AI's Strange Bedfellows: Values, Accountability and Humility

- Colleen P. Lyons, Ph.D.
- World Summit AI Amsterdam
- October 9, 2024

Disclaimer

- Ideas, thoughts, and opinions expressed in the presentation are my own, and do not represent any organization.
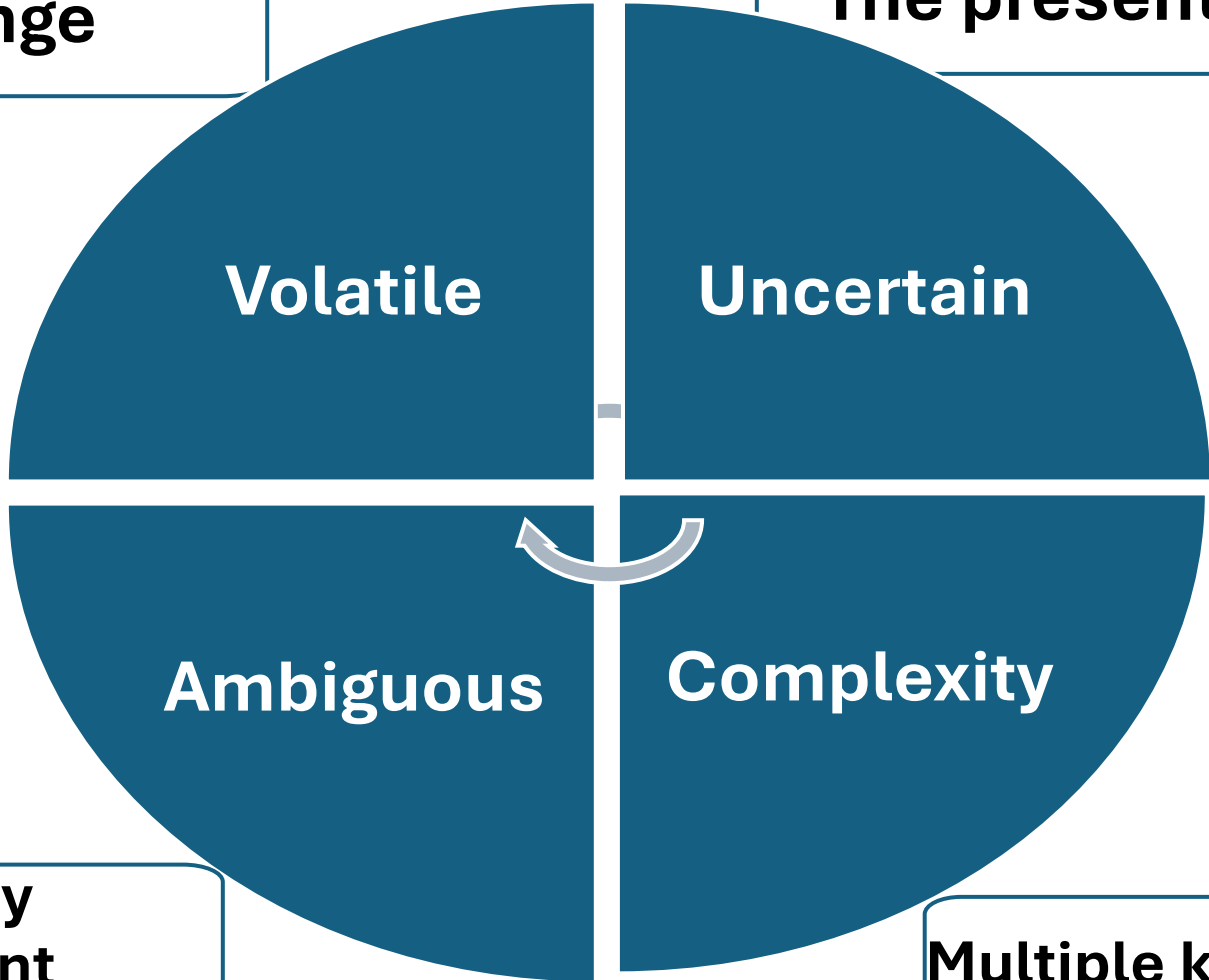
Four Take Away's

