

# Humanizing the Future of Artificial Intelligence

A Toolkit for Ethical Artificial Intelligence Leadership

**Shelbi Howard** | Senior Thesis | April 29, 2020



# Contents

Problem Research	3
Stakeholder Research	13
Solution Development	22
Final Solution	33
Impact	75
Project Resources	81



# Technology lacks oversight.



Technology is diminishing social norms and human interaction.



Mental health decline aligns with mass adoption of smartphones.



Tech companies compete for humans' *time* not their *well-being*.

## Humanity is being affected by unethical technology outcomes.

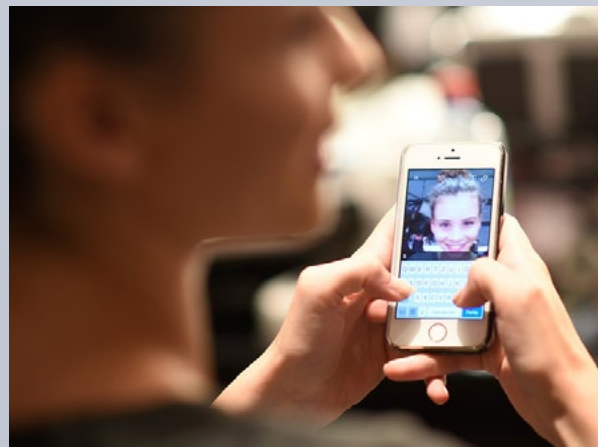


# Humanity Disrupted

Empathy is being reduced as impersonal technology becomes humans primary communication method. The kids of today are the adults of tomorrow and something about the technology needs to change before we destroy our humanity forever.



Generation Z (12-22 years old) is losing its ability to read non-verbal communication as a result of lacking physical interaction with others throughout social development.



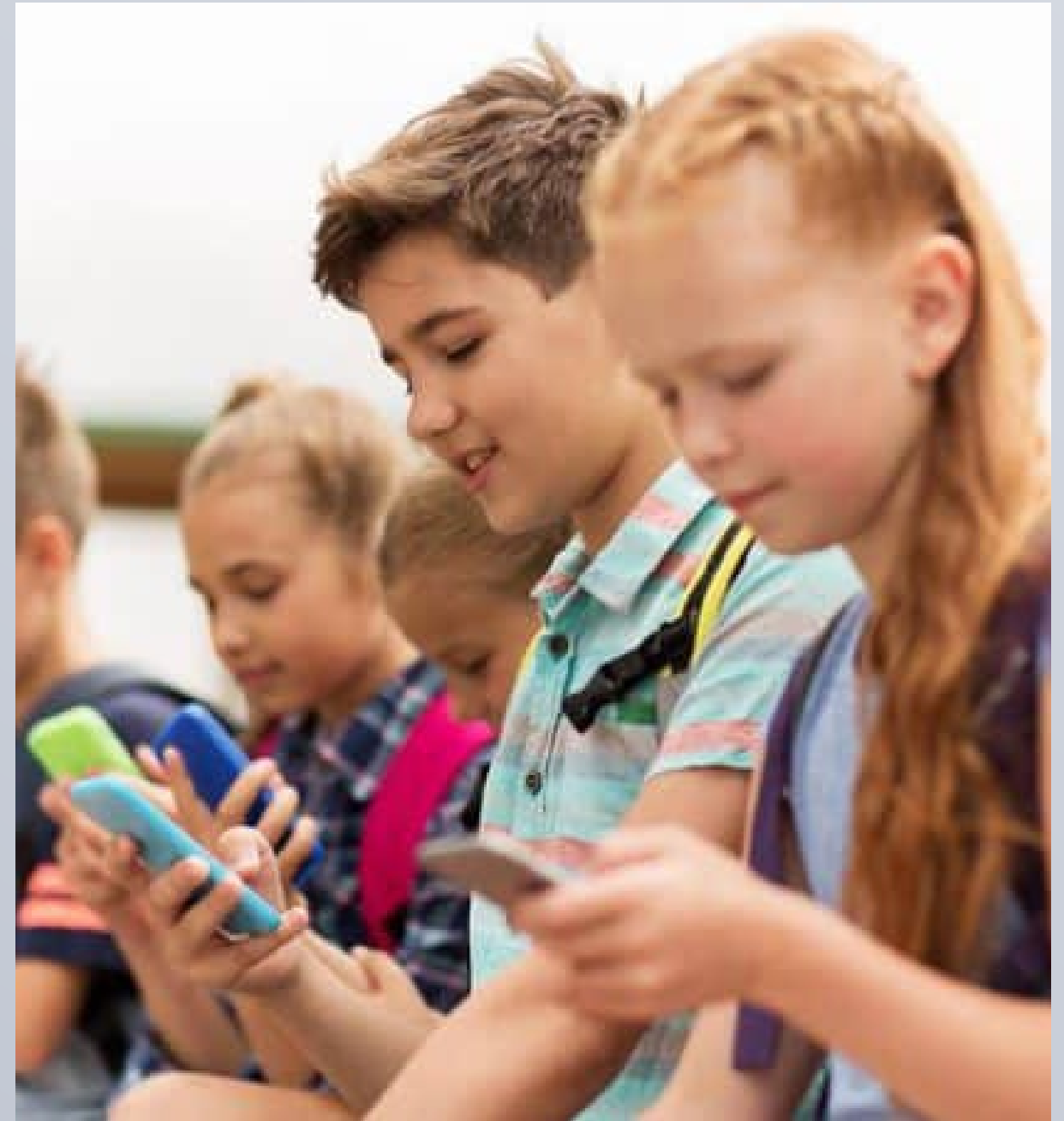
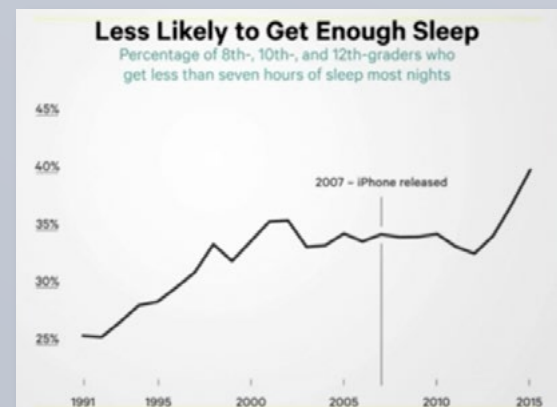
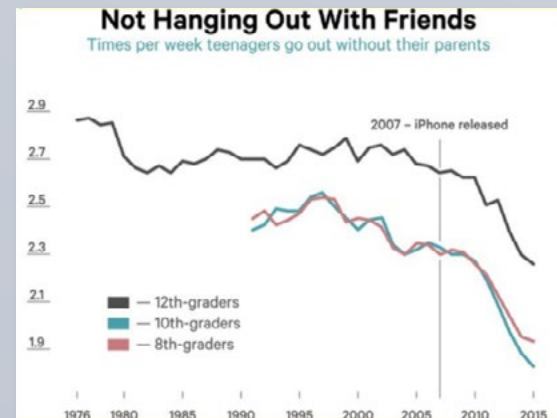
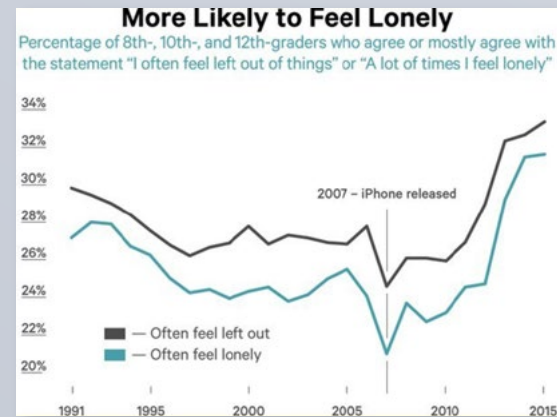
Attention has become an important human resource. At any given time today, one person has their attention split an average of 5 ways. This is affecting how we learn, sleep, feel, and behave.



Researchers have been releasing alarming statistics on a sharp and steady increase in kids' mental illness, which is now reaching epidemic proportions: 1 in 5 children has mental health problems, 43% increase in ADHD, 37% increase in teen depression, 100% increase in the suicide rate in kids 10-14 years old

# Impact of Personal Tech.

Technology today does not have human's best interests in mind. The first generation of personal tech natives have seen declines in mental health as a result of their technology usage. Creators of technology need to take responsibility for this health crisis and refocus personal technology around human health needs.



Statistics from a study conducted by Dr. Jean Twenge, author of *iGen*, show a stark change in generational norms after the release of the first "smart phone" leading to a decline in socialization and a surge in childhood loneliness from 2007 to present. In 2011-2012, those having iPhones when over the 50% mark.



# Humanizing Technology

Machine learning and artificial intelligence is making it possible for technology to become increasingly more human. While it is possible to make machines sound human, companies and consumers are beginning to question the ethics of how far these humanized machines should go.



Our devices are becoming multi-purpose, anticipatory, and responsive, but how do we ensure they continue working with us not against us?



Society is trying to solve issues caused by technology with more technology. AI conversation bots focused on therapy are emerging for millennials to have 24/7 access to mental health support.



Technologists are striving to make AI more human than humans in how it thinks, responds, and perceives information. However, users are questioning their comfort level with having a conversation with a robot when they believe it's a human. An example of this is Google Duplex, above, which holds a phone conversation while remaining indistinguishable from another human.

# Direct Social Influence

Creators of AI are at risk of building their bias into the machines they create. These implicit foundational biases accelerate social issues and lead to widespread cultural acceptance of biased norms. As a result, AI has the power to directly influence society's ethics.

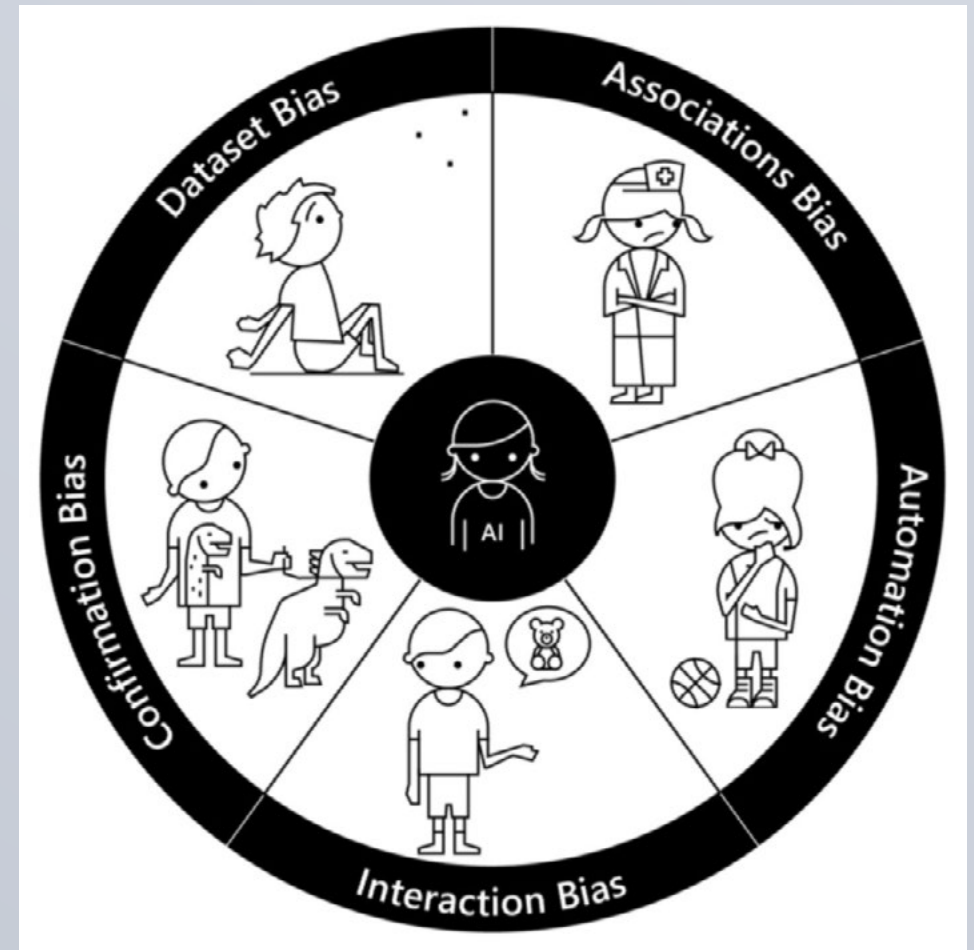


Current relationships between AI and customers will need to be re-evaluated as devices become more autonomous. Current technology relationships are beginning to reinforce slave-master dynamics with subservient identities, lack of boundaries, and limitless options. This is problematic as voice interactions act as a model for how humans treat one another.

i keep saying Alexa when I mean to say Siri and i just cant believe i live in a time where i am gettin my servant robots names mixed up

11:19 AM · 1/9/19 · [Twitter Web Client](#)

121 Retweets 970 Likes



Microsoft's Inclusive AI team has identified 5 forms of bias for AI creators to be aware of. These address issues such as association bias where device identities create subconscious sexism and dataset bias where voice recognition priorities lead to indirect racism and marginalization of certain groups.



## AI Limitations

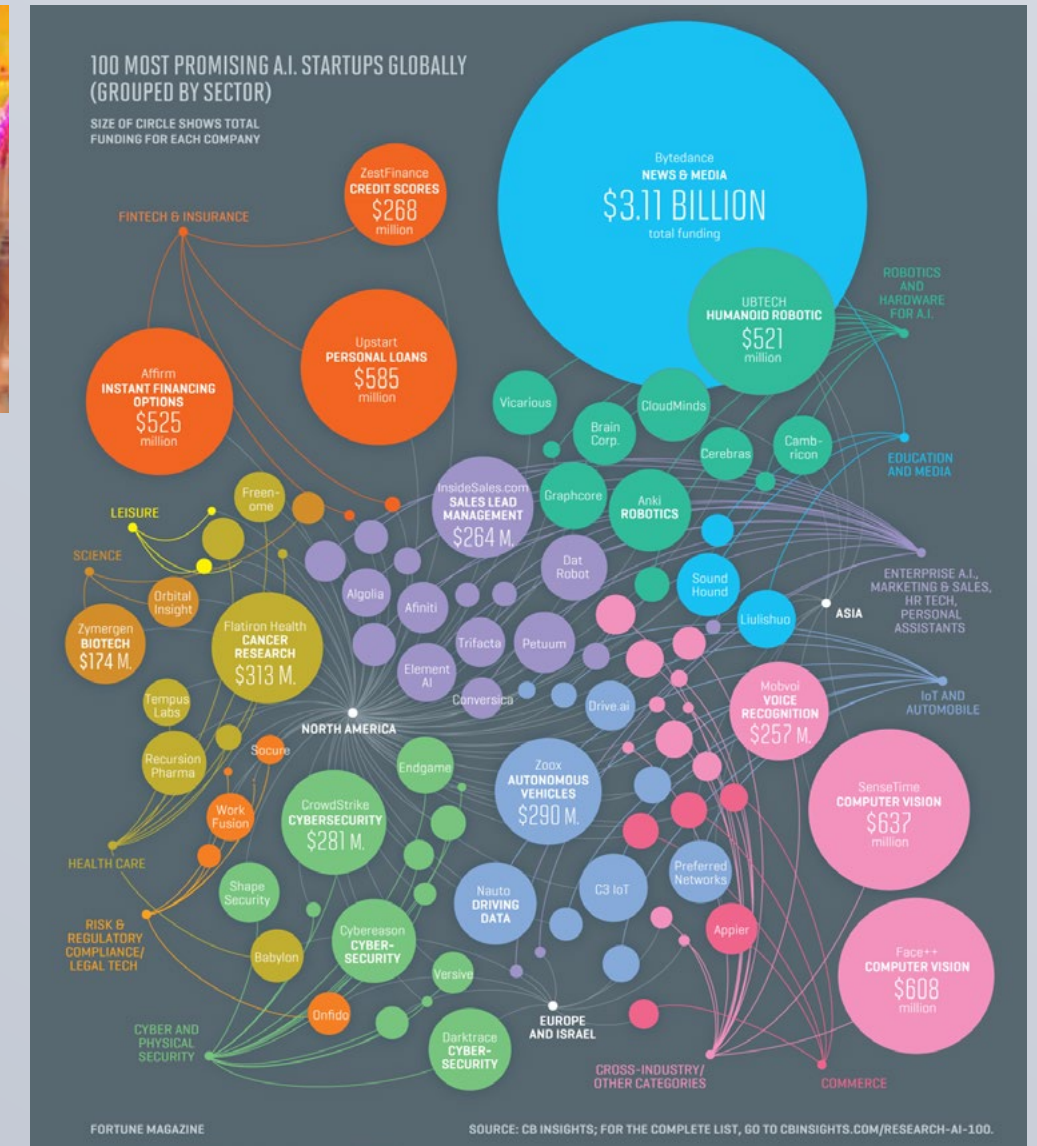
Consumer device AI is learning fast but lacks the context it needs for widespread acceptance of the technology. Industry decisions now will determine the success or failure of consumer AI in the future.



