IFIP Workshop on Intelligent Vehicle Dependability and Security (IVDS)
January 29, 2021

Designing for Increased Autonomy & Human Control

Ben Shneiderman **@benbendc**

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

Member, National Academy of Engineering

Photo: BK Adams

UNIVERSITY OF MARYLAND

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

IFIP Workshop on ~~Intelligent~~ Vehicle Dependability and Security (IVDS)
January 29, 2021

Designing for Increased Automation & Human Control

Ben Shneiderman **@benbendc**

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

Member, National Academy of Engineering

Photo: BK Adams

UNIVERSITY OF MARYLAND

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


What is Human-Centered AI?

Human-Centered AI

Amplify, Augment, Enhance & Empower People

Human Responsibility

Supertools and Active Appliances

Visual Interfaces to Prevent/Reduce Explanations

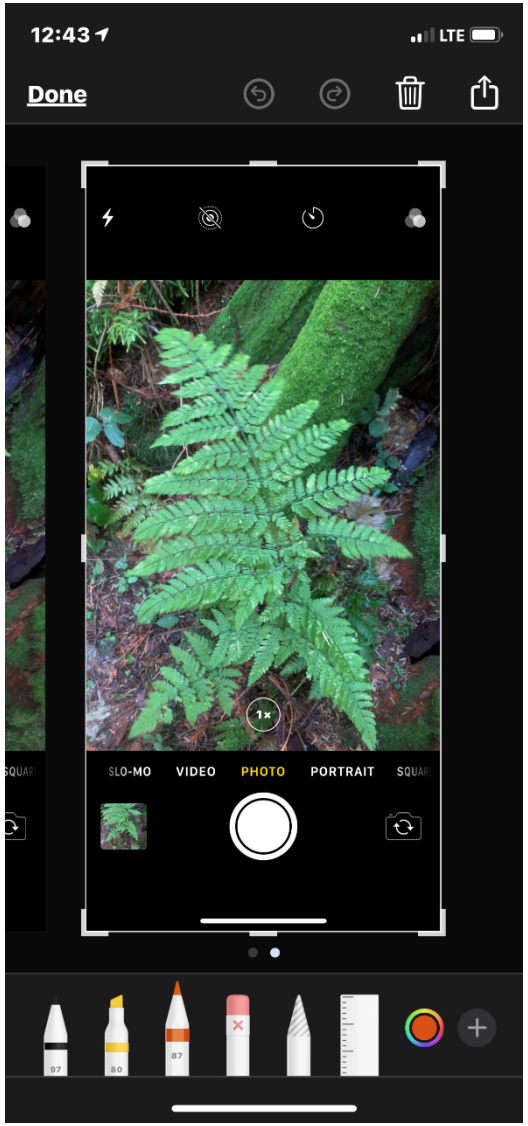
Audit Trails to Analyze Failures & Near Misses