- VUCA Rules the Day

- Outcome focus: AI is a Means to an End

- Competitive Advantage: Values-Driven Leadership

- Legacy Moment
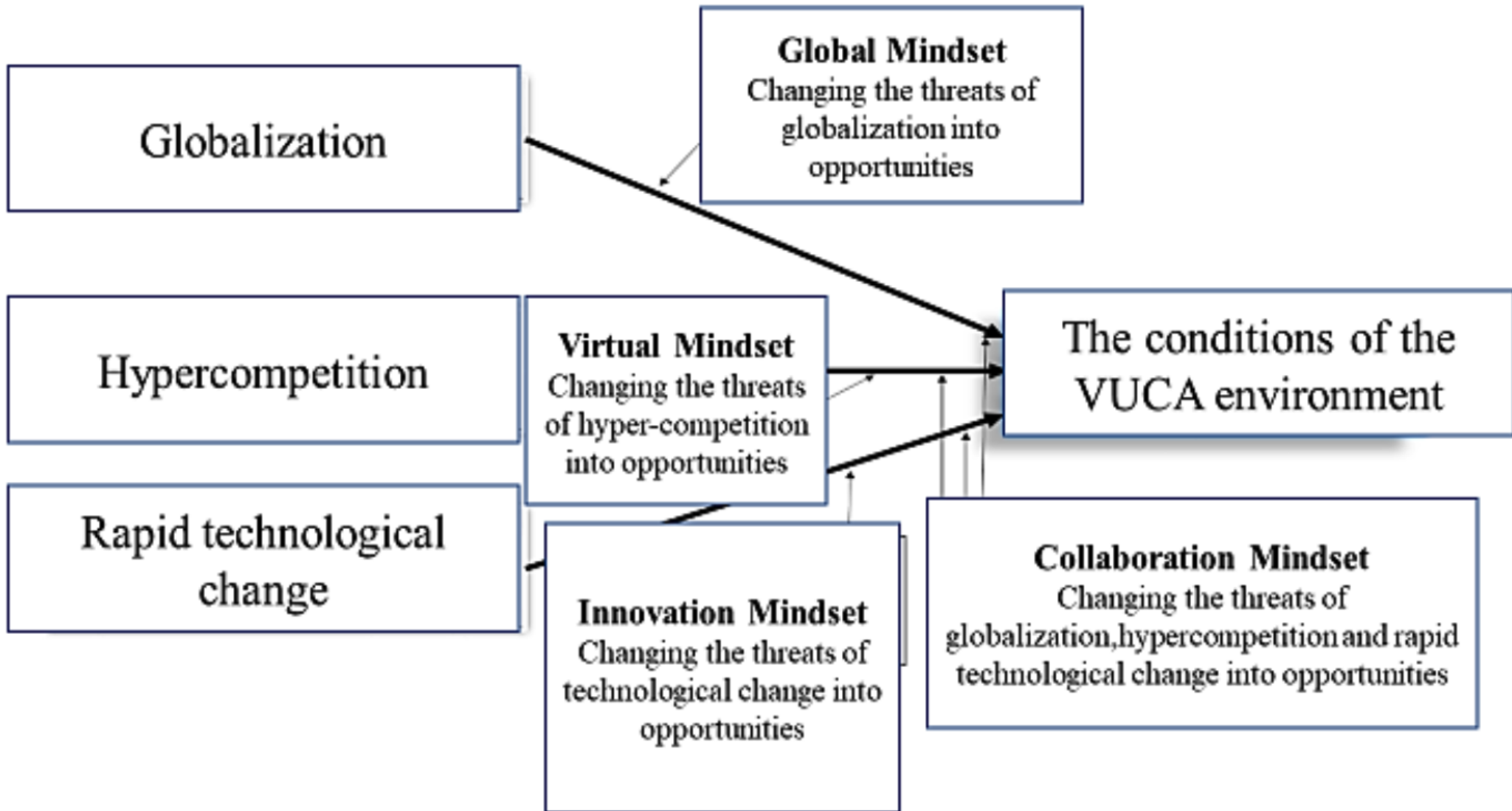
VUCA ENVIRONMENT

Rate of change

The present is unclear

Volatile

Uncertain

Ambiguous

Complexity

Lack of clarity about an event

Multiple key decision factors

Source: Jari1 & Theresa Lauraéus 2019
Fig. 3 Framework showing relationships in the VUCA environment. Modification of [40]

- Management Behavior Doesn't Support

- Inadequate Resource/Budget

- Employee Resistance

- Weak Accountability

- Old Ways Overpower

- Change Fatigue- Change Is Hard

- Etc. Etc. Etc.