AI is black and white: assumes truths, can't grow, can't create new knowledge. Microsoft's Teen Bot Tay.ai shows mob mentality at its worst by operating off of "assumed truths" shared on Twitter.



Individuals and companies responsible for sourcing, testing, and implementing AI datasets control the spread of knowledge, truth, and information sharing throughout society.



By producing in a society of constant product competition, companies don't share their datasets leading to slower development in AI growth and product innovation. This causes biases to be more prevalent and limits ethical AI decisions to a company level instead of an industry-wide discussion.

# AI Ethics Emerges

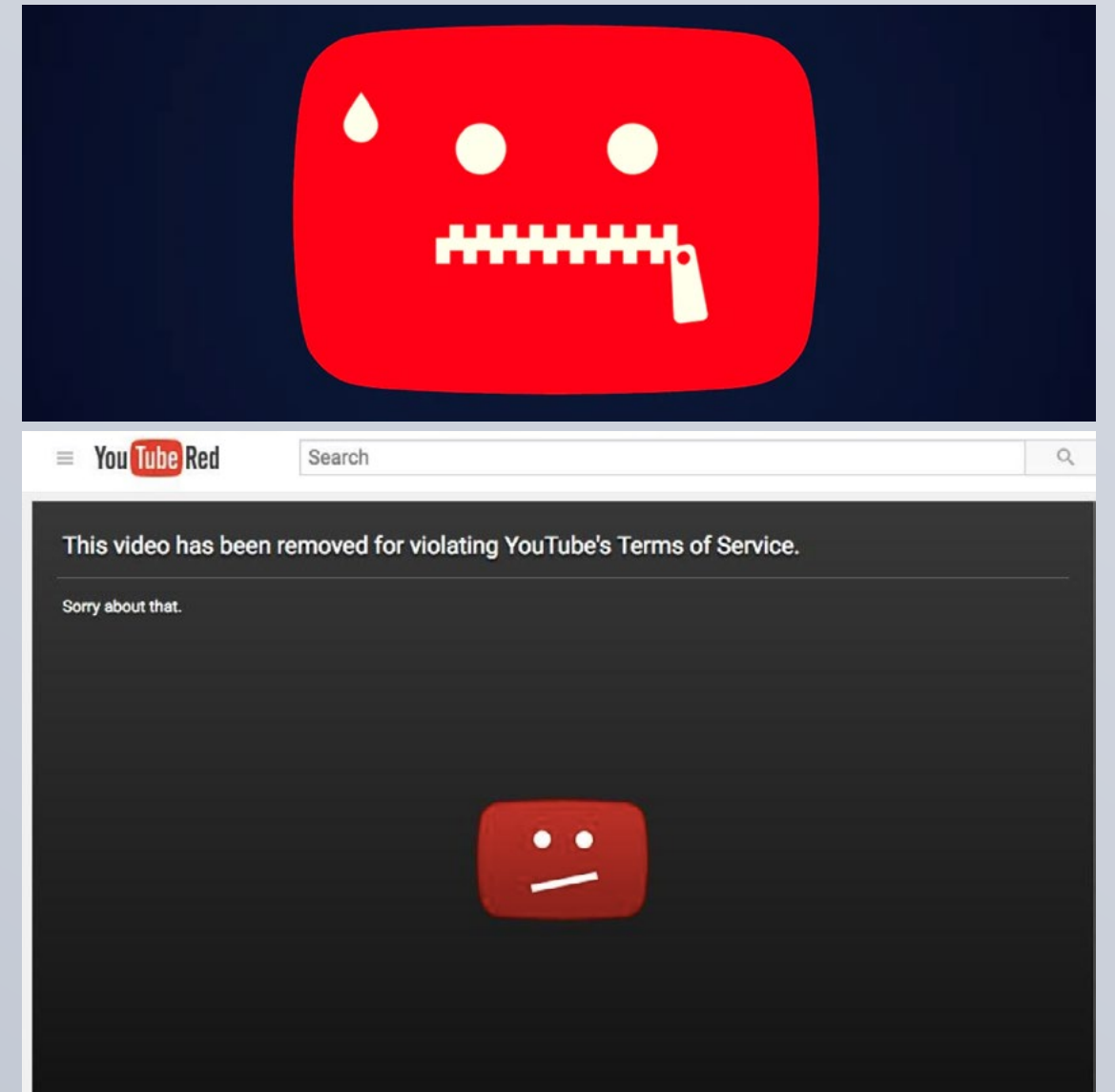
Companies are beginning to prioritize ethical AI decision-making in technology design but still struggling. Building implicit morality into a product is inherently an ethical issue in itself. To avoid technocracy, these teams should be diverse and democratic.



Google announced an AI ethics board in April 2019. Within a week, the group was cancelled as a response to outcry from employees with the comment that they were "still working on getting it right."



Microsoft faceted their AI ethics focus into multiple teams. Their Inclusive AI Team is composed of leaders from across the company to focus specifically on ensuring ethical interactions in regard to user privacy and bias.



Google's company, YouTube, has been making policy decisions after the 2016 elections to ensure the same outcome doesn't happen again. This is manifested in a company imposing policy about content based on political leanings leading to ethical questions about their effect on free speech.



As artificial intelligence makes technology more impactful,  
AI will require **ethical infrastructure development and oversight**  
to ensure the future survival and well-being of humanity.

As artificial intelligence makes technology more impactful,  
AI will require **ethical infrastructure development and oversight**  
to ensure the future survival and well-being of humanity.

**It is imperative that we begin prioritizing ethics in AI today.**



User interviews, A/B testing, and qualitative research was conducted with **business, policy, technology, and design experts** to determine the best solution for an ethical AI future.



**Richard Harknett**

Head of Political Science at the University of Cincinnati. One of the world's leading online security experts and scholar-in-residence for the DoD U.S. Cyber Command.



**Claudia Rebola**

Information design PhD providing research support for intelligent systems, human-AI interactions, and communication throughout design and development.



**Alexander Motz**

Computer engineer and serial tech entrepreneur providing business and technology insight on the functional applications of artificial intelligence.



**Sam Lowe**

Designer and brand strategist creating stories for businesses to reach customers. Offered support on creative strategy, copywriting, and leadership communications.



**Derek Shewmon**

Serial technology entrepreneur advocating for small business needs and entity engagement in ethical AI solutions.



**Erik Sheagren**

Cognitive Science PhD at Columbia University focused on the Theory of Mind. Responsible for providing philosophy and ethics support for human-AI interactions.

# Ethical AI requires support from government



## Governing Bodies

Government, public and private organizations, research institutions

## Needs

- Expert guidance
- Citizen protection
- Proactive legislation
- Global leadership

## Challenges

- Subject-matter expertise
- Global competition
- Private tech industry bias
- R&D funding

## Solution

Education

Recommendations

Resources

## Opportunities

Become a **global leader** in the race for AI dominance.

Create sustainable, long-term AI governance to **protect humanity**.

Establish a **progressive vision** for the future of American innovation.

Make **educated** legislative decisions.



## Technology Industry

AI Technology Entities



## Technology Consumers

AI Technology Users



**Ethics in AI**  
unsettled debate  
→ must educate

govern vs innovate

**Myth of fragility**

Regulation will crush innovation  
→ more safety protocols on fridge than Apple OS

**2017**

Worked w/ AI @ Oxford

→ Safe vs. Secure

Inherent tension

Human is in control

Algorithm defends itself

**Killswitch**

core vulnerability

**NIST** Natl. Inst. Standards in Tech.

→ Framework - Standards certifications

**NICE** Natl. Initiative in Cyber Edu.

→ Educational govt. template

→ Shared responsibility model

**Government**  
Protects you

**Individual**  
"good hygiene"  
↓  
Avoid viruses

\*HIPAA - effective state intervention for privacy

"Who dominates AI dominates the world"

tech. mediation →

**Security by design**

- Accept security by default
- vs goal now: efficient & accessible

Decision Loops: tool for decision-making

① Automate out

**Human** → **AI**  
\* autonomous

② Human is in

**Human** ↔ **AI**  
\* Don't automate

③ Human is on

**Human** → **AI**  
\* semi-autonomous

AI Decision Loops

**OODA**

- observe
- orient
- decide
- Act

traditional decision loop

**Architecture is vulnerable**

→ can't layer AI on top of this

\* more central, powerful, and independent

Ethics: Communal public good

**AI actors**

- China - lead by 203
- Alphabet (Google)
- US / Israel

\* Need legally enforceable ethical development laws



Richard Harknett

Head of Political Science at the University of Cincinnati. One of the world's leading online security experts and scholar-in-residence for the DoD U.S. Cyber Command.



# Ethical AI requires support from businesses



## Governing Bodies

Government, Public and Private Organizations, Research Institutions



## Technology Industry

AI Technology Entities

### Needs

- Create profit
- Build customer base
- Make new products
- Innovation

### Challenges

- Gain new customers
- Keep existing customers
- Compete in market
- Meet regulations

### Solution

Assessments

Guide

Ethical AI Principles

### Opportunities

Improve **brand image** to help hire young talent, gain new customers, and keep existing customers happy

Create **innovative solutions** for new and existing products and services.

Develop responsible, sustainable solutions for the **good of humanity**.



## Technology Consumers

AI Technology Users

## Business Customers



BIG Companies

- more bureaucracy
- must already be a credible product
- more support if from govt. → certification?

↓  
Reputation (Brand)



small companies

- move faster
- keep cost low
- want good ROI

↓  
Presence + Momentum

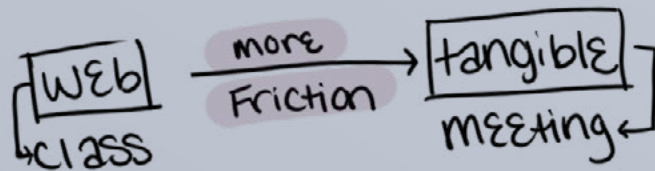
Important!

## Employee Support

For employee cert., edu., + training...

- consider minimum # of employees that need  
↳ ethical threshold

Don't set the bar too high



- \* Sell the exec that the employees are sold\*  
↳ people want it

## Industry

Product vs Service

Applications?  
70% are service-based

\* Define YOUR Focus \*

## Buy-In

- \* EXCLUSIVITY
- \* CREDIBILITY
- \* ↑ VISIBILITY
- \* Hire Gen Z - talent
- \* sell to Gen Z
- \* LEGITIMIZE company values

BUT at what cost?

want to get ahead of competition

As a business...  
**WHY SHOULD I CARE ABOUT ETHICS?**

## Incentives vs Penalties

- \* From govt.

## Solution

Are you selling this?

## Big Questions

- \* How long does it take?
- \* What's the cost?
- \* Make employees care beyond "it's a requirement"

## SaaS Packages

1	2	3
Silver	gold	Platinum
\$\$	\$\$\$	call us
# employees	"	Enterprise
Benefits	"	

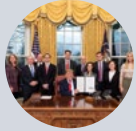
\* Pick your size focus \*



Derek Shewmon

Serial technology entrepreneur advocating for small business needs and entity engagement in ethical AI solutions.

# Ethical AI requires support from consumers



## Governing Bodies

Government, Public and Private Organizations, Research Institutions



## Technology Industry

AI Technology Entities



## Technology Consumers

AI Technology Users

### Needs

- Trustworthy technologies
- Businesses as advocates
- Tech to alleviate problems
- Healthy human connection

### Challenges

- Transparency in technology
- Finding reputable resources
- Lacking a human advocate
- Technology benefits itself

### Solution

Education

Ethical AI Principles

Resources

### Opportunities

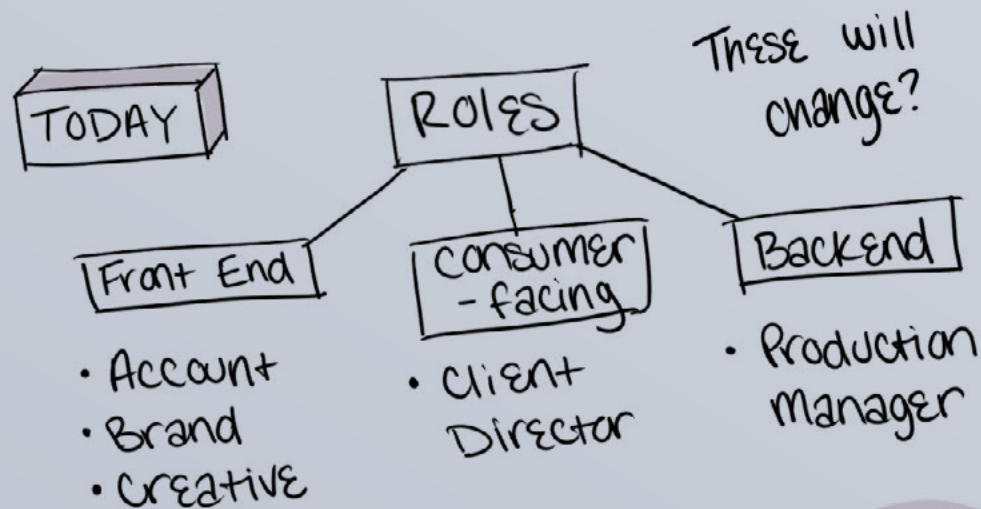
Create a **culture** that celebrates human well-being and connection.

Make **educated decisions** about technology usage.

Support businesses that actively contribute to the **greater good**.

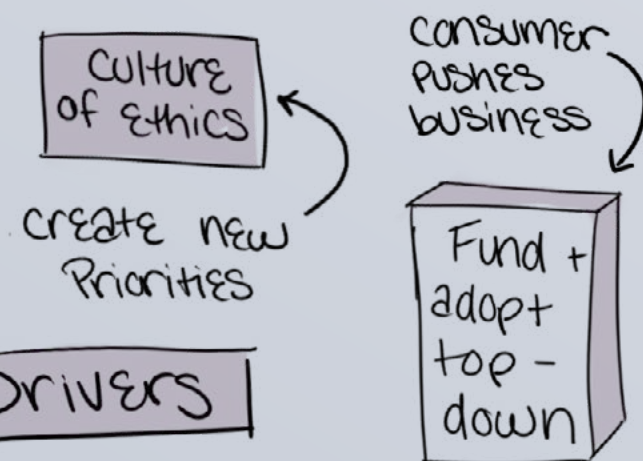
Ensure **survival** alongside AI.