Independent Oversight

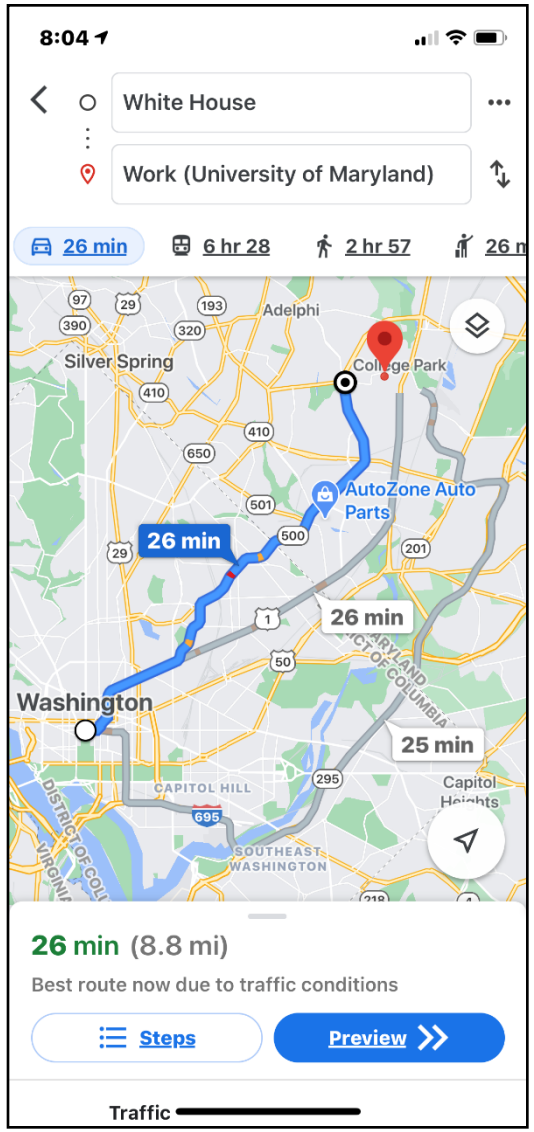
→ Reliable, Safe & Trustworthy

Supertools

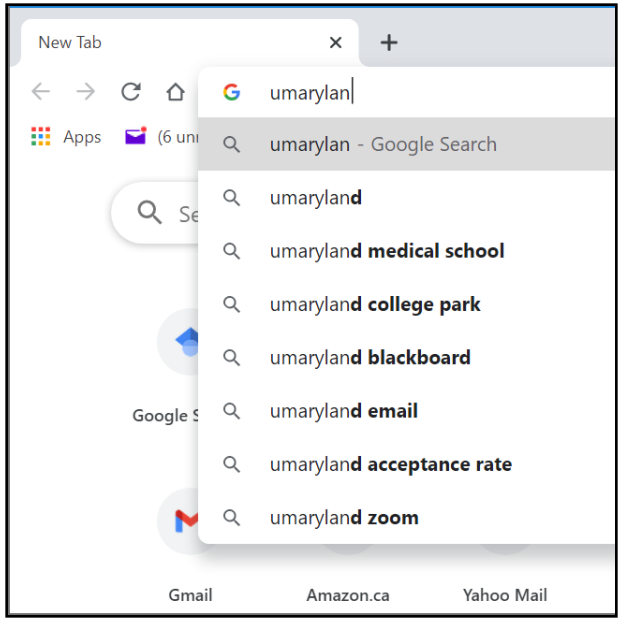
Digital Camera Controls



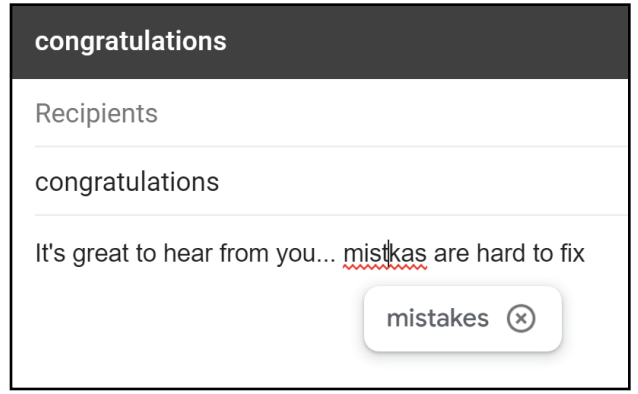
Navigation Choices



Texting Autocompletion

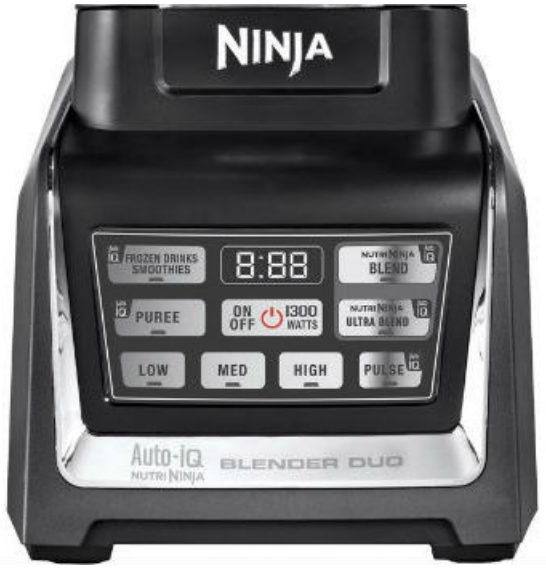


Spelling correction

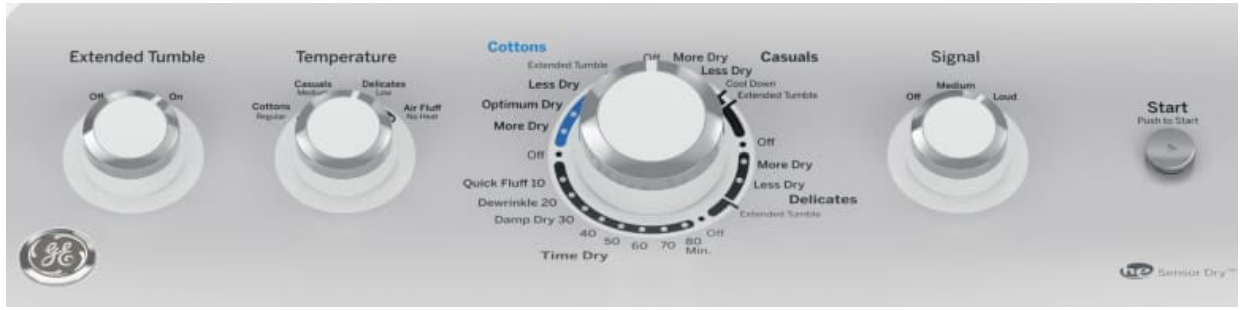


Active Appliances

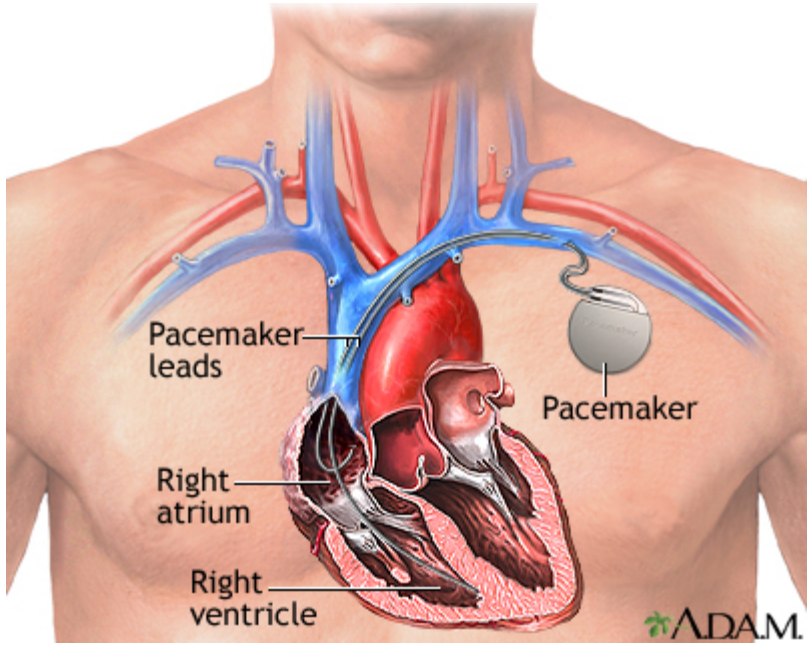
Coffee maker, Rice cooker, Blender

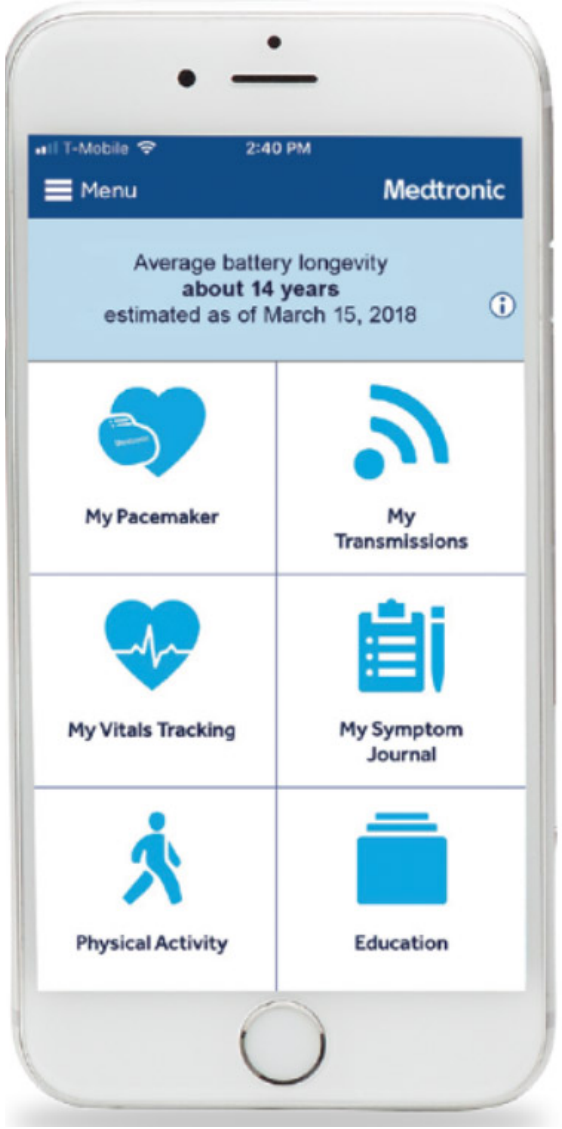


