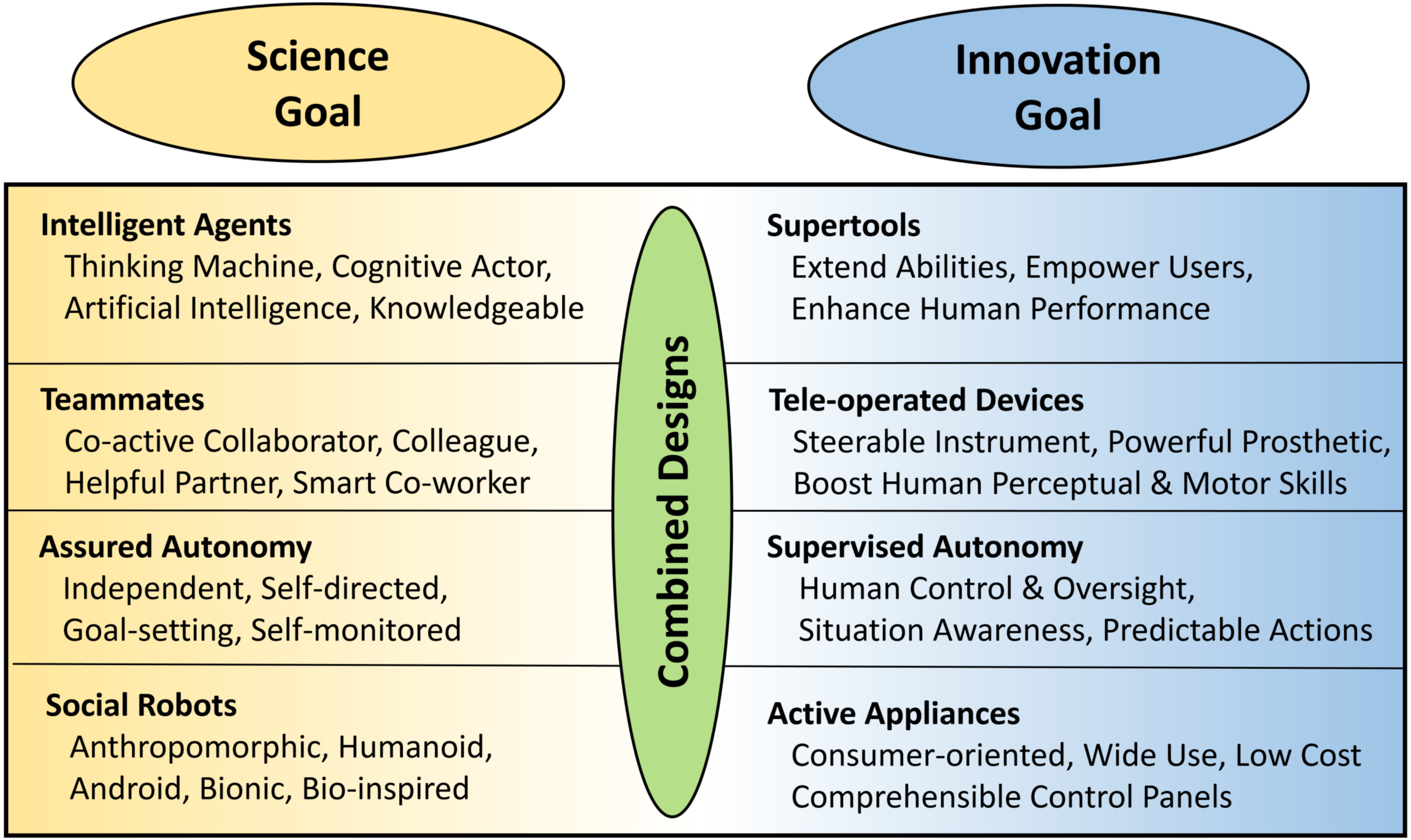
Low

Low

Computer Automation

High

Design Metaphors



Science Goal

Innovation Goal

Intelligent Agents
Thinking Machine, Cognitive Actor,
Artificial Intelligence, Knowledgeable