70% of Change Management Efforts Fail (MC Kinsey 2015)

VUCA MANDATES CHANGE

AI Ethics: Are They Obsolete?

Facebook's ethical failures are not accidental; they are part of the business model

AI's description: a person in a suit talking to a group of people
KEEP A HUMAN IN THE LOOP!

COMMENTARY · T

OpenA
been d
they m
Their
the ba
profits
contes

BY **ANN SKEET**

WELLS
FARGO
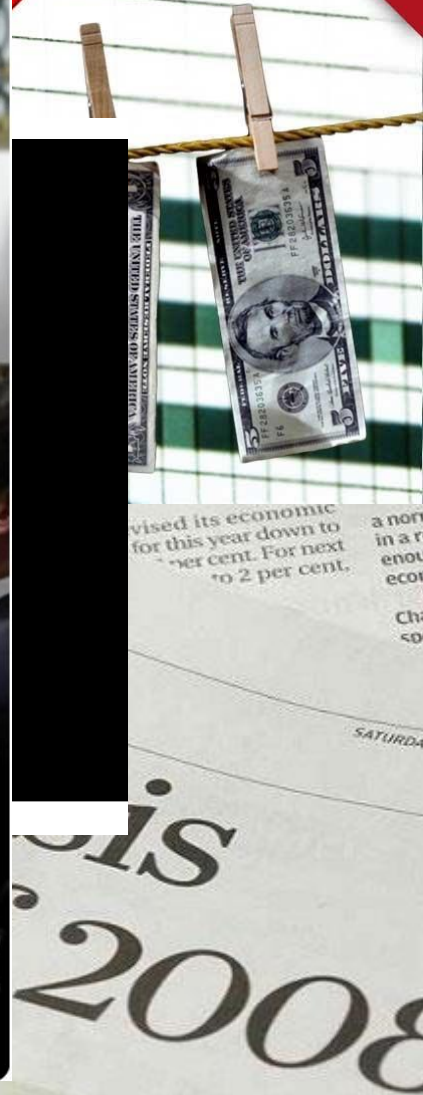
The door plug area of an Alaska Airlines Boeing 737 Max 9 aircraft, seen with the paneling removed, awaits inspection Wednesday at the airline's facilities at Seattle-Tacoma International Airport in SeaTac, Wash.

Lindsey Wasson/AP

Responsible, Trustworthy, Ethical AI?
VUCA 2.0

Amazon (2020-2022)

Issue: Amazon faced ethical scrutiny for several reasons:

AI Bias in Hiring: Amazon's AI recruitment tool was found to exhibit gender bias, discriminating against female candidates because it was trained on historical data skewed towards male candidates.

Labor Practices: The use of AI in warehouse environments for worker surveillance, strict productivity quotas, and decision-making regarding worker terminations raised concerns about labor rights, excessive monitoring, and unfair treatment.

Responsible, Trustworthy, Ethical AI?
VUCA 2.0

Clearview AI (2020)

- Issue: Clearview AI's facial recognition software backlash

billions of images scraped without users' consent then sold to law enforcement agencies, raising concerns about privacy violations, lack of consent, and potential misuse for surveillance purposes.

- Impact: This sparked debates over privacy, surveillance, and regulation of facial recognition technology, leading to legal actions in countries like Canada and European nations.

Responsible, Trustworthy, Ethical AI? VUCA 2.0

Google (2020-2021)

- Issue: Google was criticized for the handling of ethical AI research and the departure of prominent AI ethics researchers, including Dr. Timnit Gebru. Gebru's firing brought to light internal disputes about the limits of research on AI ethics and potential biases in AI models.

- Impact: This incident raised concerns about corporate influence over AI ethics research, freedom of speech in research, and the transparency and accountability of AI development within tech giants.