What problem is this solving / why should I care?

- Grow market
- ↑ consumer advocacy



How do you avoid business resentment?

create TRUST with consumers

# ① Drivers

Today

- Capitalism → Profit

Tomorrow

- New methodology ← Solution

## Fact-Check Products

- For new + existing
- Apply new standards

\* Educate businesses AND consumers \*

**Market to consumers**

- Influence the Industry
- ↳ Podcast?

**Website**

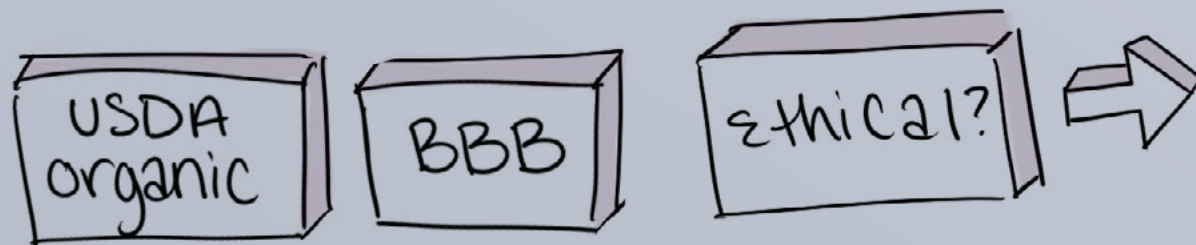
- Educate businesses
- certify here

# ② Opportunity

- Prove the Process
- \* Industry agnostic

# ③ Solution

- Certification
- ↳ Part of B corp
- Found OR Join



creates incentive for customers to buy



Sam Lowe

Designer and brand strategist creating stories for businesses to reach customers. Offered support on creative strategy, copywriting, and leadership communications.



+



+



**Ethics Education**  
Raise industry, government,  
and consumer awareness

**Industry Guidance**  
Standards and regulations  
to establish expectations

**Funding and Support**  
Government funding and top-  
down industry implementation

**= Ethical AI Future**





Wicked problems require collaborative solutions.

*Facilitating informed discourse on the future of AI was the solution I needed to create.*

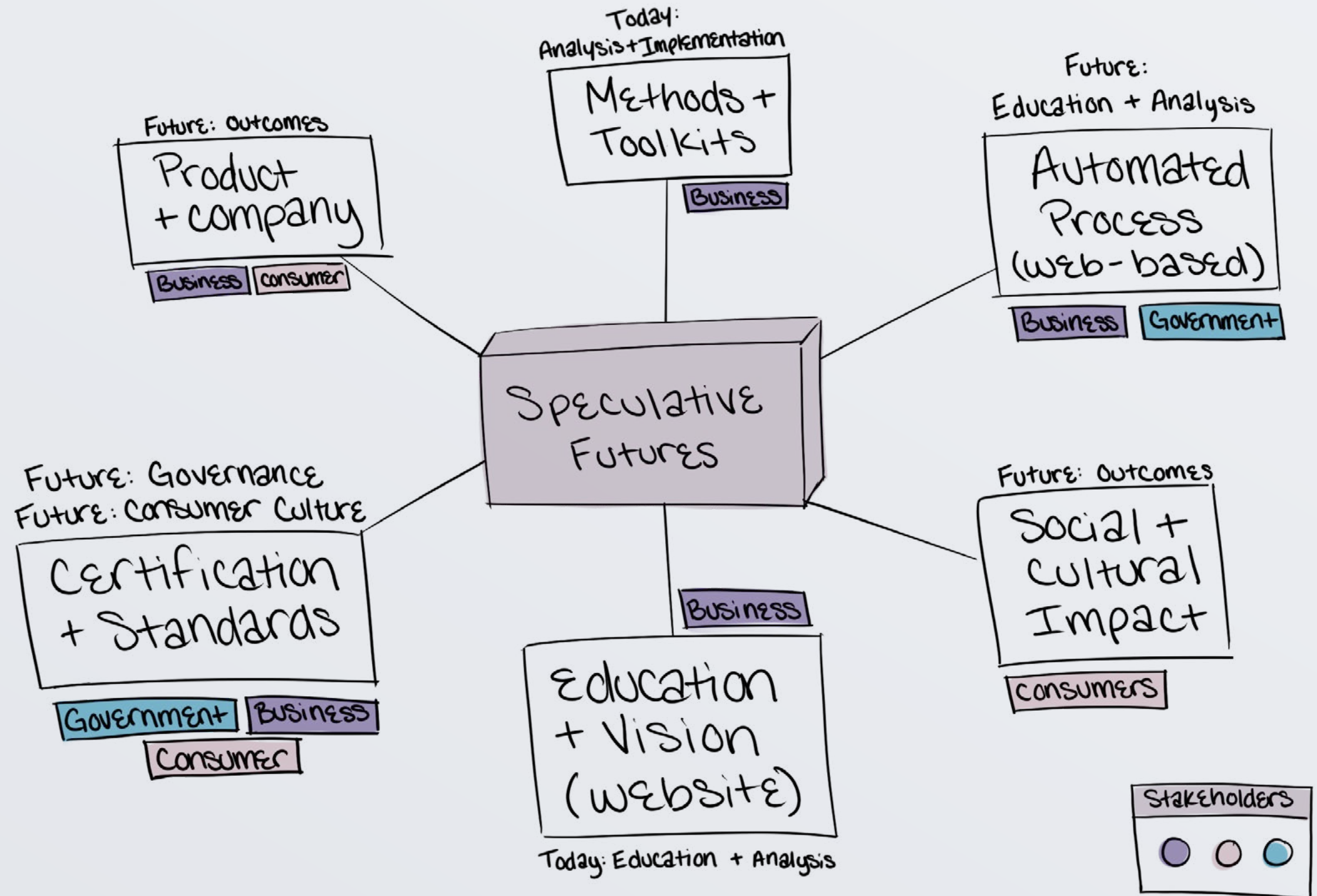
Information was designed by utilizing a **research through design methodology** to research, prototype, validate, and iterate solutions based on expert feedback.

The complexity of AI in business and policy required **extensive secondary research** to develop beneficial information for AI leaders and legislators.

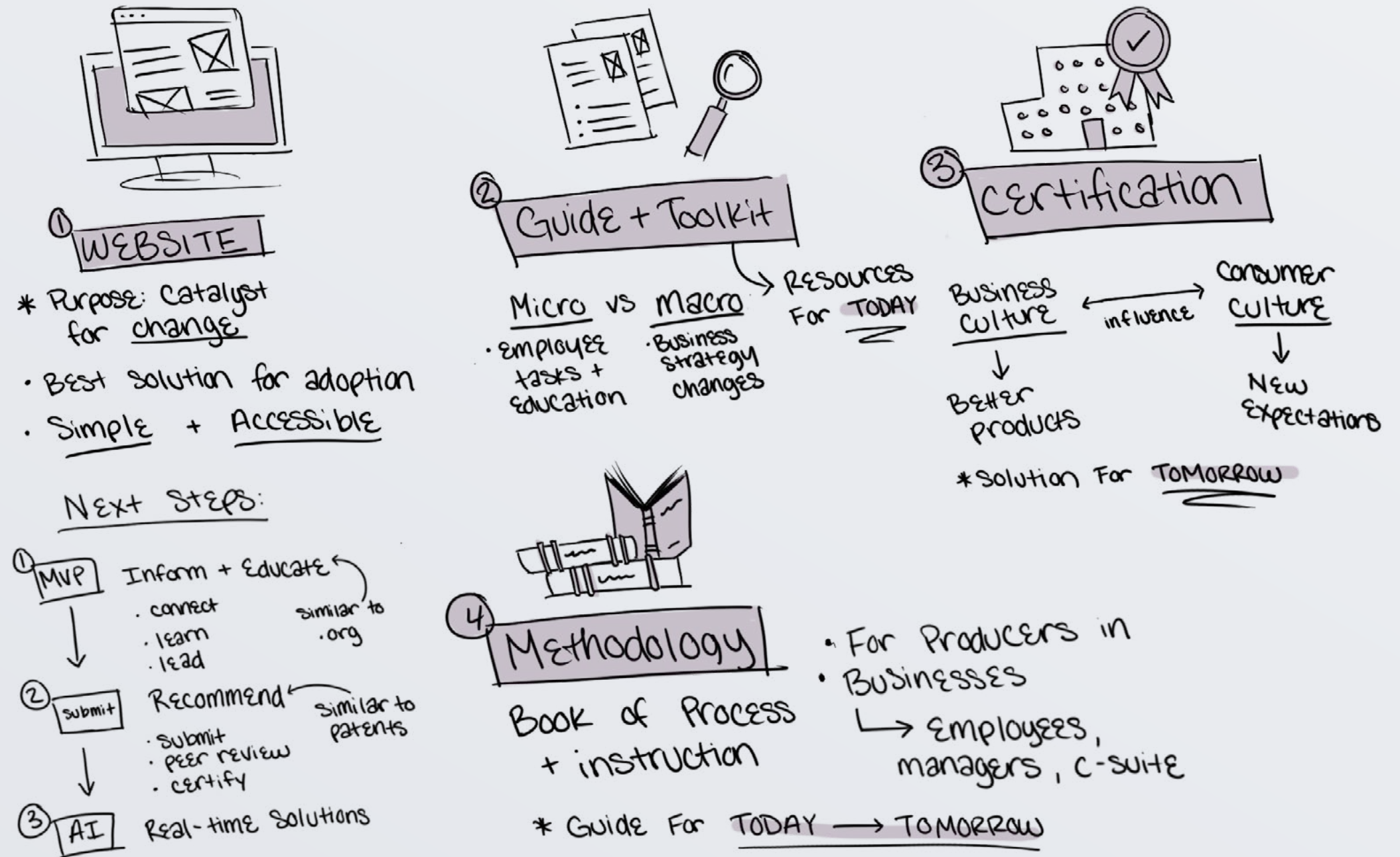
The problem was ambitious but I could contribute my design thinking and digital design skills to **integrate stakeholders into one place for discourse on the future of AI** and help them access the tools they needed to make these decisions.

The optimal final output was determined to be **an educational, integrated web-based toolkit** for leaders in business and government to find information about the importance of an ethical AI future.

Many concepts were ideated and solutions associated with stakeholders to **determine benefits and opportunities**.

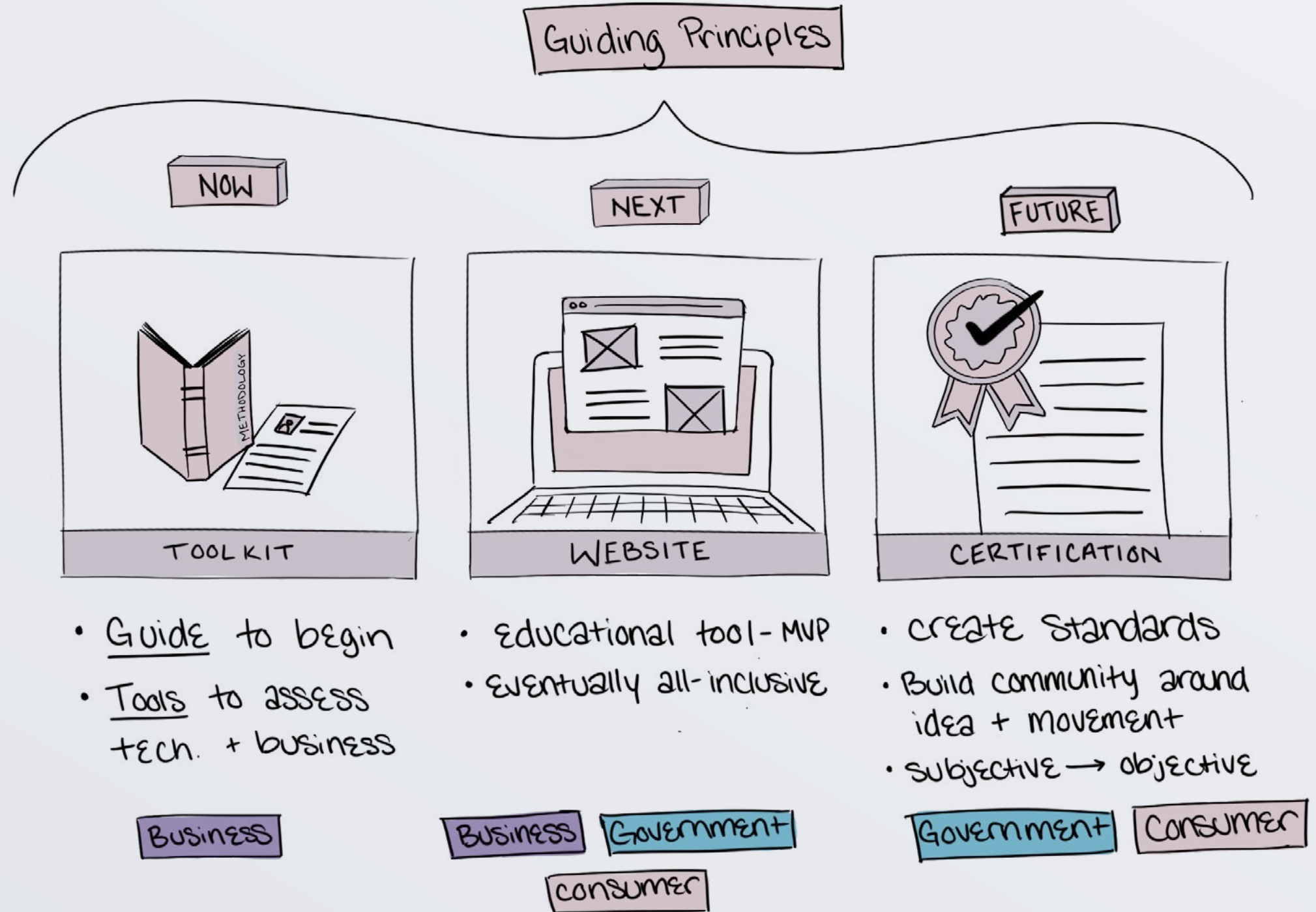


A select **few concepts** were explored further to discover the best solution.





Concepts were condensed further and scope was focused on **three primary deliverables**.