Dishwasher, Clothes Washer/Dryer



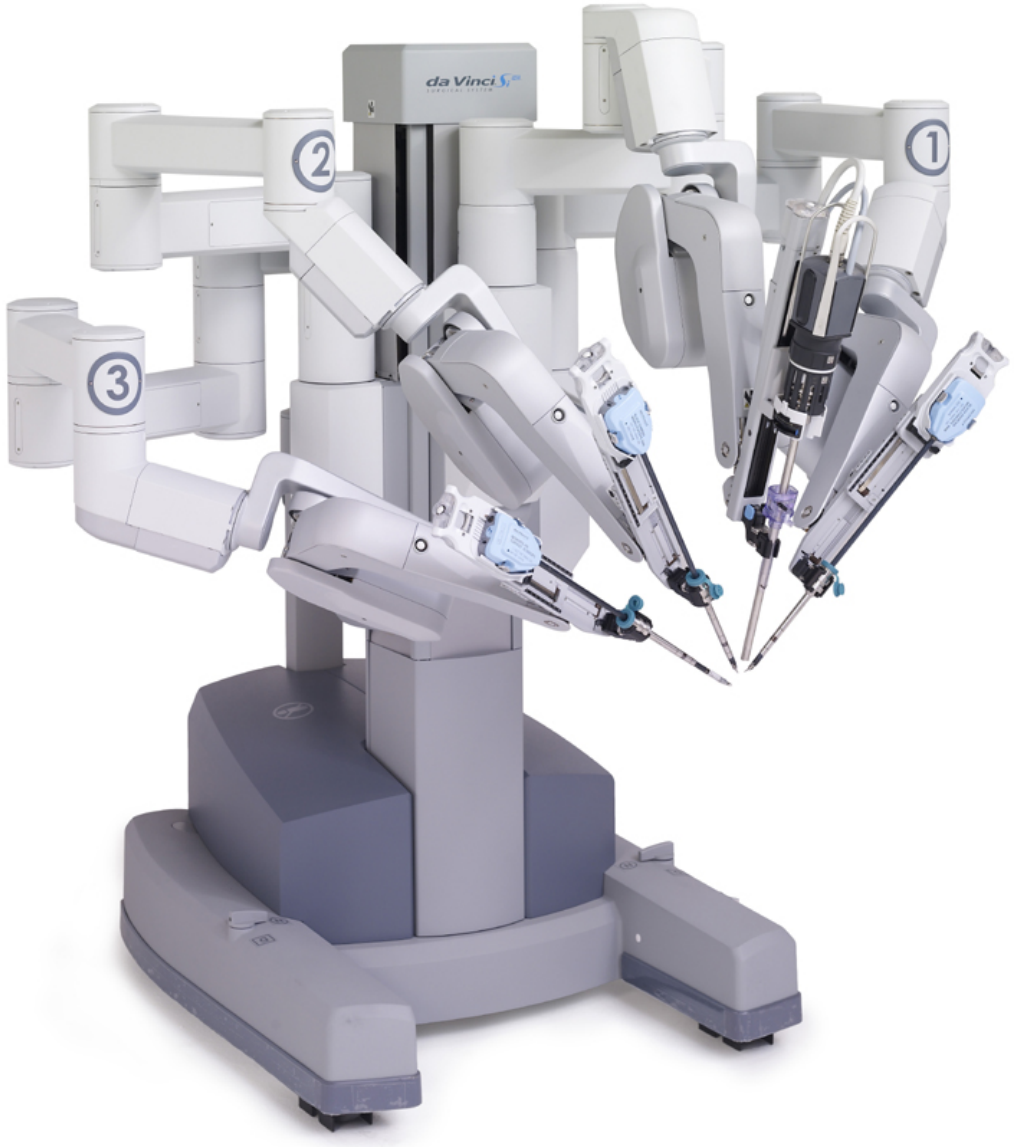
Implanted Cardiac Pacemakers





NASA Mars Rovers are Tele-Operated

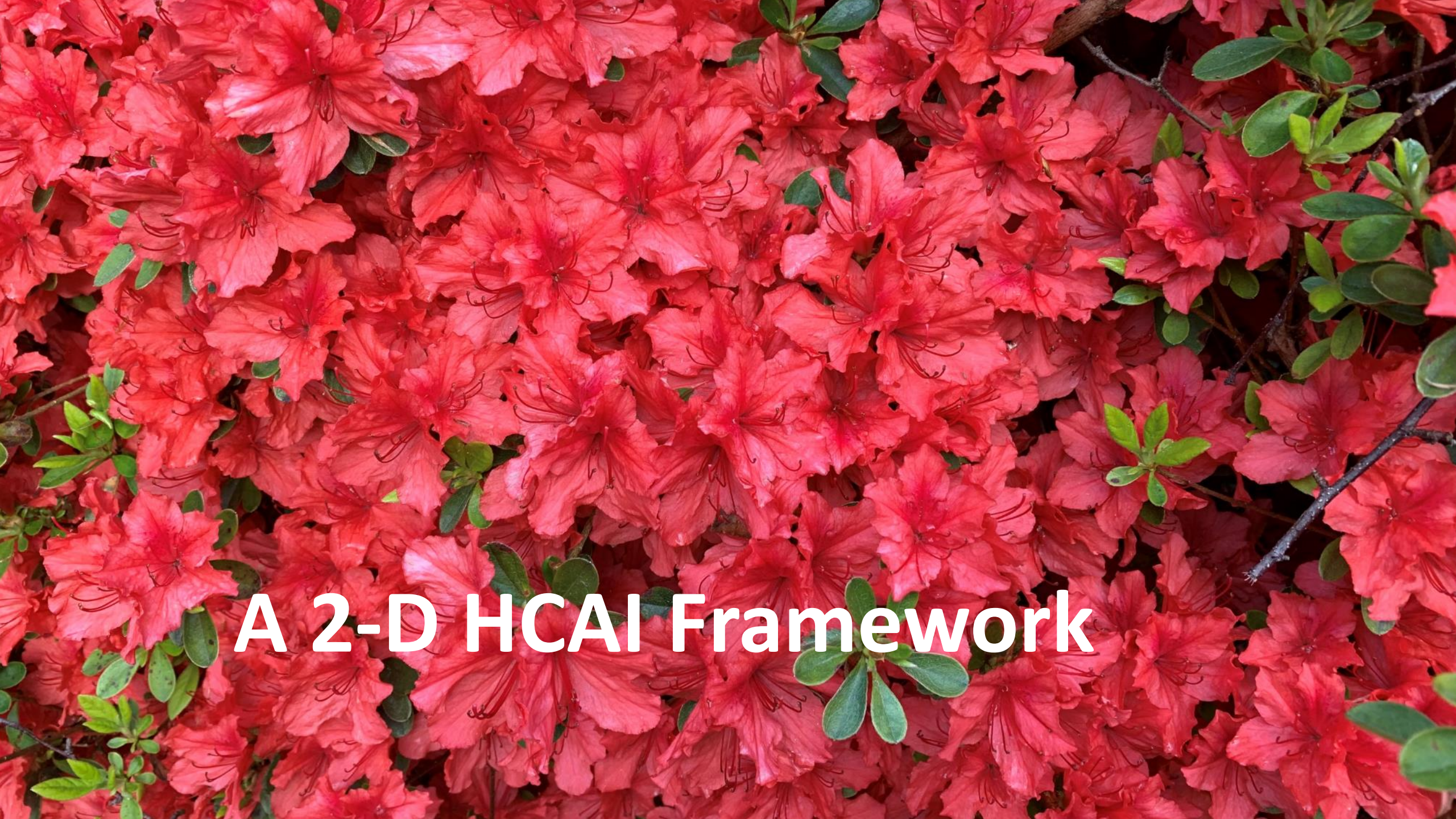
DaVinci Tele-Operated Surgery



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

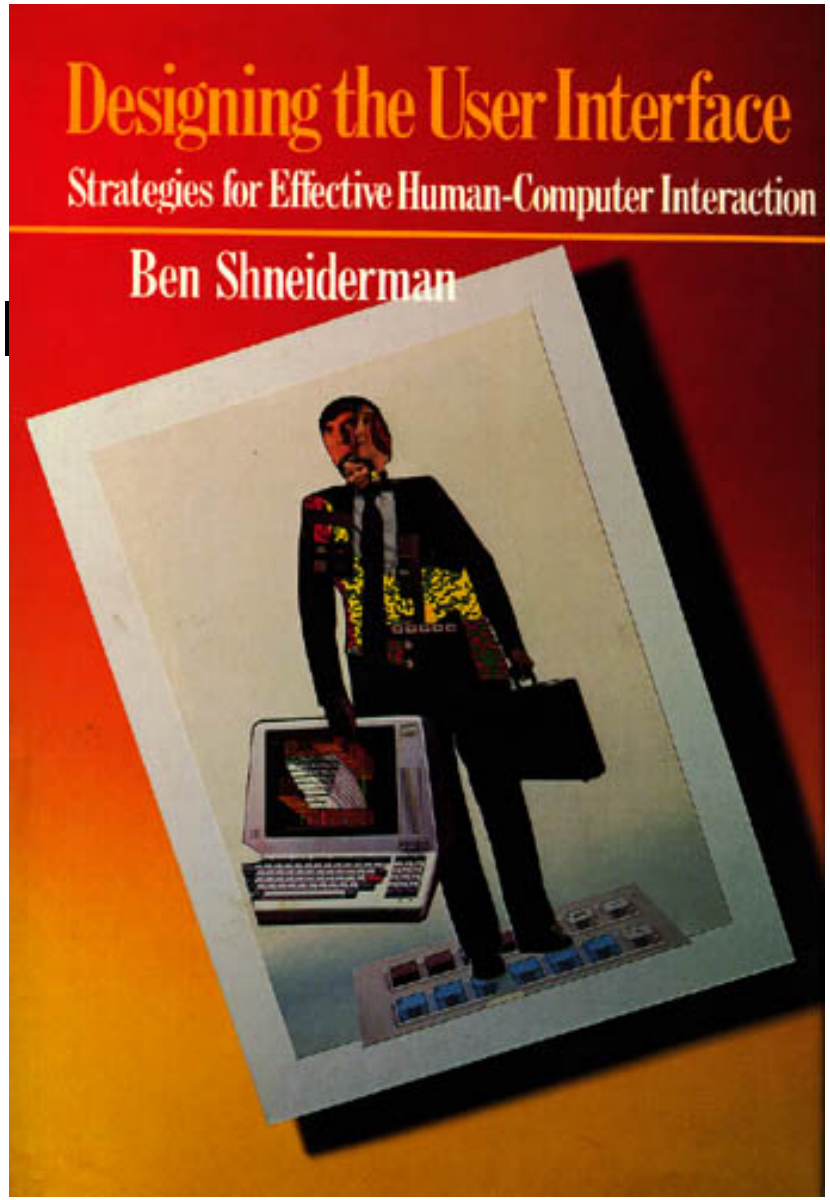
https://www.davincisurgery.com/

Bloomberg Terminal


A 2-D HCAI Framework

Designing the User Interface

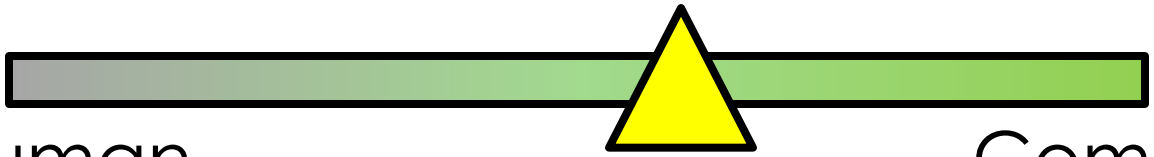
Balancing automation & human control