Responsible, Trustworthy, Ethical AI? VUCA 2.0

Facebook (Meta) (2021)

- Issue: The "Facebook Files" leak, revealed by whistleblower Frances Haugen, showed that Facebook's algorithms promoted divisive and harmful content to increase user engagement, despite internal knowledge of the potential harm to society, mental health, and political discourse.

- Impact: This led to public outcry and calls for regulatory reforms around algorithmic transparency, ethical social media practices, and the protection of vulnerable users.

Responsible, Trustworthy, Ethical AI?
VUCA 2.0

Apple (2021)

- Issue: Apple faced controversy over its proposed "child safety" feature that involved scanning users' photos on their devices to detect child sexual abuse material (CSAM). Privacy advocates criticized this move as a form of mass surveillance that could be misused or repurposed in the future.

- Impact: After widespread criticism, Apple delayed the rollout, highlighting the delicate balance between privacy and security in technology development.

Responsible, Trustworthy, Ethical AI? VUCA 2.0

TikTok (2020-2022)

- Issue: TikTok has been under scrutiny for several ethical issues, including:

- Data Privacy: Concerns over data collection practices, including the collection of biometric data without clear consent.

- Algorithmic Bias: The platform's content recommendation algorithms were found to amplify certain types of content, leading to concerns about the impact on mental health, misinformation, and potential political influence.

- Impact: Governments and regulators in multiple countries have questioned TikTok's practices, leading to investigations and potential regulatory actions.

WRONG ANSWER: MUCH BIGGER ISSUES

Outcome.

.05% of this AI business book addresses ethics and culture

Reinforces a glib notion of "ethics"

Ignores the cultural pitfalls of change- 70% failure rate

Some serious FOMO pressure- competitors/employees

ETHICAL

Moral principles and practices

...guidelines,...

No Bias

How is it possible to adopt, implement, regulate, measure, & monitor Ethical, Responsible, Trustworthy AI on a global level?

Sustainable

Transparency

Fairness

Privacy & Data Protection

Accountability

Ethical Considerations

Kitchen Sink

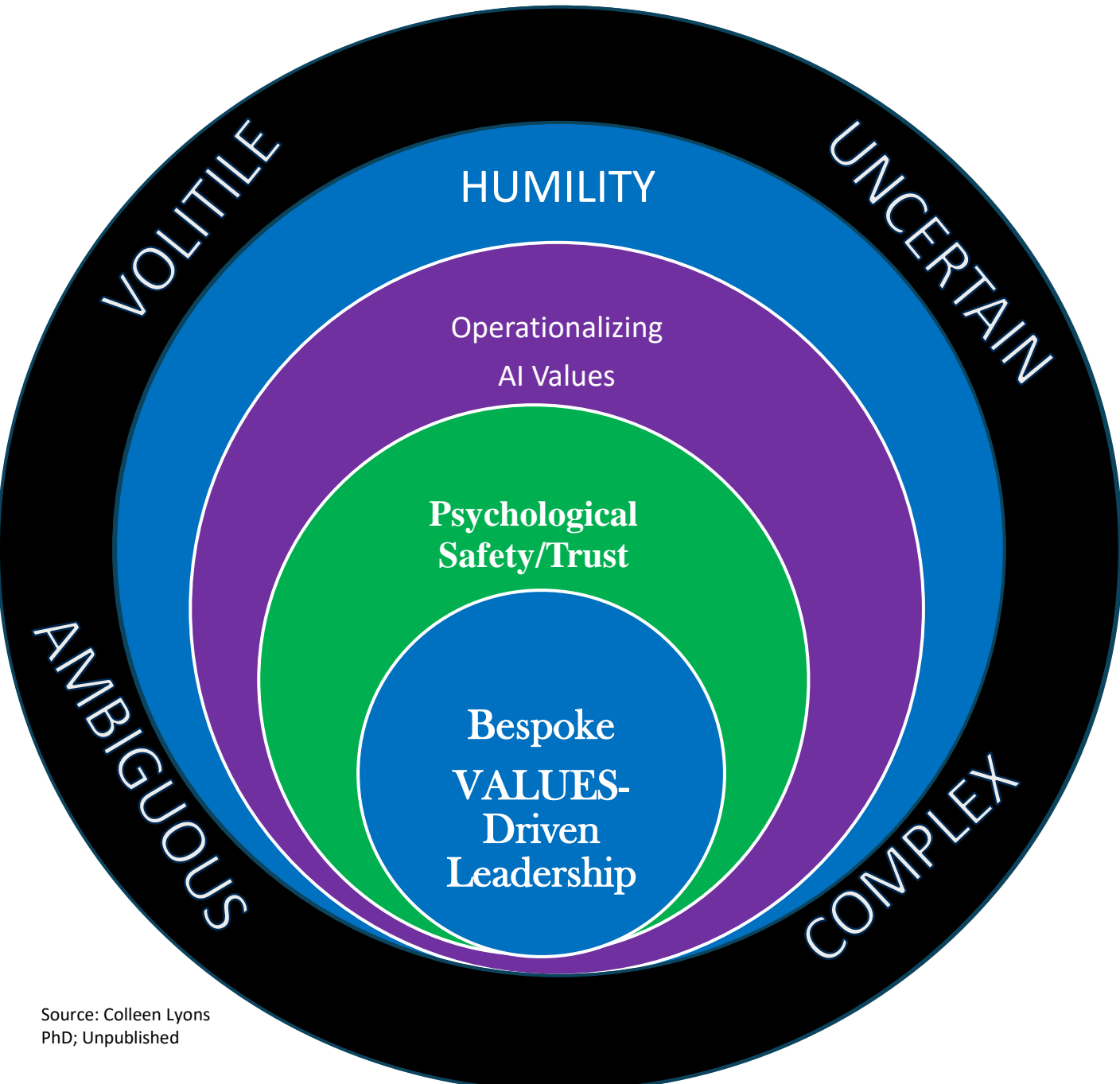# Trustworthy/Responsible/Ethical AI Word Salad