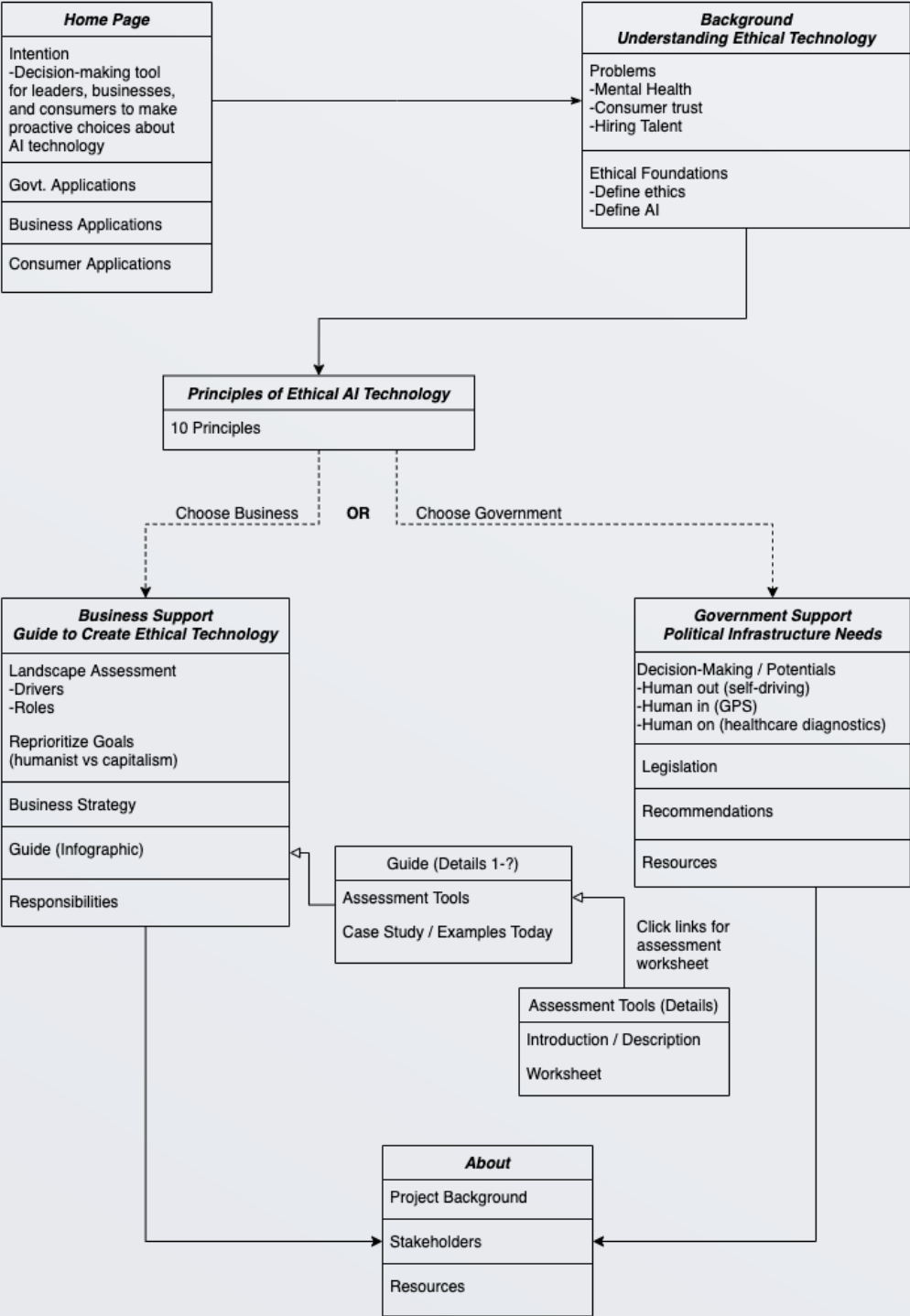


Tabs communicate **guiding principles and relevant tools** for users and stakeholders in industry and government.



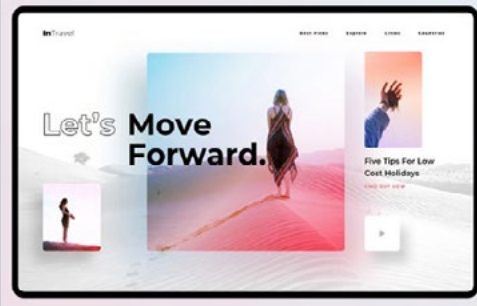
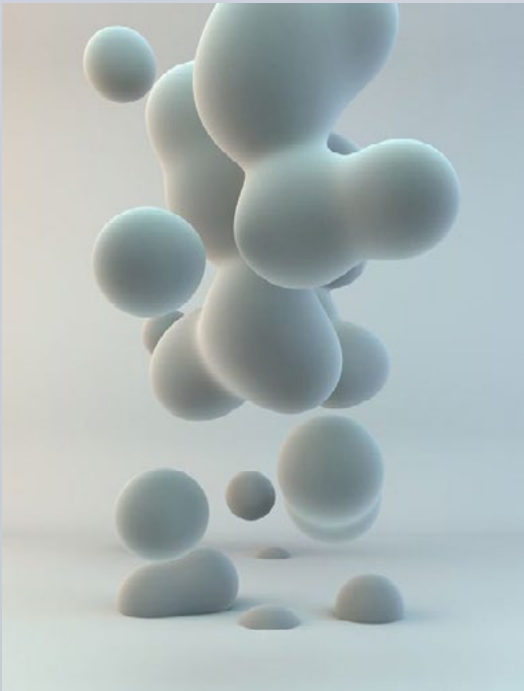
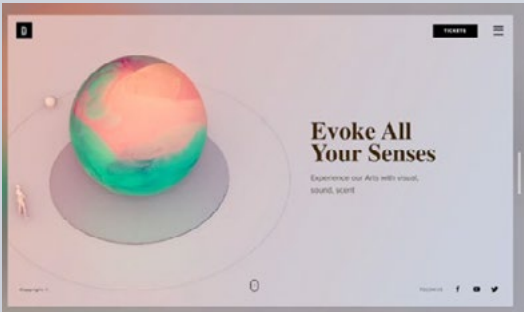
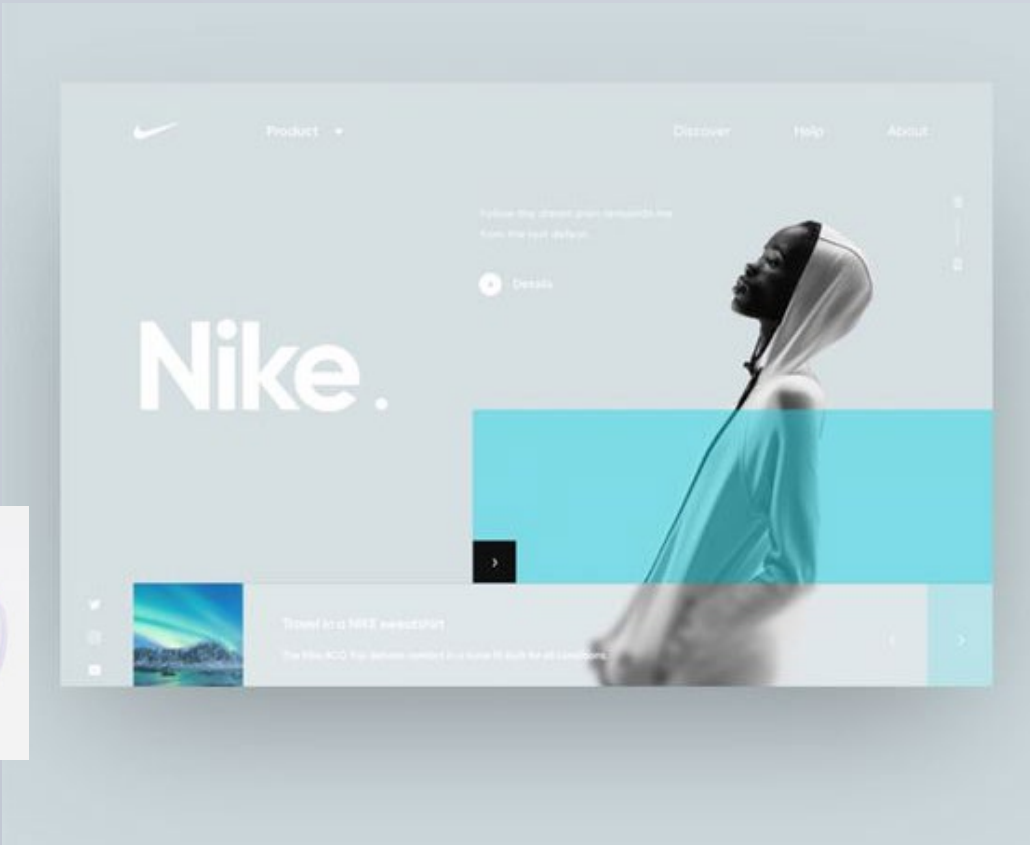
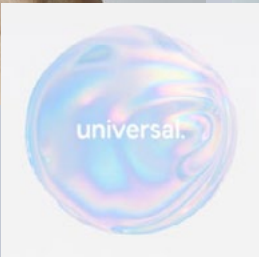
This educational **website** is the most simple, accessible solution to ensure the most visibility for all user groups.

The **website architecture** is optimized for the three user groups: government, business, and consumer.

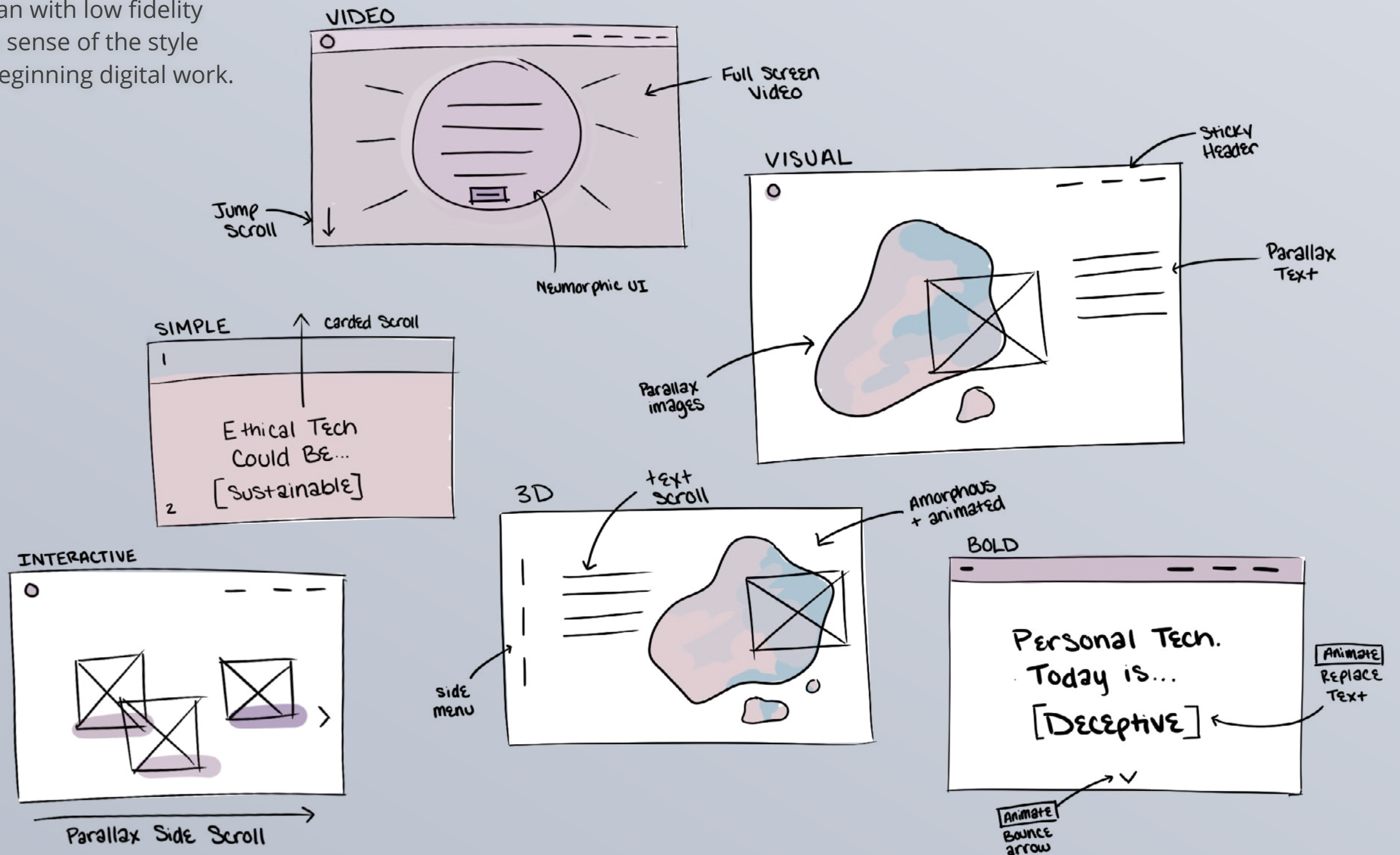




The website style was curated to provide users a sense of **transparency, trust, and innovation.**



Website design began with low fidelity **wireframes** to get a sense of the style and layout before beginning digital work.





# The website underwent many iterations and user feedback sessions over the 3-month research and design period.





The website branding utilizes color and transparency to depict the values of a transparent, human-centered future.



Header 1  
Open Sans Bold 48pt #ffffff

Header 2  
Open Sans Semibold 24pt #000000

Subheader 1  
Open Sans Semibold 24pt #3b456f

Button 1

Button 2

Link >



#aeb4c2 #ffffff #000000 #3b456f #3d4972 #35a0da  
#dee2e9 60% opacity #6a87a9 #78c7e9

SUBHEADER 2  
Open Sans Semibold 16pt #ffffff

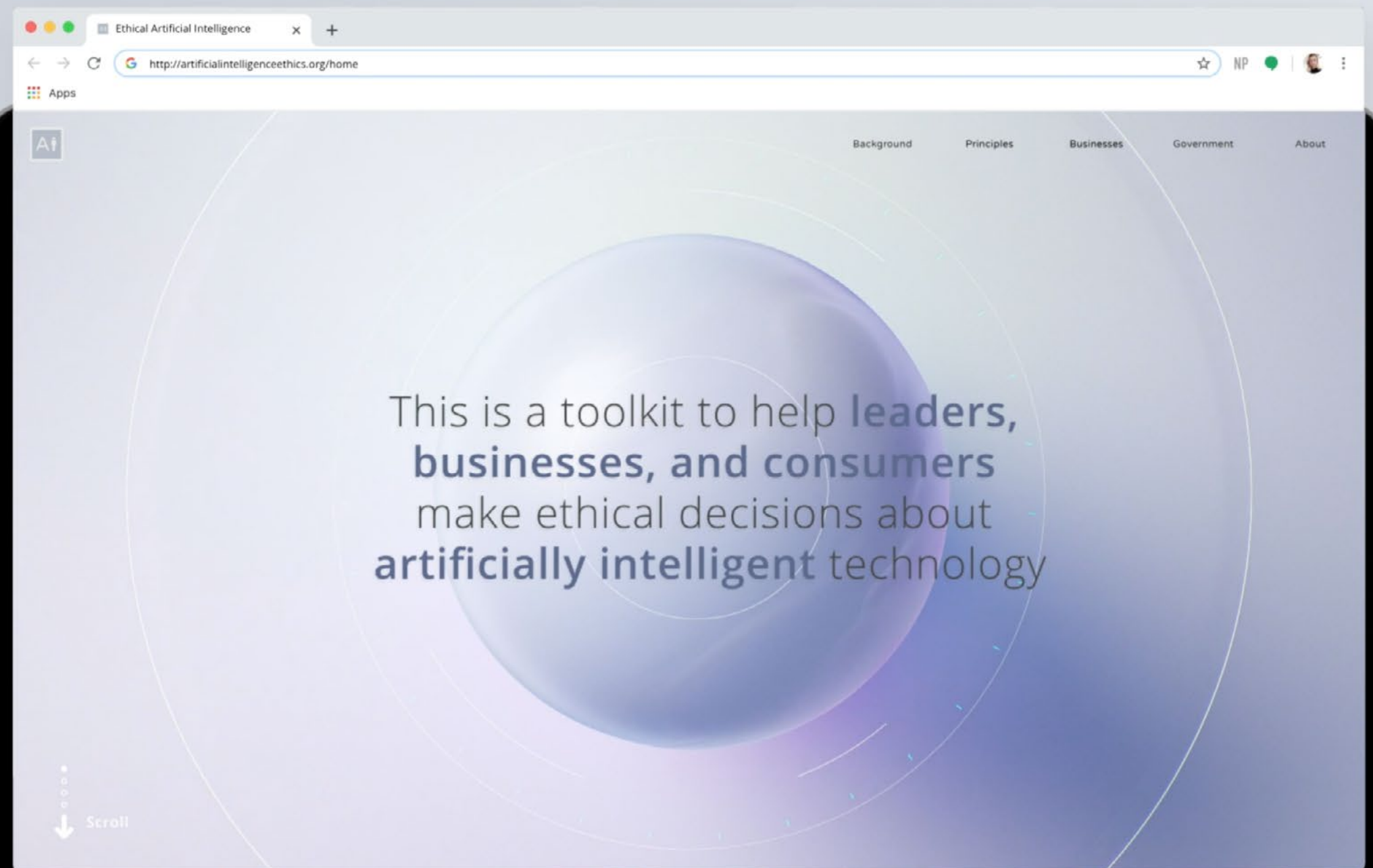
Body  
Open Sans Regular 18pt #000000

Links  
Open Sans Bold 21pt #3b456f





The **home** page guides users through the purpose of the website and how to get the most out of it based on who they are.