First Edition: 1986

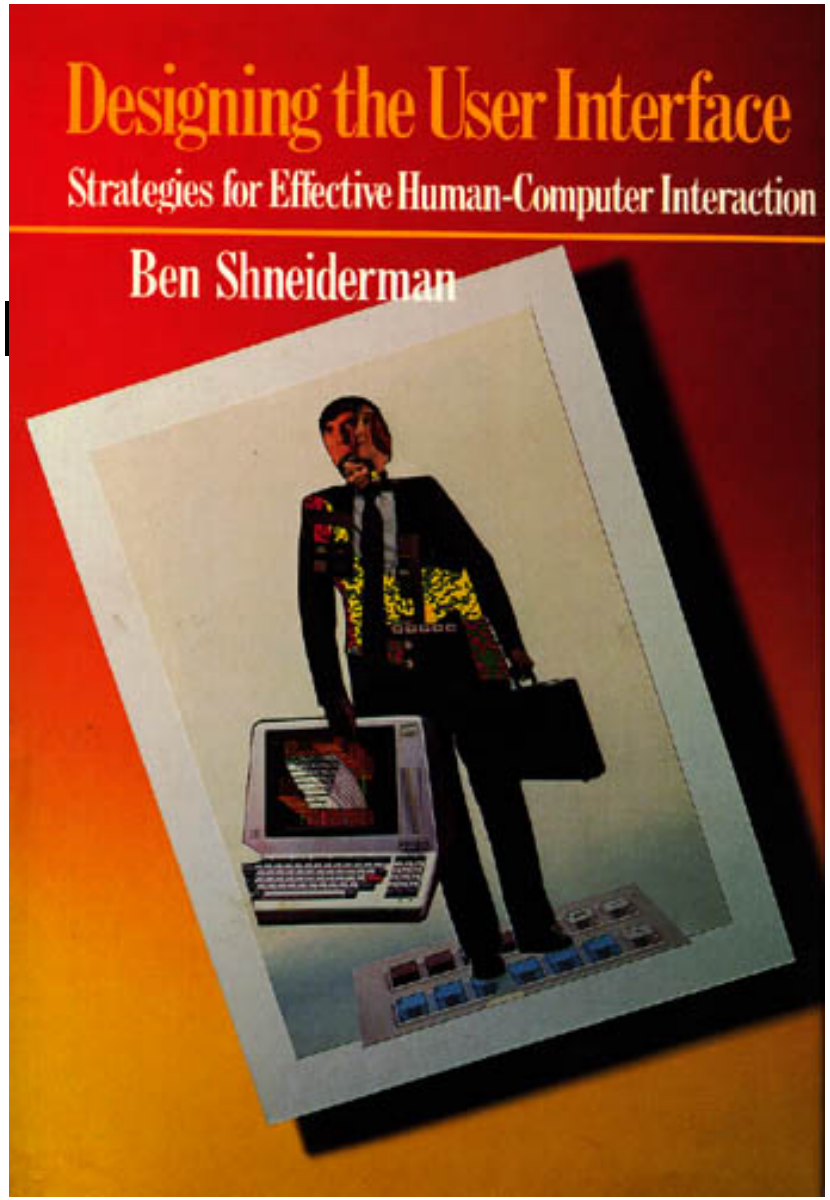
Designing the User Interface

Balancing automation & human control



Human control

Computer automation



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

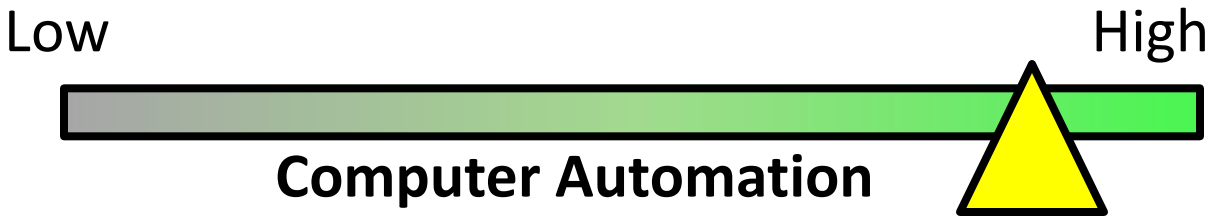
Designing the User Interface

Ensuring human control
while increasing automation

Sixth Edition: 2016

Designing the User Interface

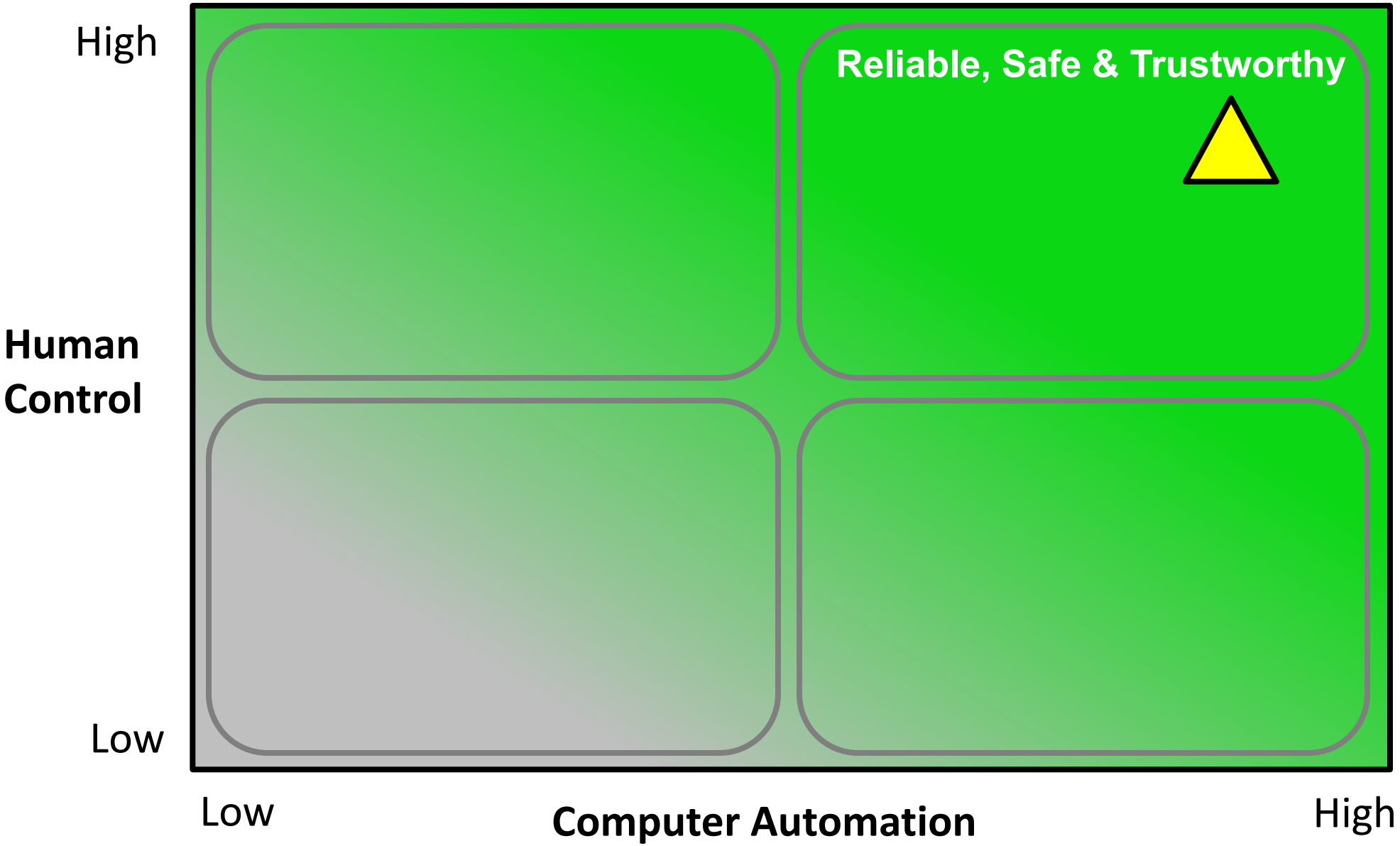
Ensuring human control
while increasing automation