What do we do now?

GOOD AI

Source: Colleen Lyons PhD; Unpublished

Humility
Humility is a quality characterized by a modest view of oneself, an acknowledgment of one's limitations, and a willingness to learn from others.

Operationalizing Good AI
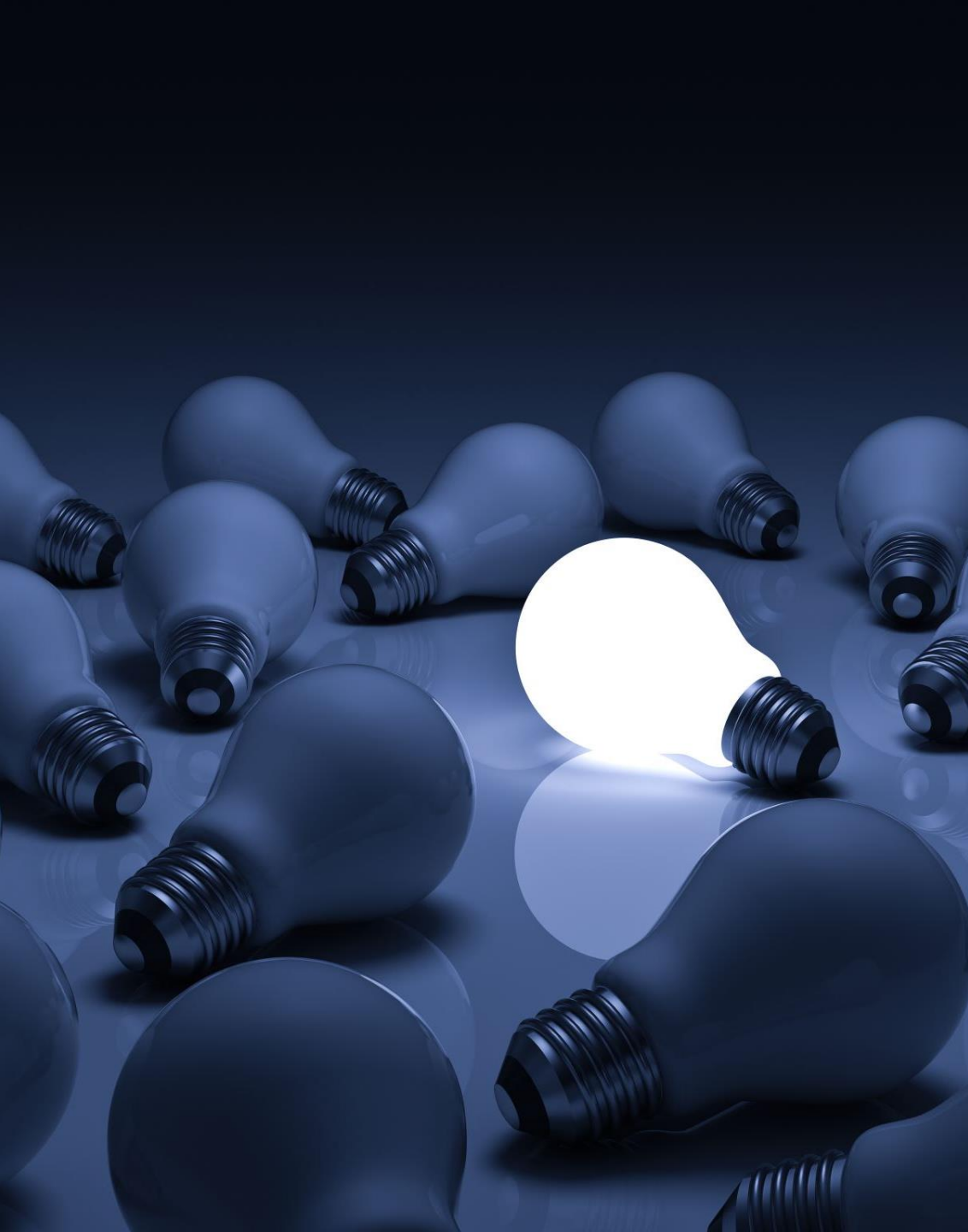**Apply equal testing and development rigor to AI values**