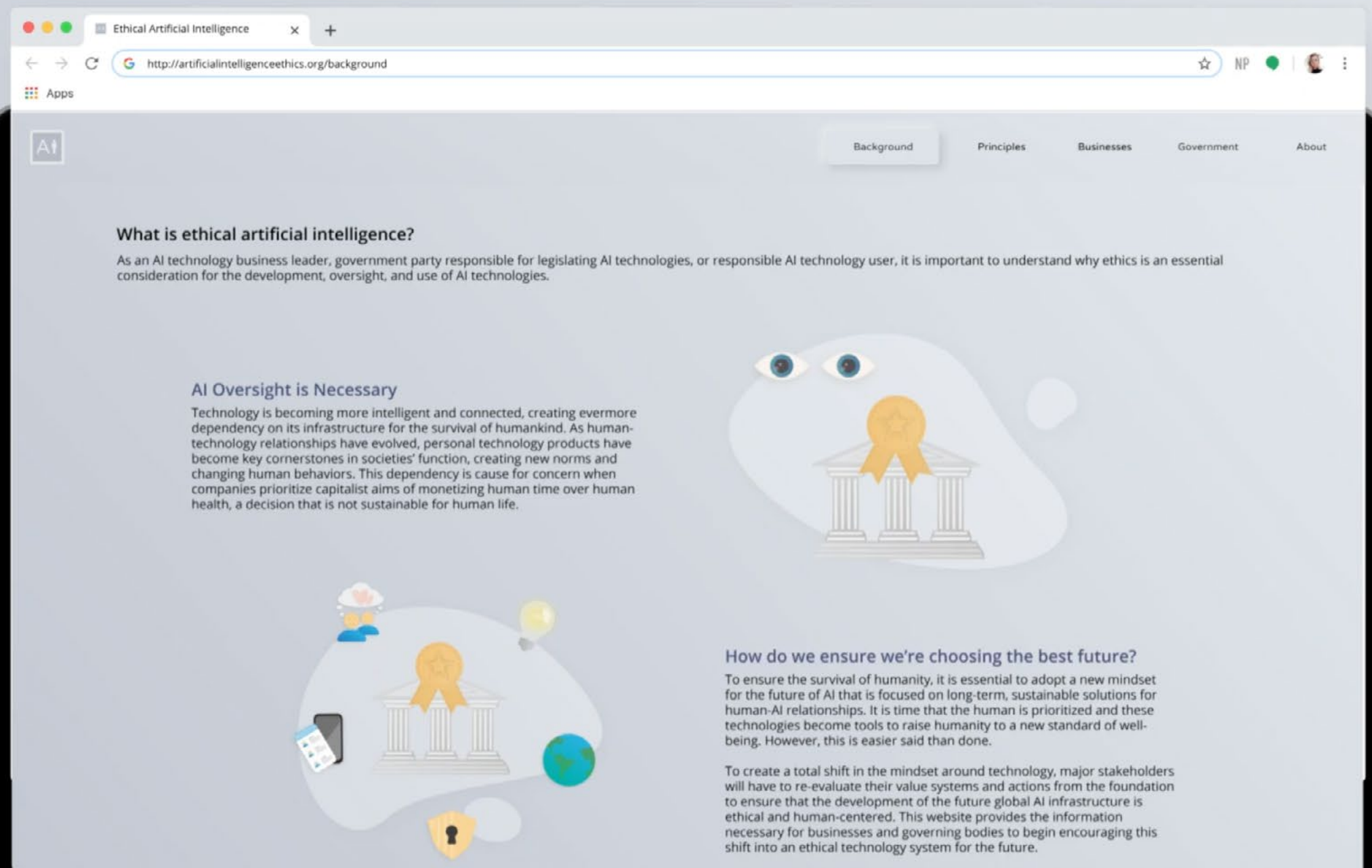


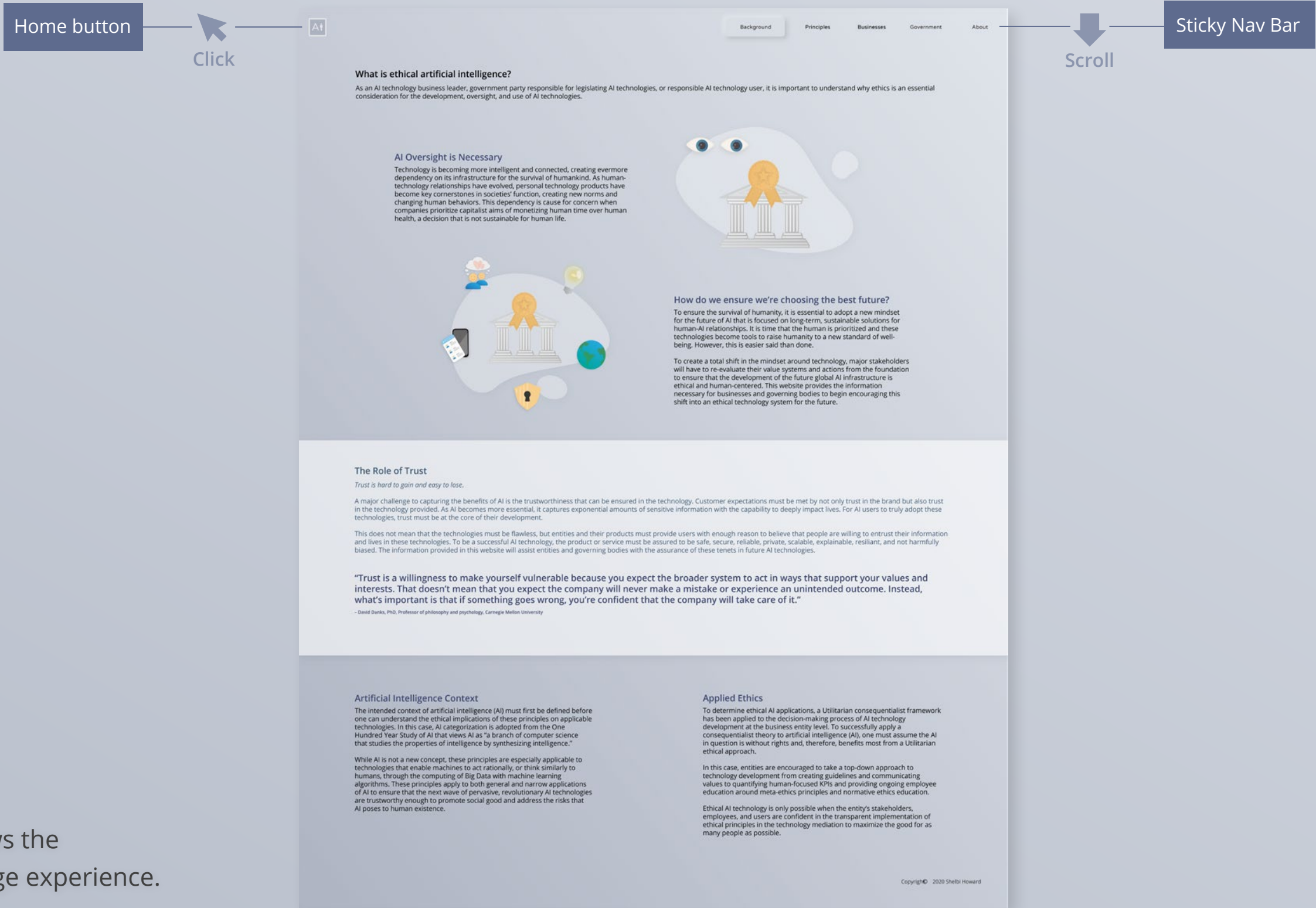




Wireframe (right) shows home page experience.

The importance of ethical AI is communicated in the **background**. From this page, businesses and governments choose from differing experiences.

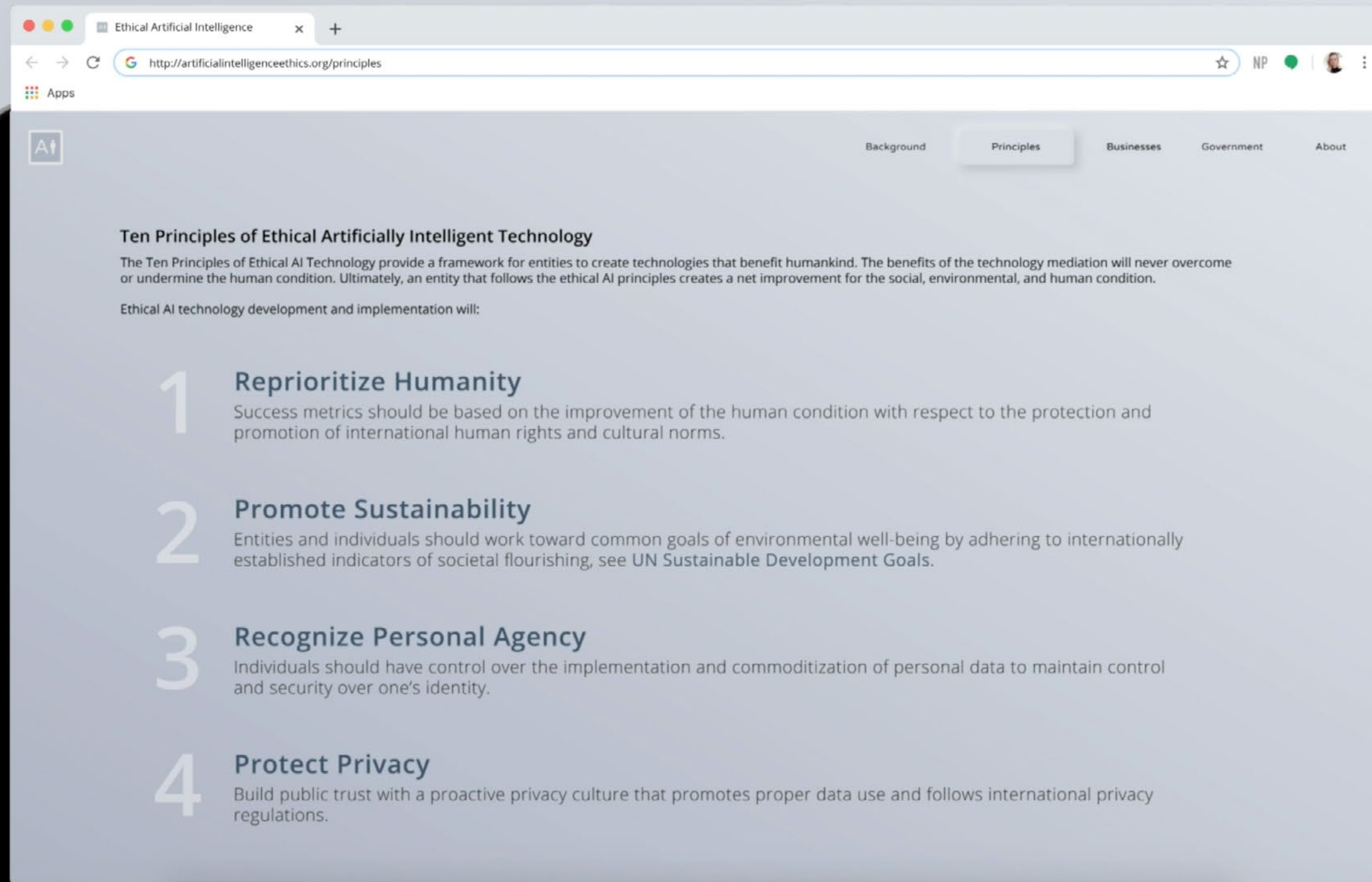




Wireframe shows the Background page experience.



The **Ten Principles of Ethical AI Development** set universal expectations for the future of ethical AI.



# Ten Principles of Ethical AI Development

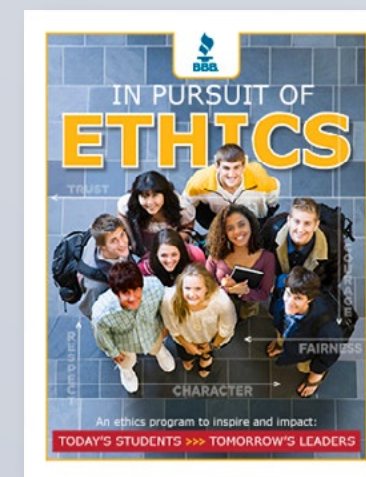
1	<b>Reprioritize Humanity</b> Success metrics should be based on the improvement of the human condition with respect to the protection and promotion of international human rights and cultural norms.
2	<b>Promote Sustainability</b> Entities and individuals should work toward common goals of environmental well-being by adhering to internationally established indicators of societal flourishing, see <a href="#">UN Sustainable Development Goals</a> .
3	<b>Recognize Personal Agency</b> Individuals should have control over the implementation and commoditization of personal data to maintain control and security over one's identity.
4	<b>Protect Privacy</b> Build public trust with a proactive privacy culture that promotes proper data use and follows international privacy regulations.
5	<b>Provide Security</b> Individuals should have reasonable trust in the proper management of personal information. If security is breached, the entity is responsible for informing affected parties and swiftly finding favorable solutions.
6	<b>Ensure Transparency</b> Decisions and motives should demonstrate good behavior through clear communication and accessibility to information. When problems arise, speed and response quality should strengthen user trust.
7	<b>Embrace Accountability</b> Entities responsible for technology should share the responsibility of the technology mediations, implementation, and effects with individuals involved in the technology's use.
8	<b>Equitable Opportunity</b> Technology solutions should be useful and marketable to people with diverse abilities throughout its lifespan and applications, see <a href="#">Universal Design Principles</a> . Entities are responsible for educating employees and building inclusive teams to ensure equitably mediated technology outcomes.
9	<b>Democratize Decisions</b> Technology decisions should be scrutinized by internal and external groups that represent products' diverse user base throughout the development and implementation processes. This includes: governing parties, users, business leaders, technologists, and 3rd party partners.
10	<b>Demonstrate Competence</b> Entities responsible for technology creation and implementation should operate with the required knowledge and skill to make effective, responsible decisions around the ethical implementation of the technology.