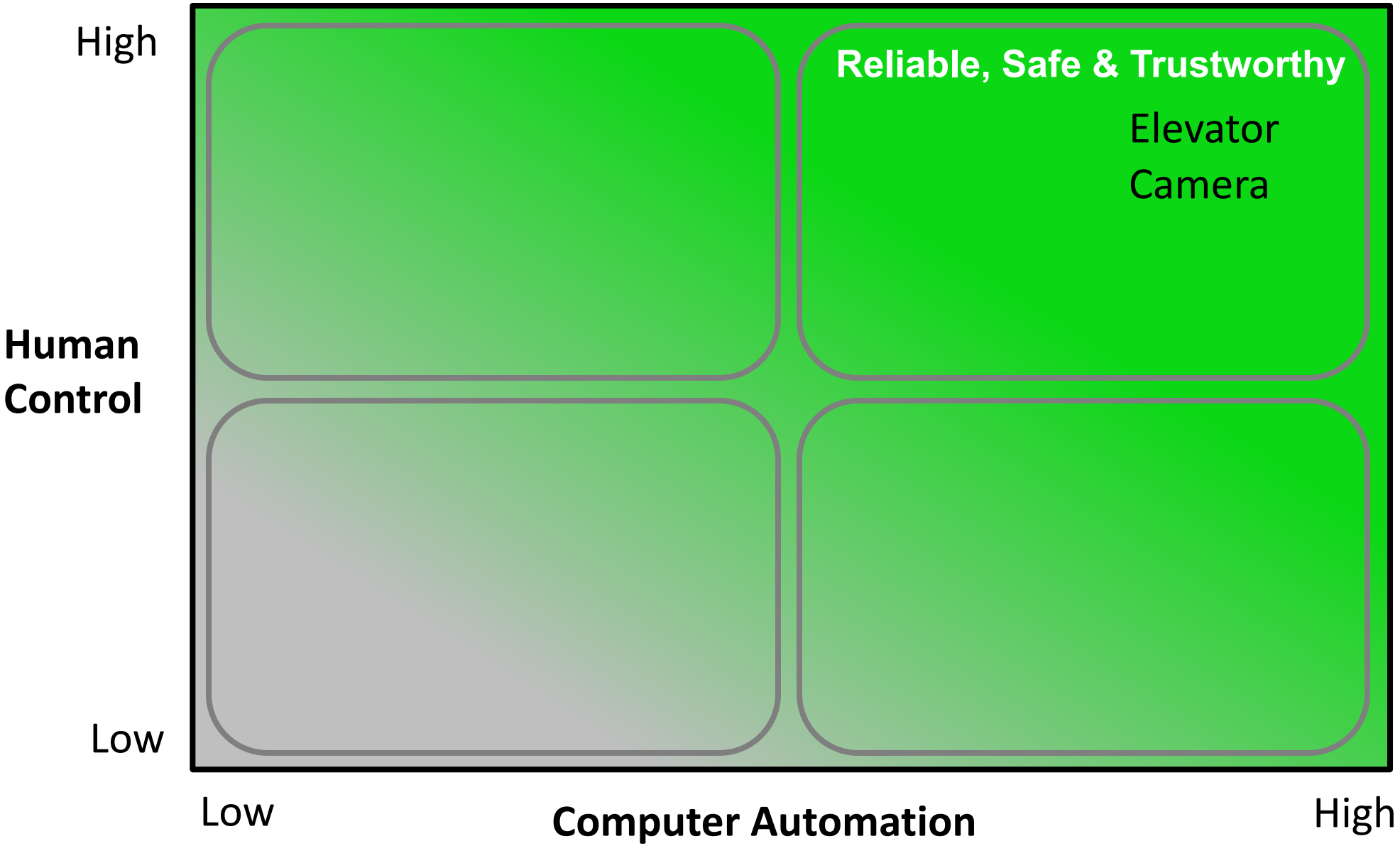
Sixth Edition: 2016

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

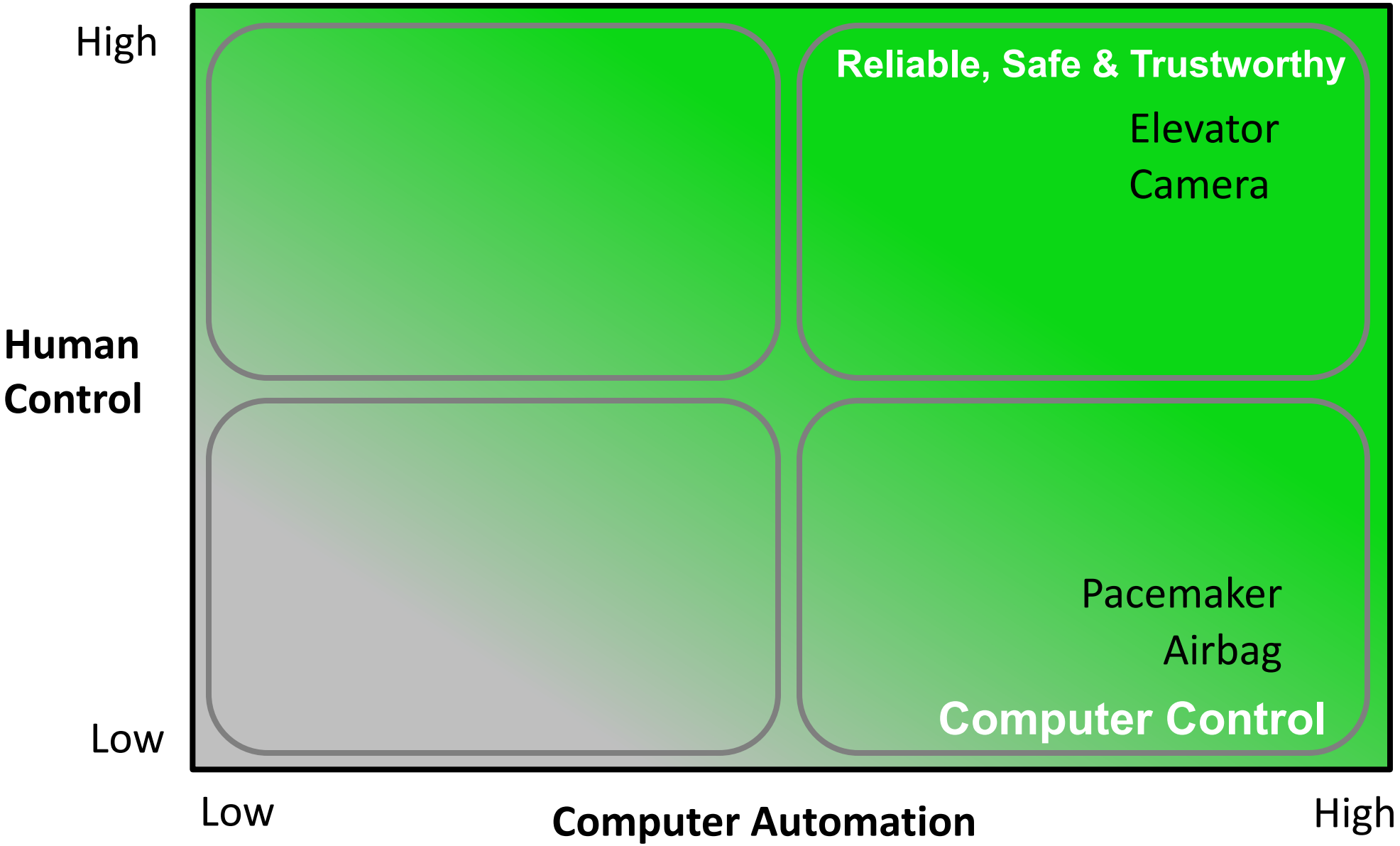
Human-Centered AI



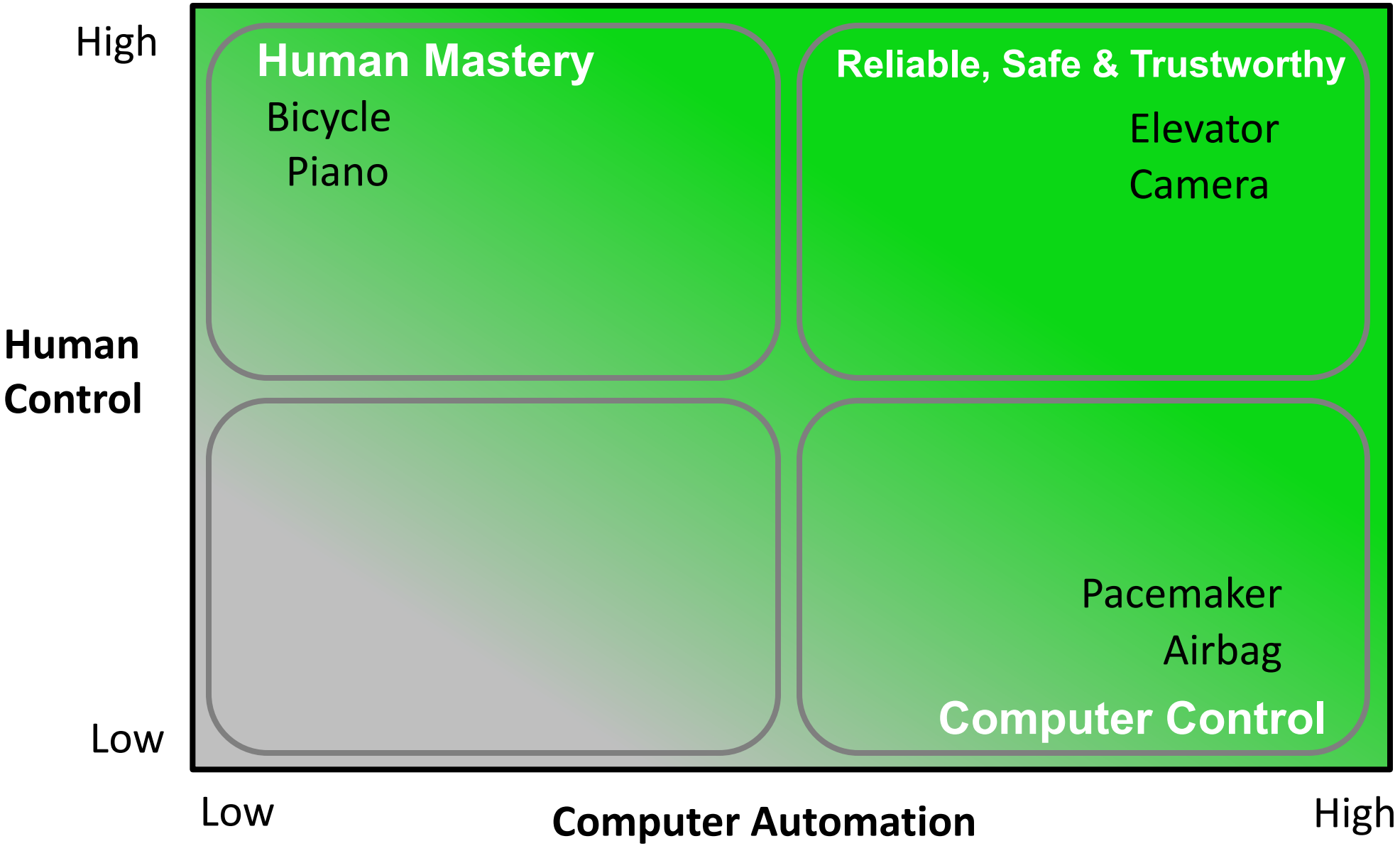
Human-Centered AI



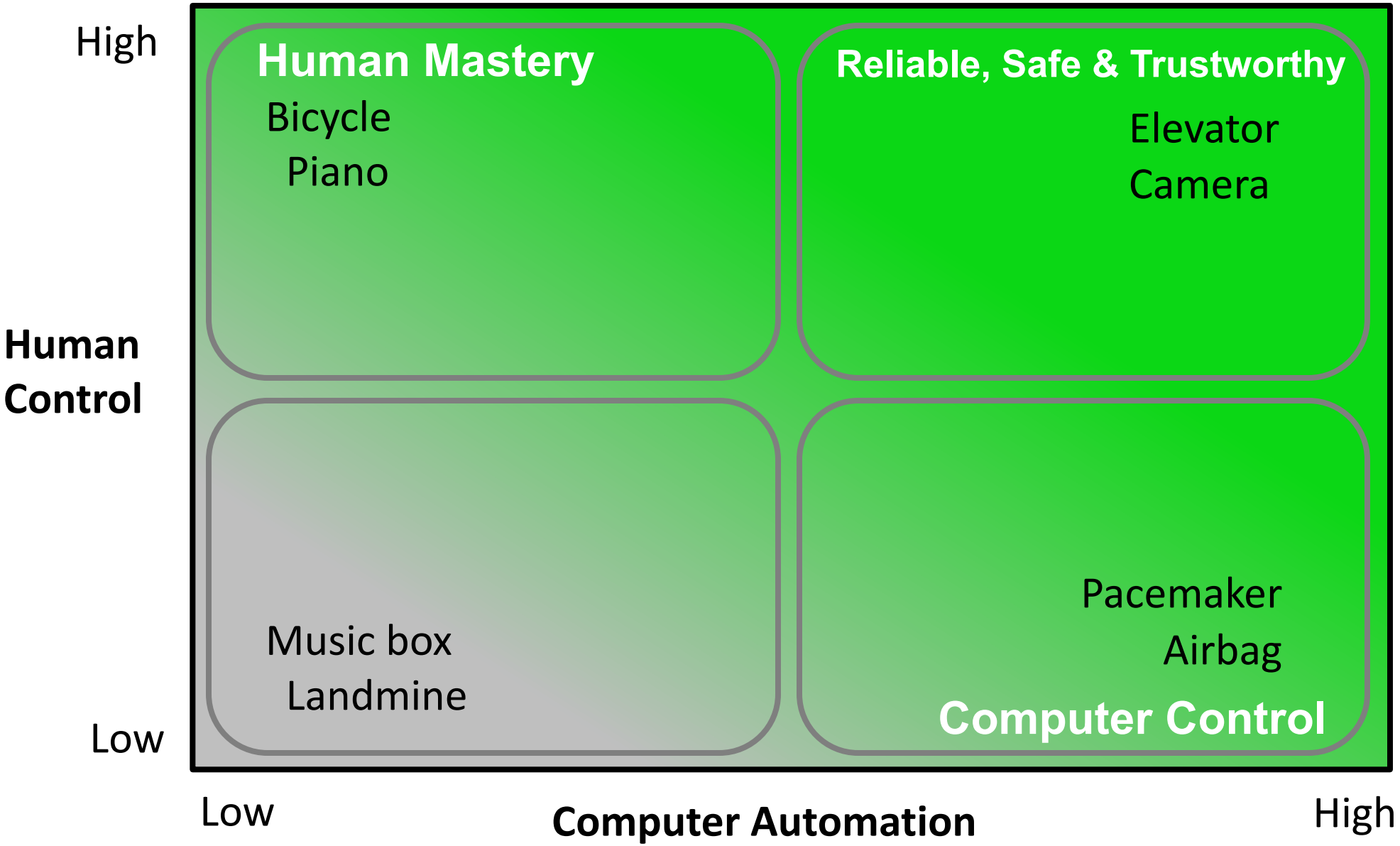
Human-Centered AI



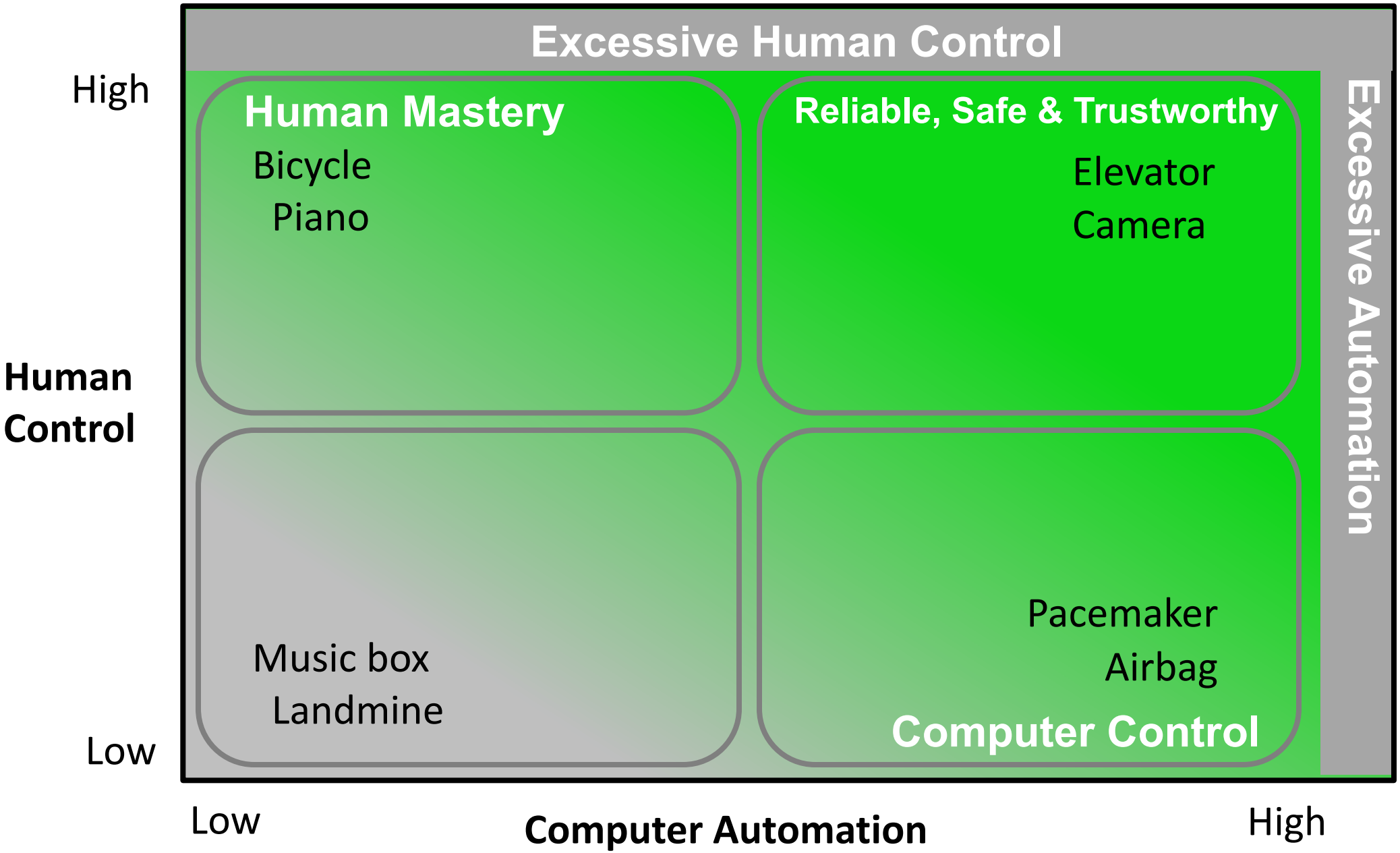
Human-Centered AI