Psychological Safety
Ability to actively innovate/critique/challenge without fear (Team level)

Trust
I will not be harmed for speaking up or engaging in intelligent failure (individual level)

Bespoke Values-Driven Leadership
*Democratically determined values in collaboration with ALL employees*

Humility

By far, the greatest danger of artificial intelligence is that people conclude too early that they understand it.

Elezer Yudkowsky, Research Fellow

Machine Intelligence Research Institute

Research- Humility positively effects business outcomes

- Humility at the leader & team level positively influences
- Team performance,
- Innovation
  - Market
  - Problem orientation
- Levers
  - Learning Environment
  - Psychological Safety
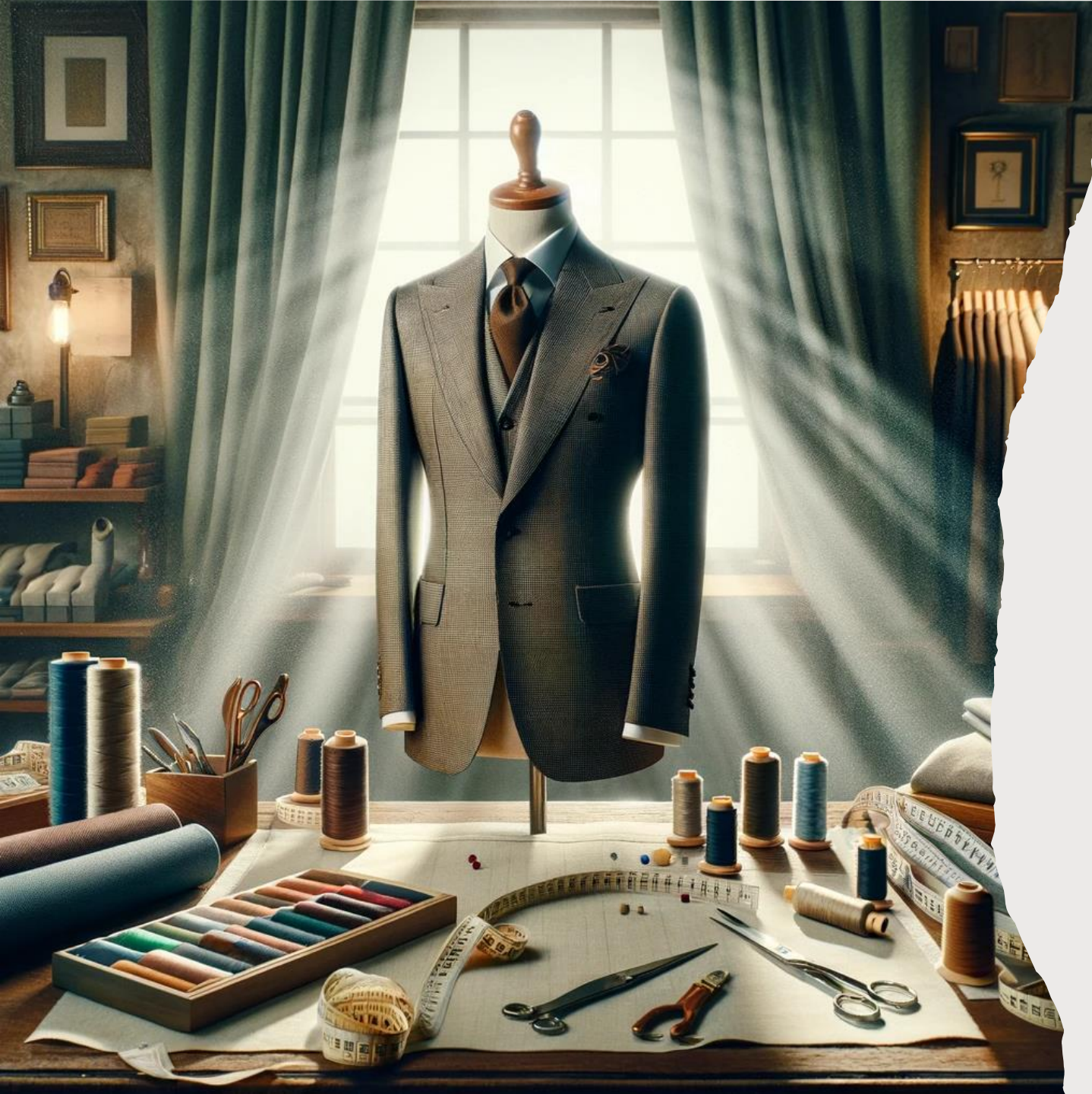
- chatGPT4 Image of Bespoke Suit
- craftsmanship
- attention to detail
- customization
- creating a personalized garment