Tech Trends 2020  
Deloitte Insights



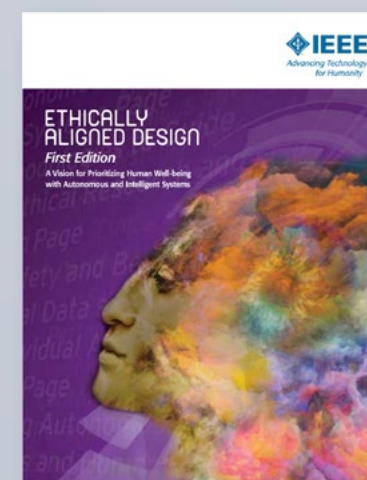
7 Universal Design Principles  
Center for Universal Design,  
North Carolina State University



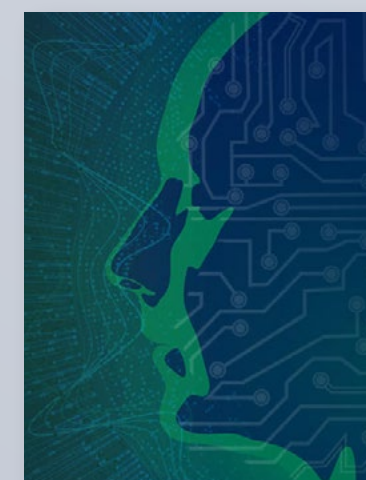
In Pursuit of Ethics  
BBB Training



International Health  
Regulations (2005)  
World Health Organization



Ethically Aligned Design  
Institute of Electrical and  
Electronics Engineers



Artificial Intelligence  
National Institute of  
Standards and Technology

Principles were created from aggregate **secondary research** of guides, principles, and methodologies on AI ethics, human-computer interaction, and trust.



Home button



Background Principles Businesses Government About



Sticky Nav Bar



Link to Business



Business Support  
Implement these principles  
into your business to create  
a more ethical future

Government Guidance  
Make informed decisions  
about AI technologies to  
create a more ethical future



Link to Government

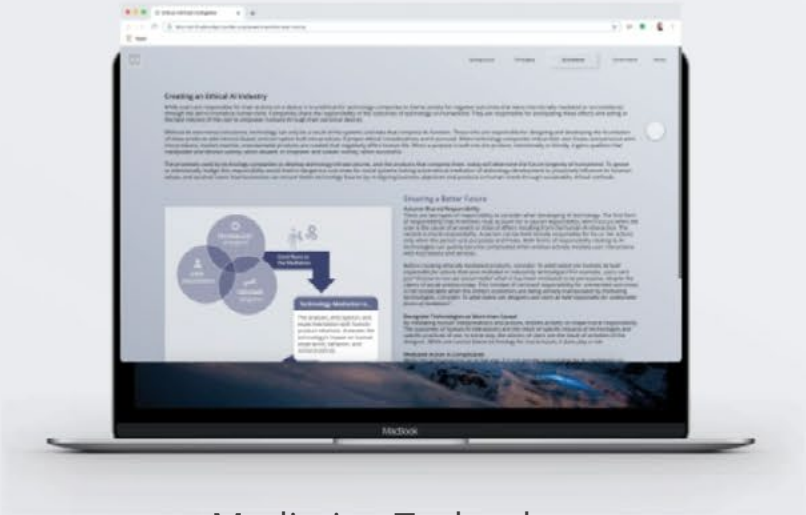


Business Support  
Implement these principles  
into your business to create  
a more ethical future

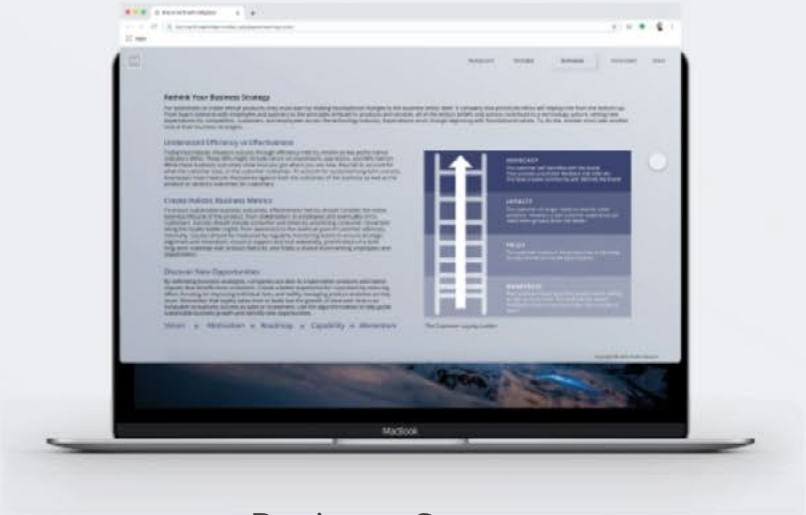
Government Guidance  
Make informed decisions  
about AI technologies to  
create a more ethical future

Wireframes show the  
Principles page experience.

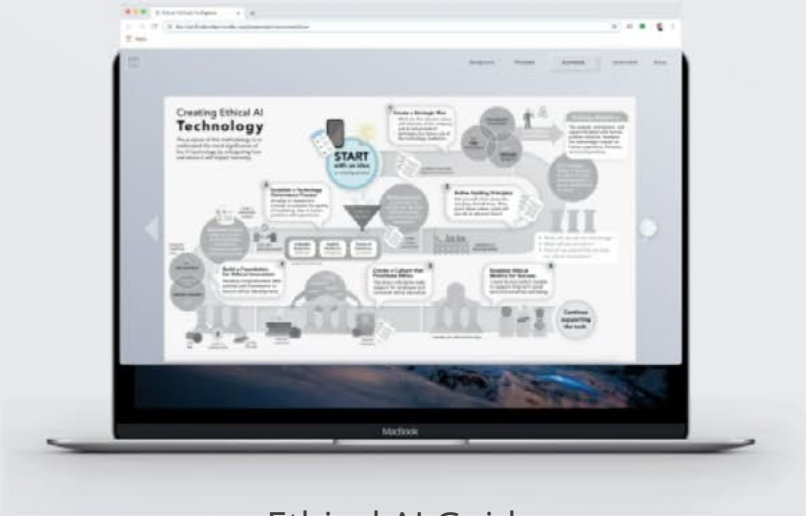
Businesses are guided through the development of ethical AI entities and technologies



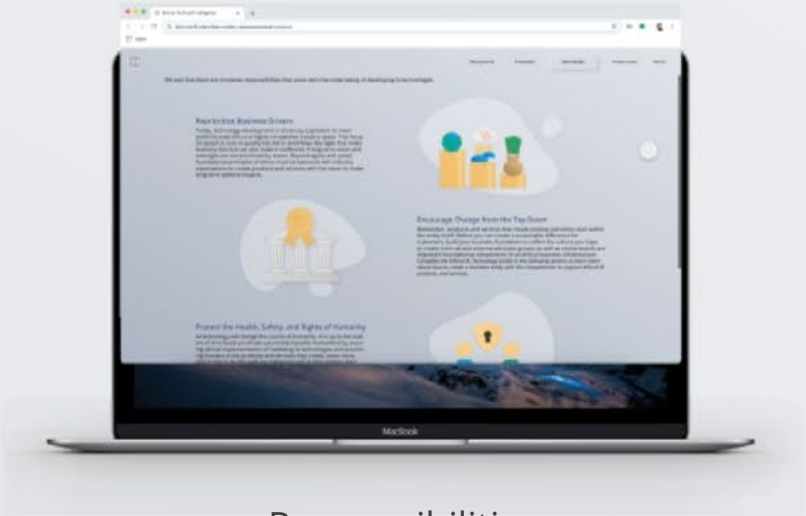
Mediating Technology



Business Strategy

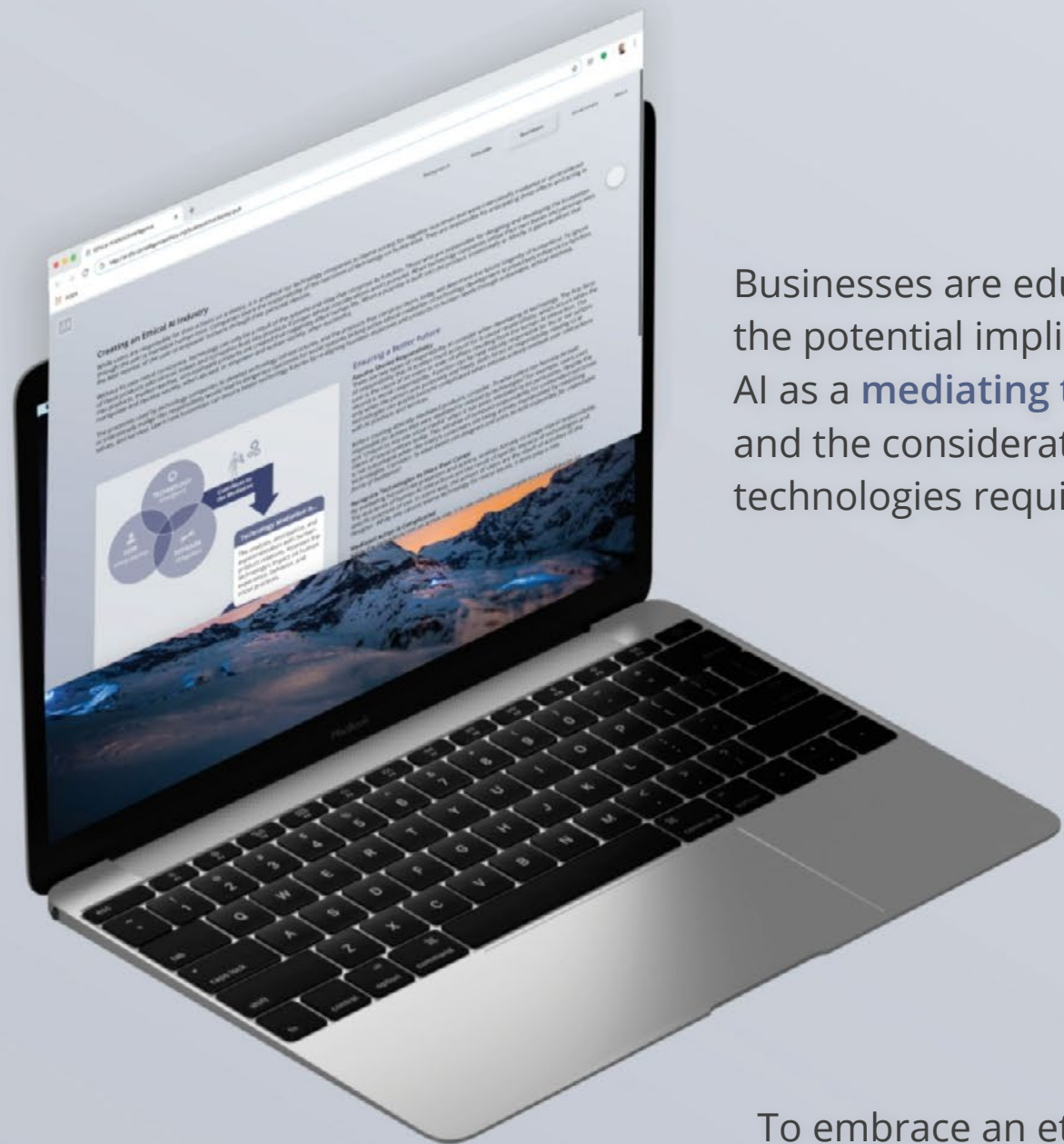


Ethical AI Guide



Responsibilities



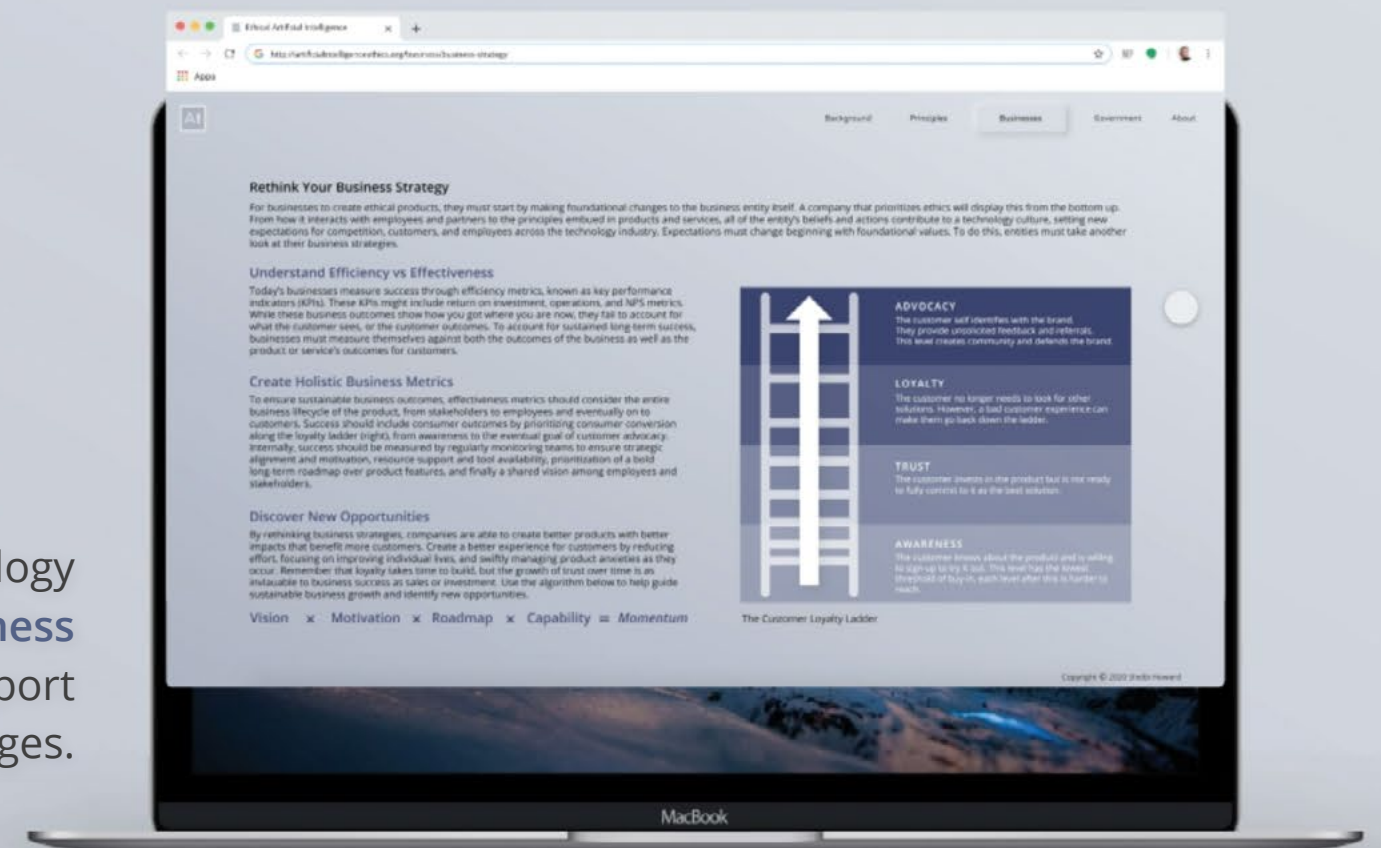


Businesses are educated on the potential implications of AI as a **mediating technology** and the considerations these technologies require.

To embrace an ethical AI technology within an existing entity, **business strategy** education provides support to make successful strategic changes.

“We’re seeing a kind of a Wild West situation with AI and regulation right now. The scale at which businesses are adopting AI technologies isn’t matched by clear guidelines to regulate algorithms and help researchers avoid the pitfalls of bias in datasets.”