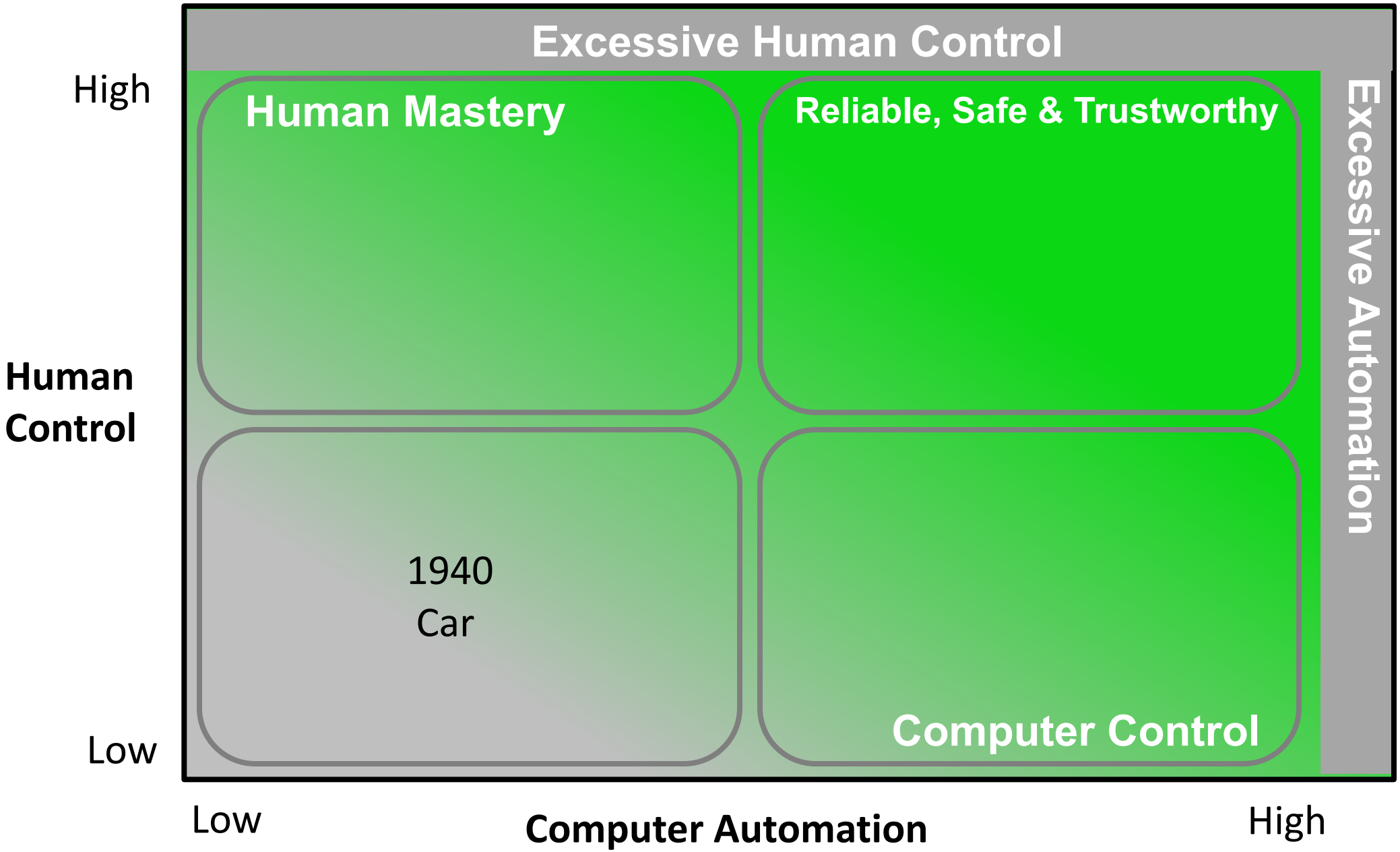
Human-Centered AI



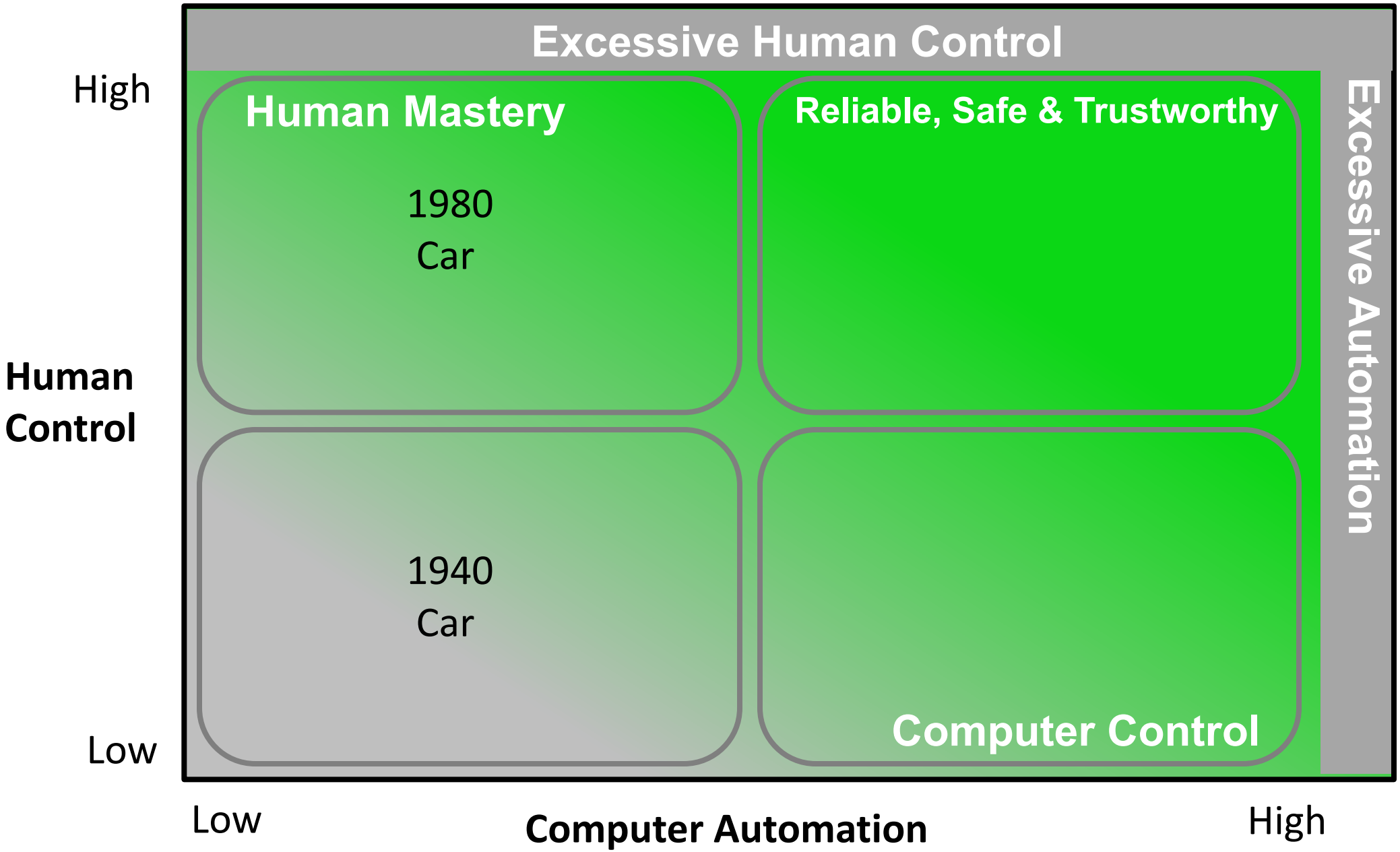
Human-Centered AI



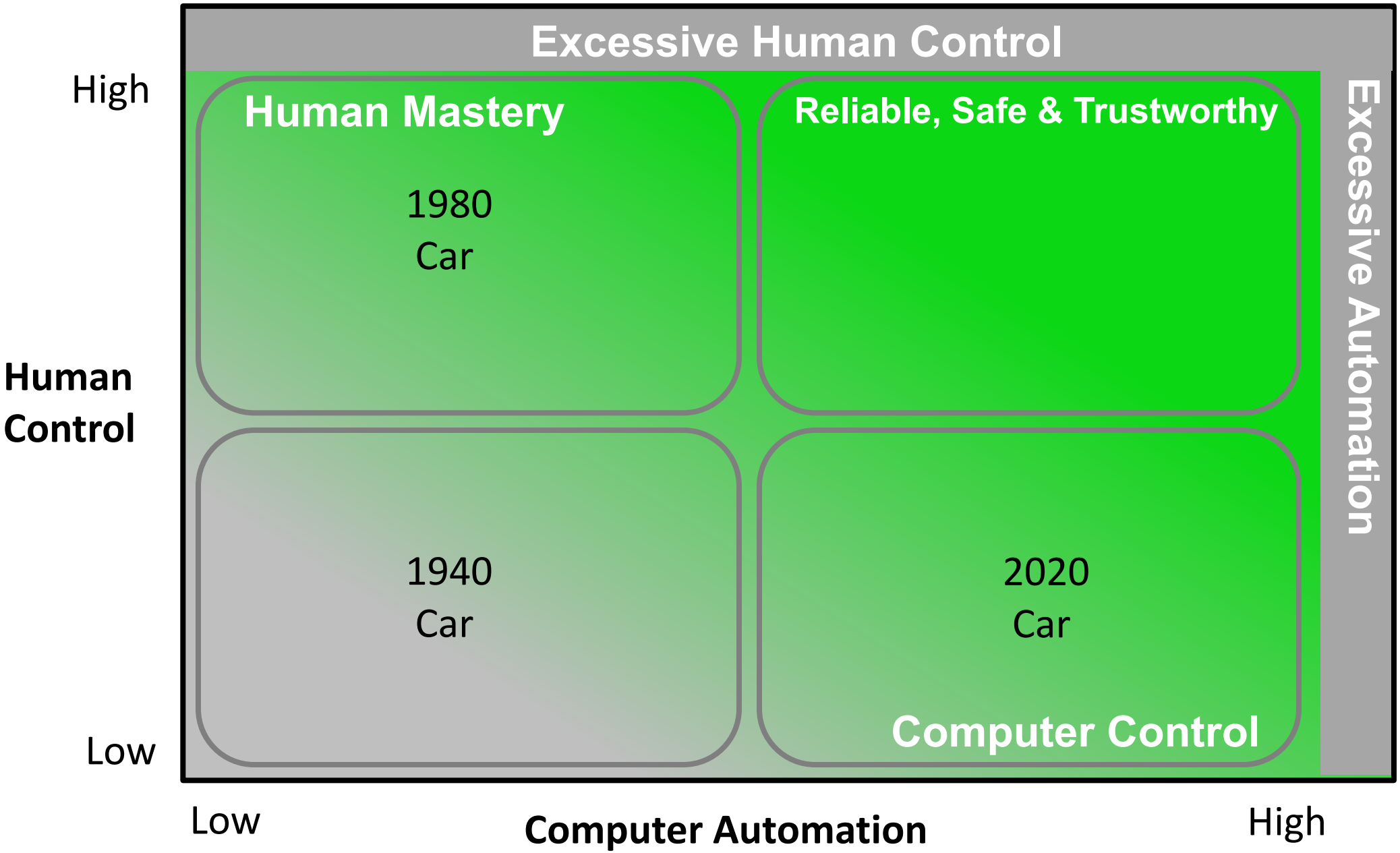
Car Control Designs



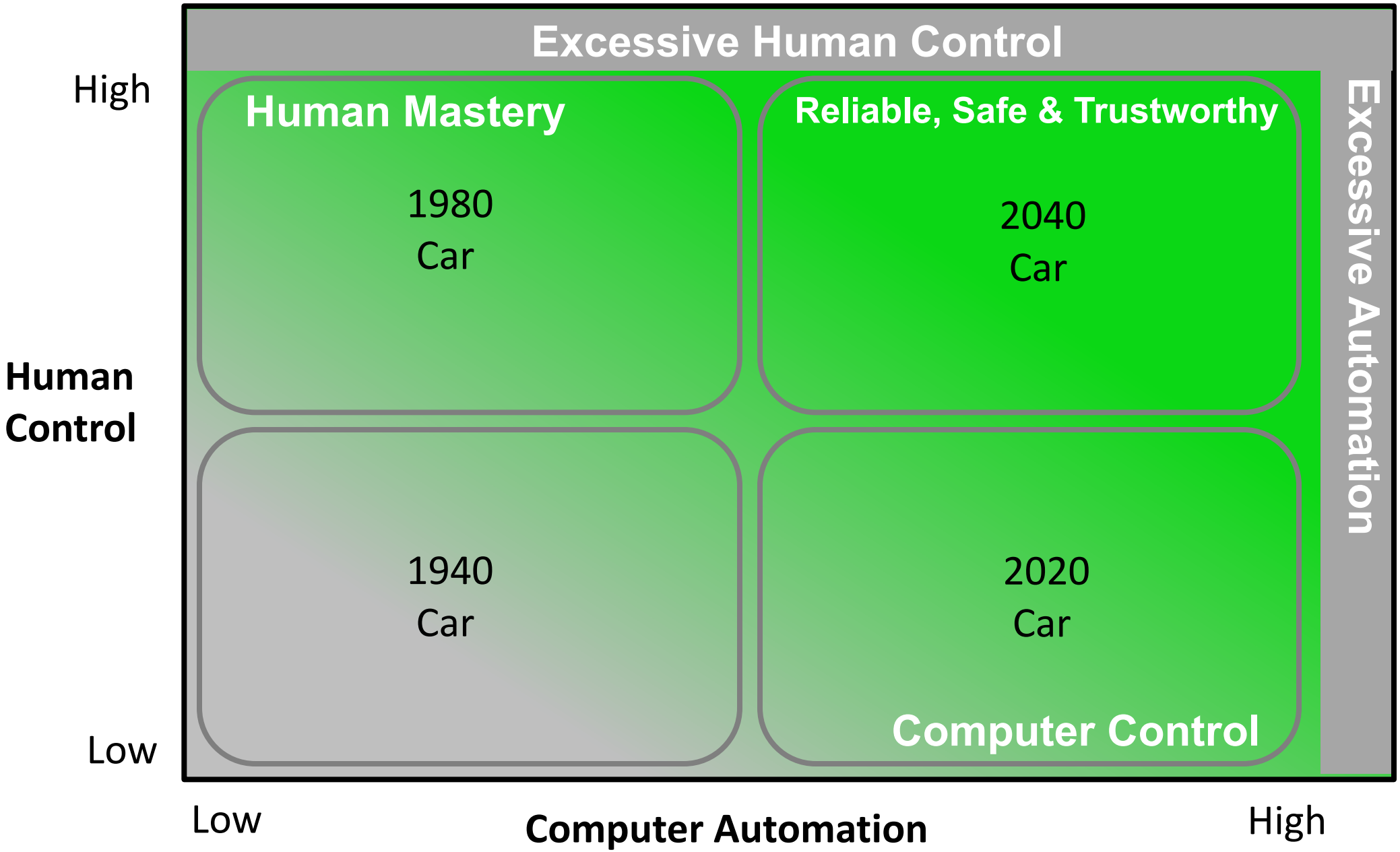
Car Control Designs



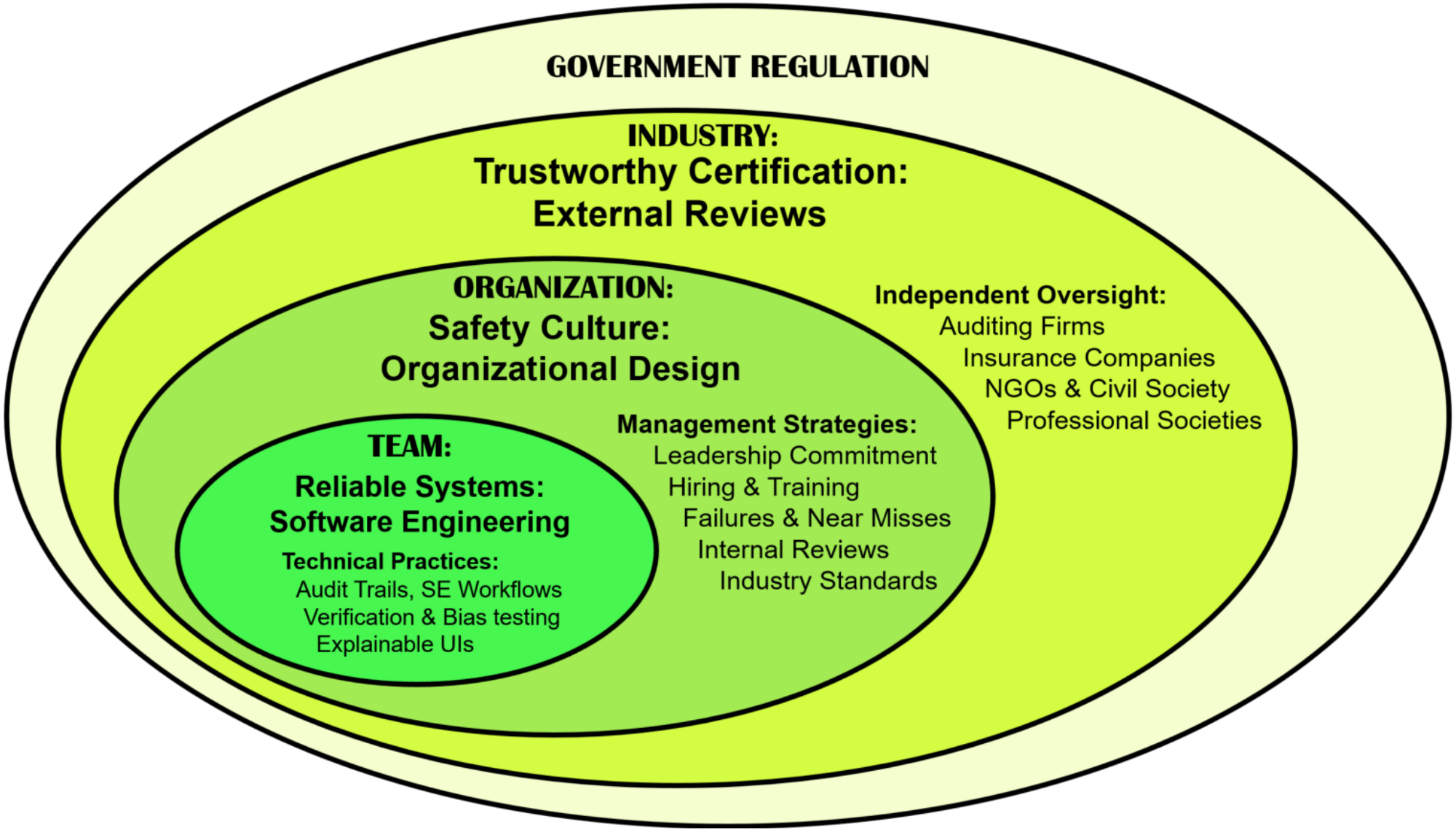
Car Control Designs



Car Control Designs



Governance Structures for Human-Centered AI