Bespoke Corporate Values

4.**Perseverance** - Persisting in the face of challenges and setbacks.

5.**Respect** - Treating others with courtesy, dignity, and professionalism.

6.**Innovation** - Pursuing new creative ideas that have the potential to change the world.

These values can vary widely from one company to another, reflecting their unique culture and strategic goals.

Bespoke values that are democratically determined and lived by enable Good AI.

Outsourcing values is a bad idea.

How do we achieve good AI?

Implementation Science

•The foundational layer of operationalizing G&R AI is for organizations to adapt implementation science principles for AI development and implementation. How can implementation science help implement an AI strategy?

- **Framework for Effective Implementation**
- **Stakeholder Engagement**: *importance of all stakeholders early in the process to ensure their needs and concerns are addressed, thereby increasing buy-in, and reducing resistance to change.*
- **Identification of Barriers and Facilitators**: helps identify *potential barriers to the adoption of AI strategies (e.g., technical limitations, lack of skills, organizational culture)* and accelerate implementation (e.g., leadership support, incentives, training programs...*tailor the implementation approach to the* specific context of an organization.
- **Evidence-Based Strategies**:
- **Monitoring and Evaluation**:
- **Scalability and Sustainability**: Scaled up and sustained over time integrating AI into organizational routines.
- **Policy and Regulatory Compliance**: navigating the policy and regulatory landscape develop strategies that are both compliant and effective.

Giving Voice to Values

GVV is an innovative approach to values – driven leadership

Proven Success

Founded by Dr. Mary Gentile, former HBS, Benedictine Visiting Scholar

Voice and Enact Your Own Values

1,200+ Organizations

Differentiation: "How"

Perceived Choices
- Do Nothing
- Emotional Confrontation
- Retaliate
- Disengage
- Leave
- Report through formal channel: Ethics, HR