Teammates
Co-active Collaborator, Colleague,
Helpful Partner, Smart Co-worker

Assured Autonomy
Independent, Self-directed,
Goal-setting, Self-monitored

Social Robots
Anthropomorphic, Humanoid,
Android, Bionic, Bio-inspired

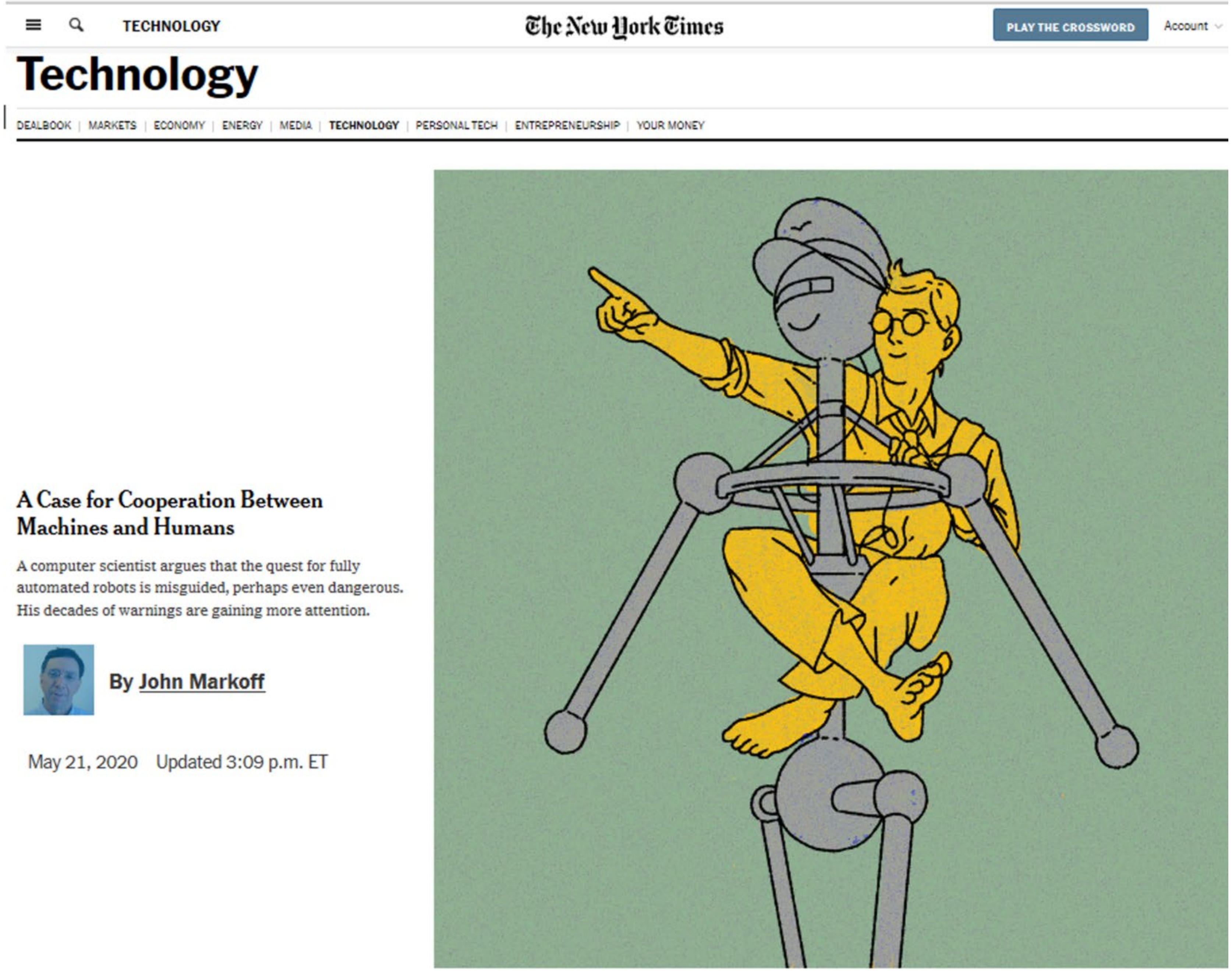
Combined Designs

Supertools
Extend Abilities, Empower Users,
Enhance Human Performance

Tele-operated Devices
Steerable Instrument, Powerful Prosthetic,
Boost Human Perceptual & Motor Skills

Supervised Autonomy
Human Control & Oversight,
Situation Awareness, Predictable Actions

Active Appliances
Consumer-oriented, Wide Use, Low Cost
Comprehensible Control Panels



TECHNOLOGY
The New York Times
PLAY THE CROSSWORD
Account
DEALBOOK | MARKETS | ECONOMY | ENERGY | MEDIA | TECHNOLOGY | PERSONAL TECH | ENTREPRENEURSHIP | YOUR MONEY