https://dl.acm.org/doi/10.1145/3419764

Summary

Human-Centered AI

Amplify, Augment, Enhance & Empower People

Human Responsibility

Supertools and Active Appliances

Visual Interfaces to Prevent/Reduce Explanations

Audit Trails to Analyze Failures & Near Misses

Independent Oversight

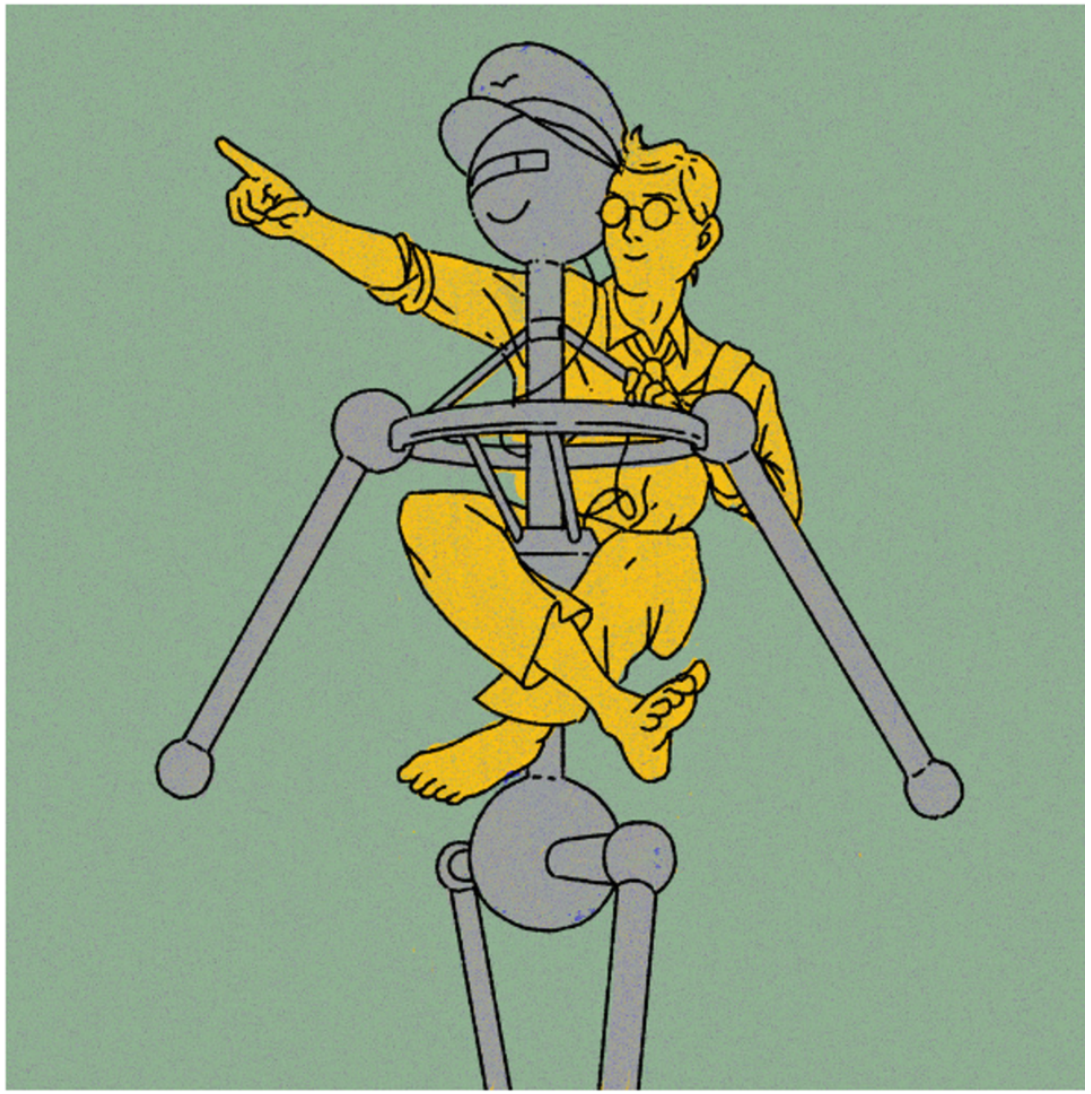
→ Reliable, Safe & Trustworthy