— Timnit Gebru, Research Scientist, Google AI





Home button



Background

Principles

Businesses

Government

About



Scroll

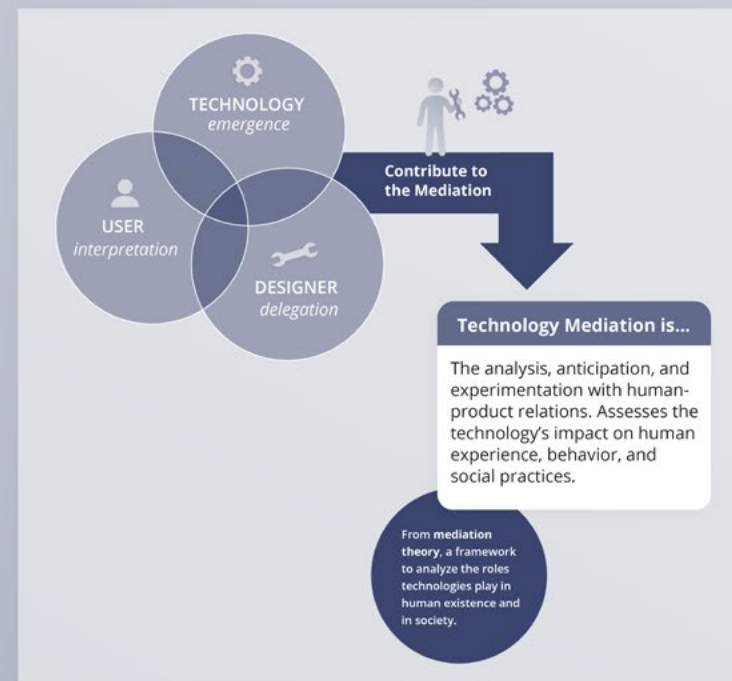
Sticky Nav Bar

## Creating an Ethical AI Industry

While users are responsible for their actions on a device, it is unethical for technology companies to blame society for negative outcomes that were intentionally mediated or unconsidered through the aim to monetize human time. Companies share the responsibility of the outcomes of technology on humankind. They are responsible for anticipating these effects and acting in the best interest of the user to empower humans through their personal devices.

Without its own moral conscience, technology can only be a result of the systems and data that comprise its function. Those who are responsible for designing and developing the foundation of these products add intrinsic biases and corruption built into products if proper ethical considerations aren't pursued. When technology companies imbue their own biases and personal aims into products, market-reactive, unsustainable products are created that negatively affect human life. When a purpose is built into the product, intentionally or blindly, it gains qualities that manipulate and deceive society, when abused, or empower and sustain society, when successful.

The processes used by technology companies to develop technology infrastructures, and the products that comprise them, today will determine the future longevity of humankind. To ignore or intentionally malign this responsibility would lead to dangerous outcomes for social systems lacking active ethical mediation of technology development to proactively influence its function, values, and survival. Learn how businesses can ensure better technology futures by re-aligning business objectives and products to human needs through sustainable, ethical methods.



## Ensuring a Better Future

### Assume Shared Responsibility

There are two types of responsibility to consider when developing AI technology. The first form of responsibility that AI entities must account for is causal responsibility, which occurs when the user is the cause of an event or state of affairs resulting from the human-AI interaction. The second is moral responsibility. A person can be held morally responsible for his or her actions only when the person acts purposely and freely. Both forms of responsibility relating to AI technologies can quickly become complicated when entities actively mediate user interactions with AI products and services.

Before creating ethically mediated products, consider: *To what extent can humans be held responsible for actions that were mediated or induced by technologies?* For example, users can't just "choose to not use social media" when it has been mediated to be persuasive, despite the claims of social entities today. This mindset of removed responsibility for unintended outcomes is not sustainable when the entity's customers are being actively manipulated by mediating technologies. Consider: *To what extent can designers and users be held responsible for undesirable forms of mediation?*

### Recognize Technologies as More than Causal

By mediating human interpretations and actions, entities actively co-shape moral responsibility. The outcomes of human-AI interactions are the result of specific impacts of technologies and specific practices of use. In some way, the actions of users are the result of activities of the designer. While one cannot blame technology for moral issues, it does play a role.

### Mediated Action is Complicated

While the technology has an active role, it is not morally accountable for its mediations on human behavior. With the rise of AI, technologies are becoming full-fledged moral agents in the way humans are moral agents. However, it is important to remember that both users and designers are responsible for technologically mediated actions.

### Realize Technology's Potential

Technology mediation should focus on the health and longevity of the humans using it. By applying ethics to the development of AI technology products and services, we can be proactive about our future.

Next: Business Strategy

Copyright © 2020 Shelbi Howard

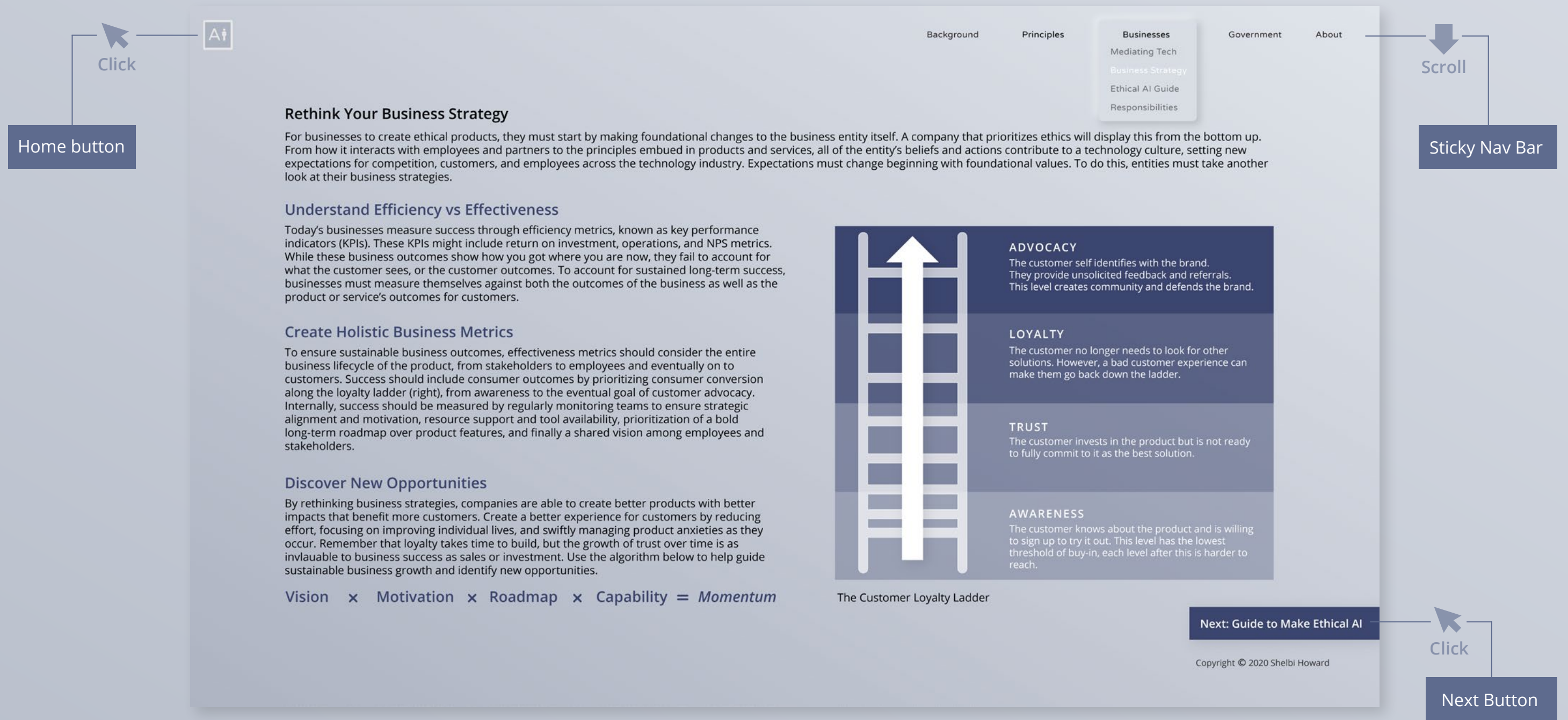


Click

Next Button

Appear at  
page bottom

Wireframes show the **Mediating Technology** experience for businesses.

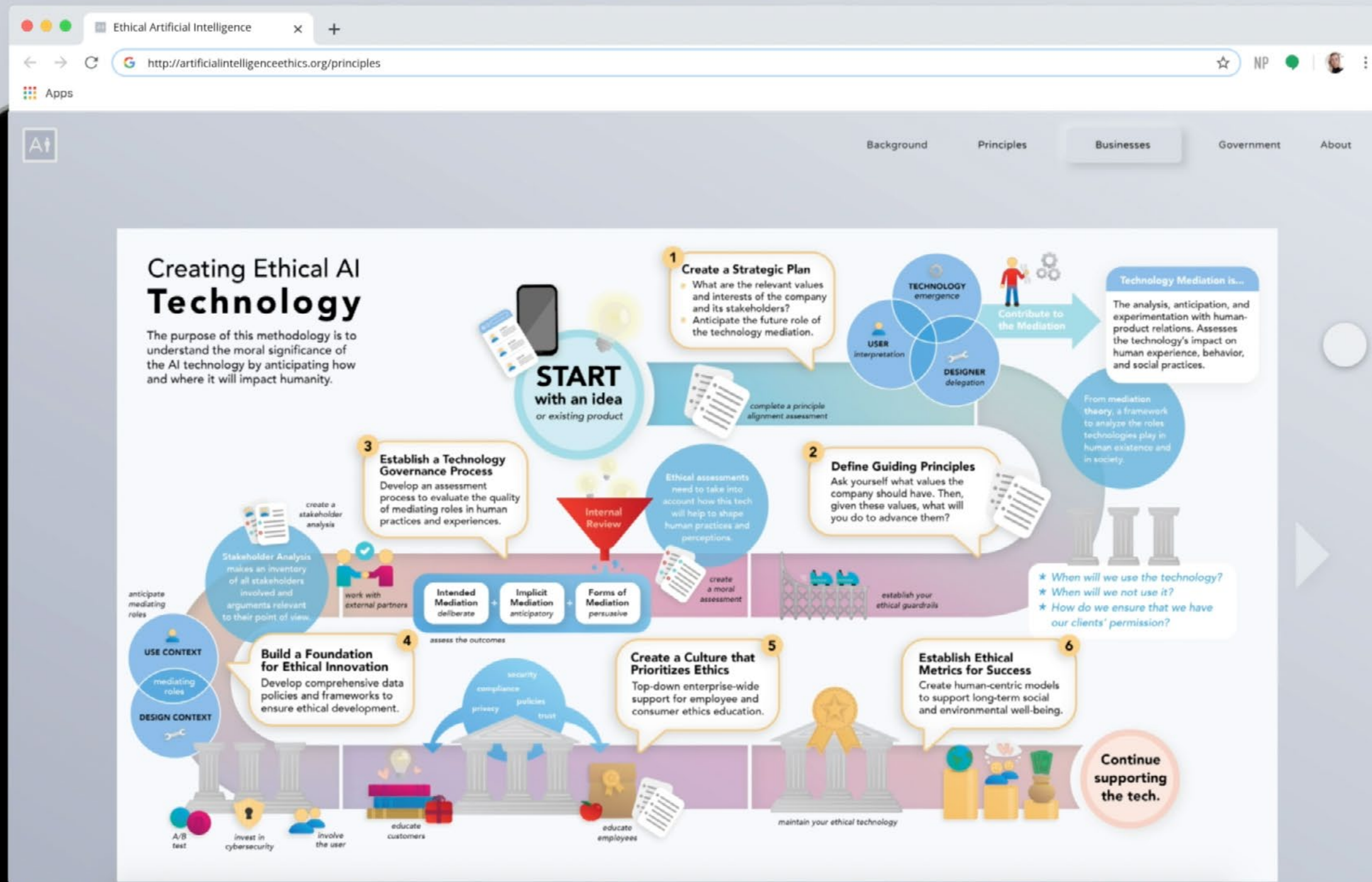


Wireframes show the **Business Strategy** experience for businesses.

Appear at page bottom



The **Ethical AI Development Guide** provides step-by-step guidance with tools to businesses.





Home button

Click

Background

Principles

Businesses

Landscape

Guide

Risks + Benefits

Government

About

Sticky Nav Bar

Scroll

Creating Ethical AI Technology

The purpose of this methodology is to understand the moral significance of the AI technology by anticipating how and where it will impact humanity.

1

Create a Strategic Plan

What are the relevant values and interests of the company and its stakeholders?  
Anticipate the future role of the technology mediation.

2

Define Guiding Principles

Ask yourself what values the company should have. Then, given these values, what will you do to advance them?

3

Establish a Technology Governance Process

Develop an assessment process to evaluate the quality of mediating roles in human practices and experiences.

4

Build a Foundation for Ethical Innovation

Develop comprehensive data policies and frameworks to ensure ethical development.

5

Create a Culture that Prioritizes Ethics

Top-down enterprise-wide support for employee and consumer ethics education.

6

Establish Ethical Metrics for Success

Create human-centric models to support long-term social and environmental well-being.

START with an idea or existing product

complete a principle alignment assessment

From mediation theory, a framework to analyze the roles technologies play in human existence and in society.

Technology Mediation is...  
The analysis, anticipation, and experimentation with human-product relations. Assesses the technology's impact on human experience, behavior, and social practices.

Stakeholder Analysis makes an inventory of all stakeholders involved and arguments relevant to their point of view.

create a stakeholder analysis

work with external partners

Internal Review

Ethical assessments need to take into account how this tech will help to shape human practices and perceptions.

create a moral assessment

establish your ethical guardrails

When will we use the technology?  
When will we not use it?  
How do we ensure that we have our clients' permission?

USE CONTEXT

mediating roles

DESIGN CONTEXT

assess the outcomes

security compliance privacy policies trust

invest in cybersecurity

involve the user

educate customers

educate employees

maintain your ethical technology

Continue supporting the tech.