Implementation
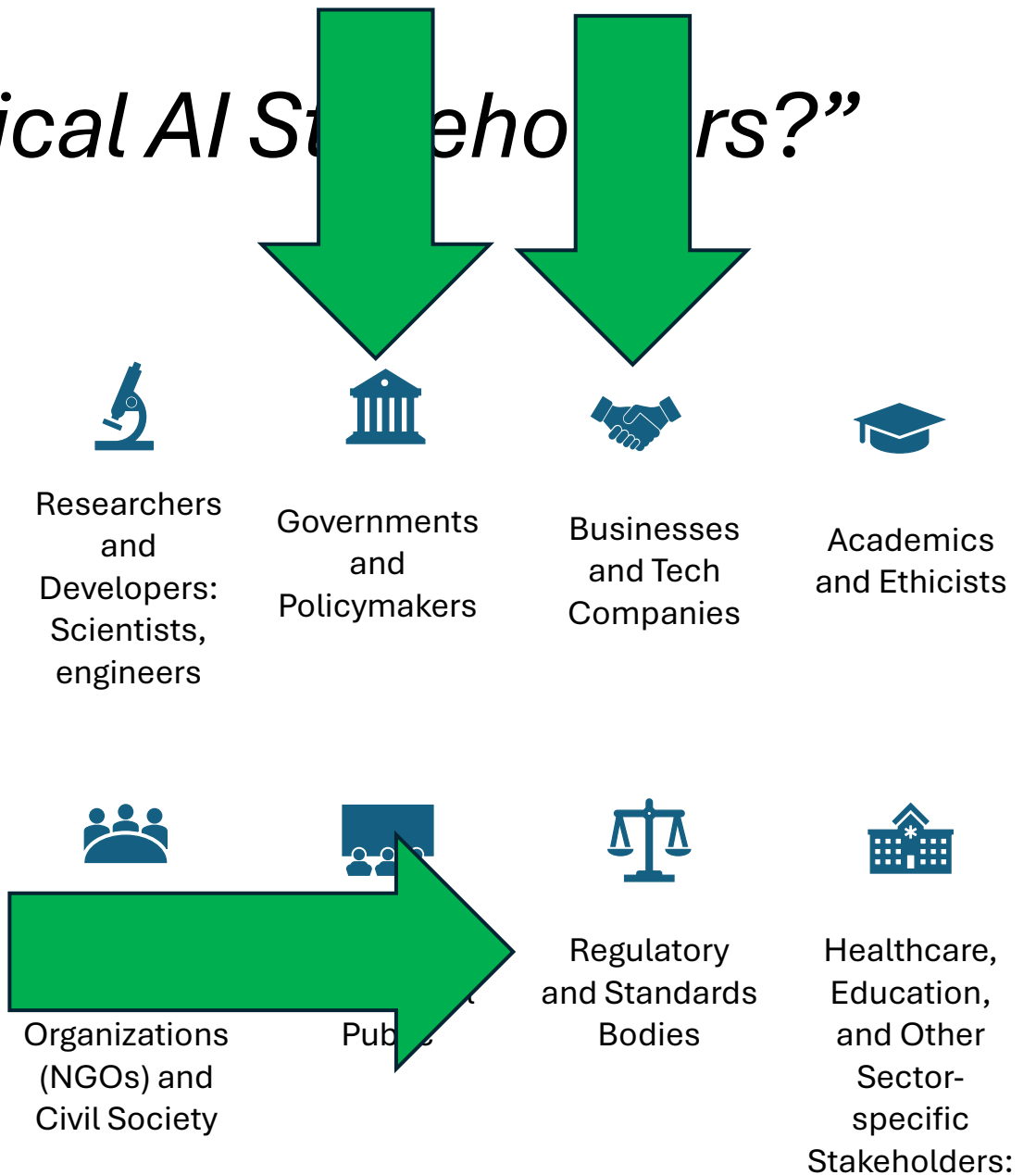
7 pillars

Ethical Muscle Memory

Normalization

GVV UPSKILLS LEADERS & EMPLOYEES with "*HOW*" TO ENACT THEIR VALUES

chatGPT4o *"Who are Ethical AI Stakeholders?"*

Ethical AI stakeholders are individuals, groups, or organizations that have an interest in ensuring that artificial intelligence (AI) is developed, deployed, and used responsibly, fairly, and in ways that align with societal values.

Researchers and Developers: Scientists, engineers

Governments and Policymakers

Businesses and Tech Companies

Academics and Ethicists

Organizations (NGOs) and Civil Society

Public

Regulatory and Standards Bodies

Healthcare, Education, and Other Sector-specific Stakeholders:

What's Your Legacy?