A Case for Cooperation Between Machines and Humans

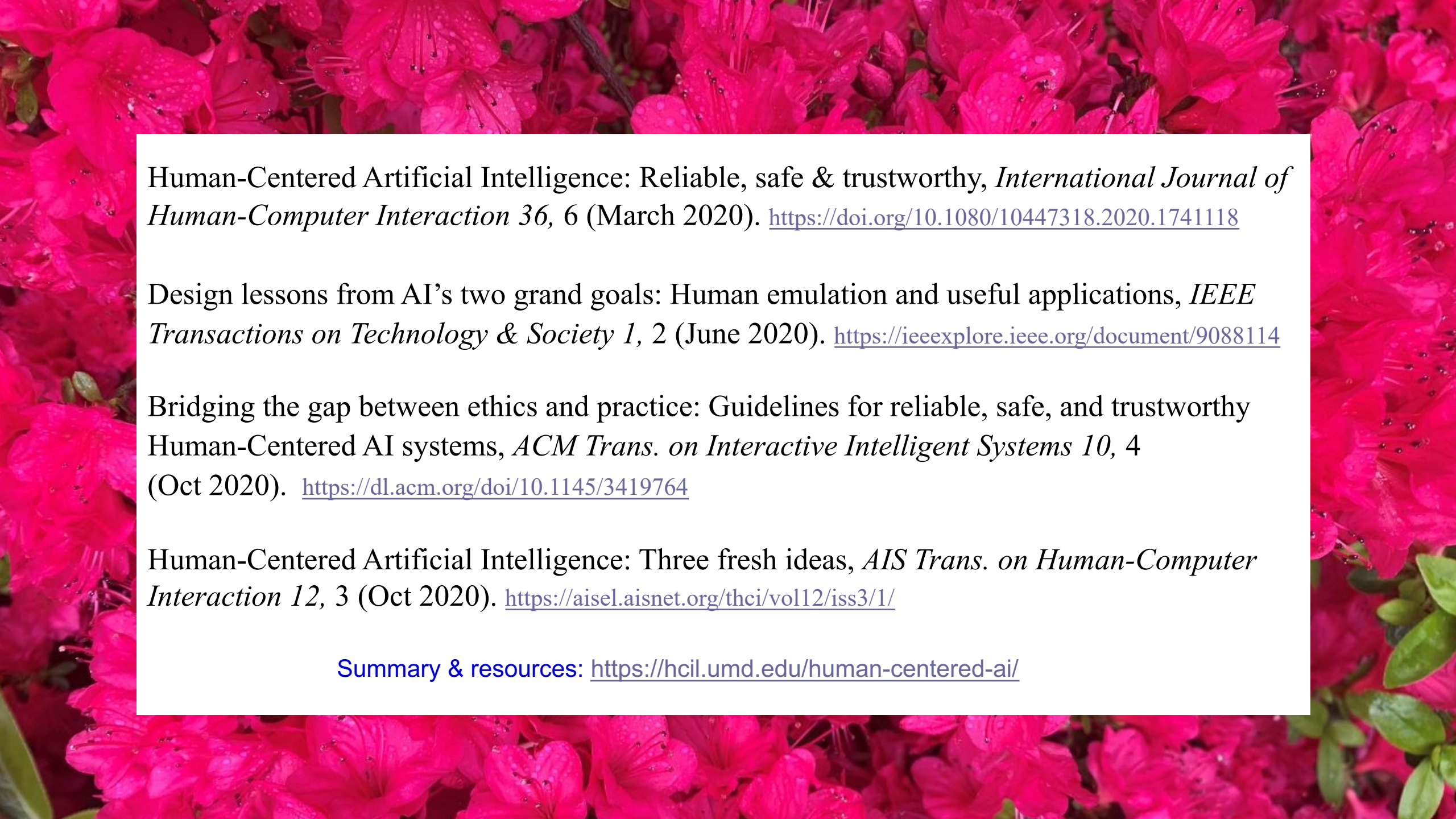
A computer scientist argues that the quest for fully automated robots is misguided, perhaps even dangerous. His decades of warnings are gaining more attention.

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/

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