Technology

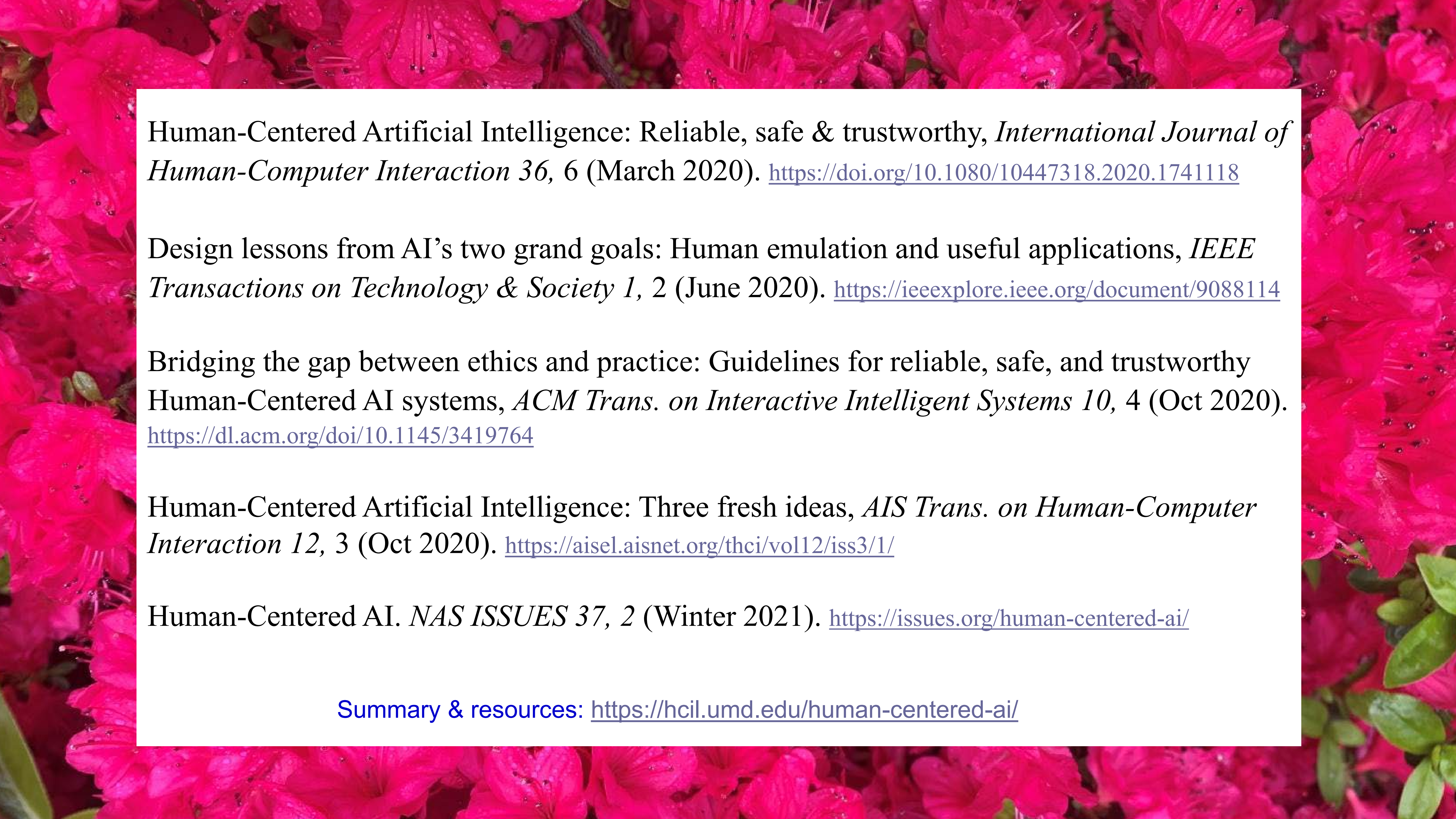
A Case for Cooperation Between Machines and Humans

A computer scientist argues that the quest for fully automated robots is misguided, perhaps even dangerous. His decades of warnings are gaining more attention.

By John Markoff

May 21, 2020 Updated 3:09 p.m. ET

https://www.nytimes.com/2020/05/21/technology/ben-shneiderman-automation-humans.html



Human-Centered Artificial Intelligence: Reliable, safe & trustworthy, *International Journal of Human-Computer Interaction 36,* 6 (March 2020). https://doi.org/10.1080/10447318.2020.1741118

Design lessons from AI's two grand goals: Human emulation and useful applications, *IEEE Transactions on Technology & Society 1,* 2 (June 2020). https://ieeexplore.ieee.org/document/9088114

Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy Human-Centered AI systems, *ACM Trans. on Interactive Intelligent Systems 10,* 4 (Oct 2020). https://dl.acm.org/doi/10.1145/3419764

Human-Centered Artificial Intelligence: Three fresh ideas, *AIS Trans. on Human-Computer Interaction 12,* 3 (Oct 2020). https://aisel.aisnet.org/thci/vol12/iss3/1/

Human-Centered AI. *NAS ISSUES 37, 2* (Winter 2021). https://issues.org/human-centered-ai/

Summary & resources: https://hcil.umd.edu/human-centered-ai/

The Future is Human-Centered