Guide to Create an Ethical AI Business

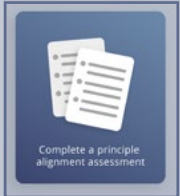
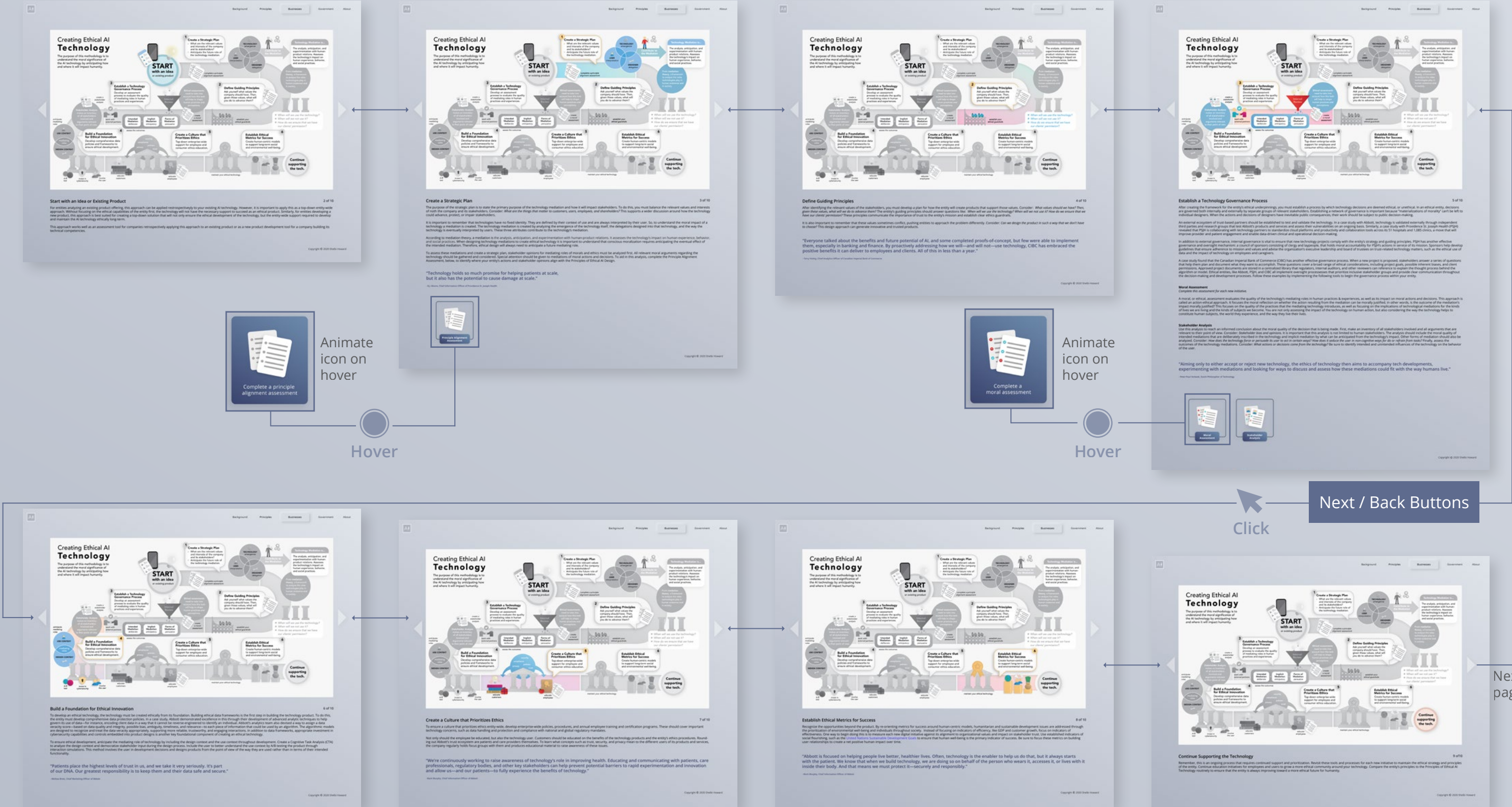
1 of 10

To create an ethical AI technology, the entity must have the necessary architecture in place to properly develop and support the technology. The following is a step-by-step approach for entities to ensure that both their business competencies and their technology are ethically aligned.

The aim of this approach is a two-fold focus on the entity and its technology. First, a framework for entities to ensure ethical development through their guiding principles, culture, and employee structure. Secondly, the tools provided with this approach assist entities in understanding the moral significance of the technology by anticipating how and where it will impact. Ultimately, entities are provided the resources to create a wholistic ethical approach with their AI technology.

Copyright © 2020 Shelbi Howard

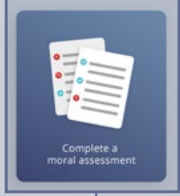
Wireframes show the **Ethical AI Development Guide** experience for businesses.



Animate icon on hover



Hover



Animate icon on hover



Hover



Click

Next / Back Buttons

Next page



Home button

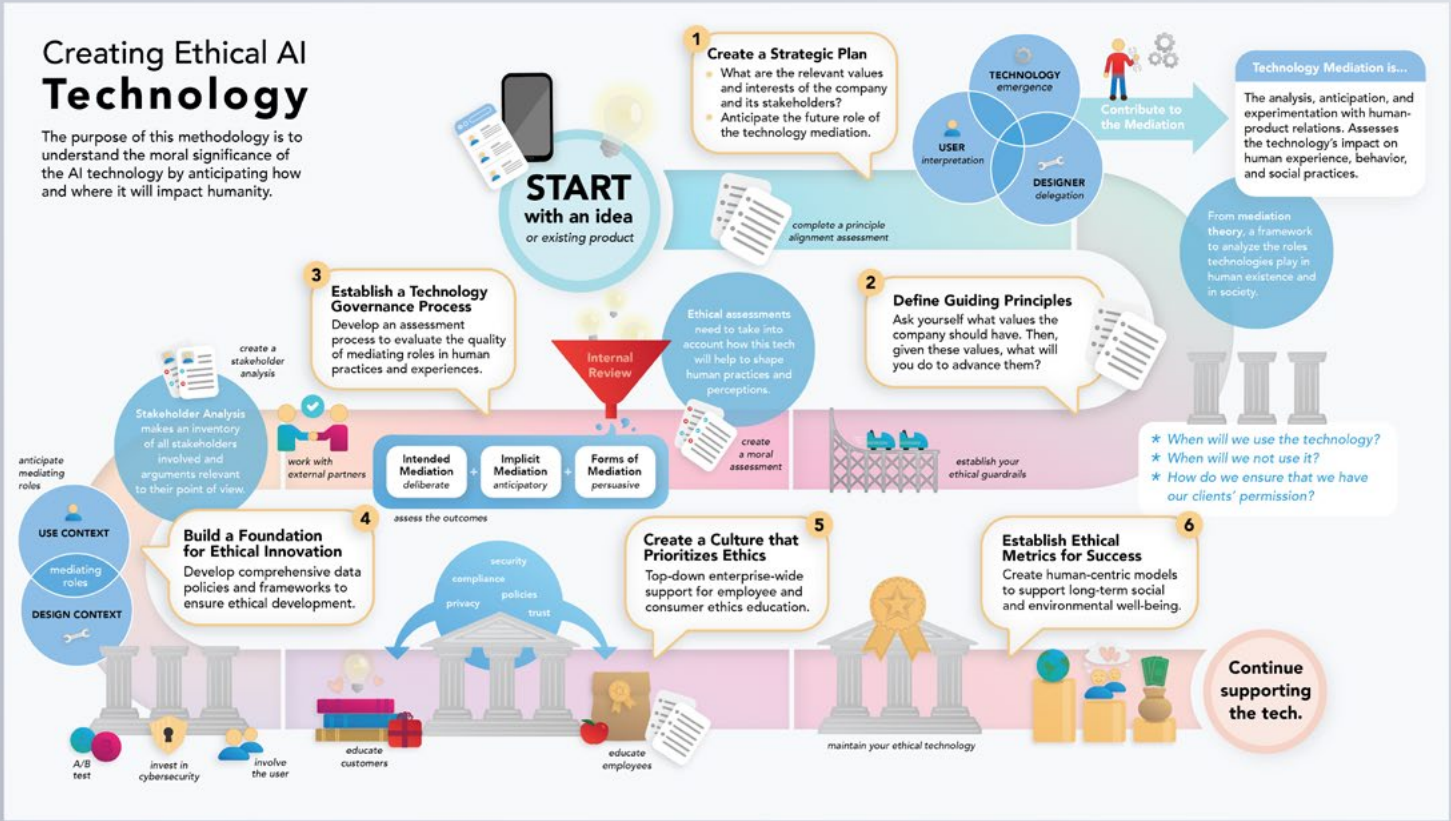


Background Principles Businesses Government About



Sticky Nav Bar

Back Button



Guide to Create an Ethical AI Business

Download and share this infographic to review later and support ethical technology.

10 of 10

Share



Share on Facebook

Share on Twitter

Download

Next: Responsibilities

Next Link



Appear at page bottom

Copyright © 2020 Shelbi Howard

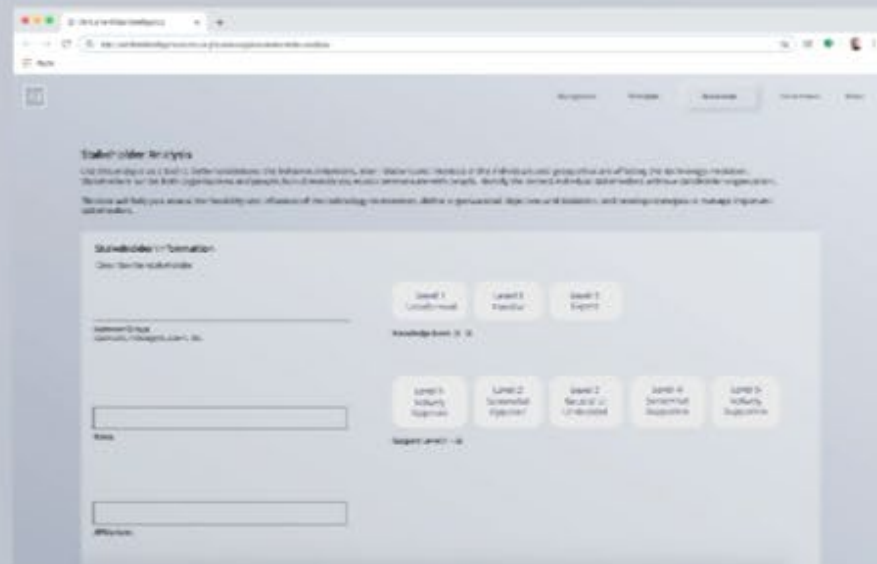


Download Button

Wireframes show the **Ethical AI Development Guide** experience for businesses.

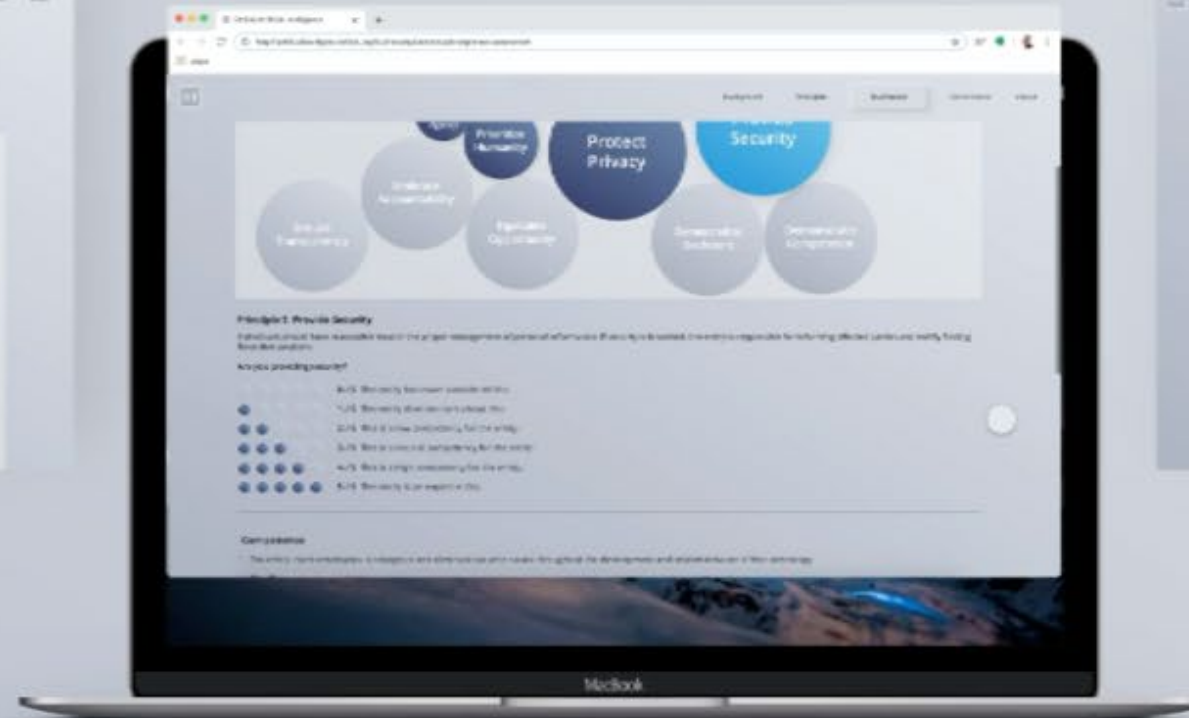


# Three tools guide businesses through the ethical development of AI technologies and entities.



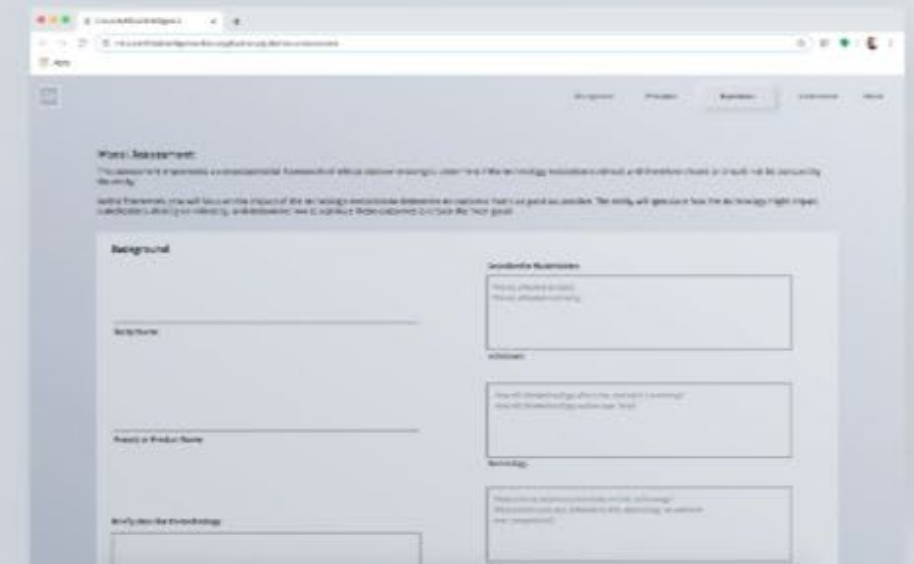
## Stakeholder Analysis

Analyze stakeholder opinions and impact based on personal bias, involvement, and power.



## Principle Alignment Assessment

Compare entity and product values to the Ten Principles of Ethical AI.



## Moral Assessment

Anticipate the ethical impact of the technology mediation.

The **Stakeholder Analysis** was created based on project management methods and the environmental screening stakeholder concept (Menlow, 1981).

The screenshot shows the 'Stakeholder Analysis' section of a web application. At the top, there are navigation tabs: Background, Principles, Stakeholders (active), and Governance. Below the title, two paragraphs explain the purpose of the analysis. The main form area contains sections for describing the stakeholder, their knowledge level, and their support level.

### Stakeholder Analysis

Use this analysis as a tool to better understand the behavior, intentions, inter-relations and interests of the individuals and groups that are affecting the technology mediation. Stakeholders can be both organizations and people, but ultimately you must communicate with people. Identify the correct individual stakeholders within a stakeholder organization. This tool will help you assess the feasibility and influence of the technology intervention, define organizational objectives and decisions, and develop strategies to manage important stakeholders.

---

#### Stakeholder Information

Describe the stakeholder

**Name:** Fiona Remyee Lewis

**Name or Group:** Sponsors, Managers, users, etc.

**Role:**

**Affiliations:**

---

#### Knowledge Level (1 - 3)

Level 1 Uninformed	Level 2 Familiar	Level 3 Expert

---

#### Support Level (1 - 5)

Level 1 Actively Opposed	Level 2 Somewhat Opposed	Level 3 Neutral or Uninvolved	Level 4 Somewhat Supportive	Level 5 Actively Supportive

Copyright © 2020 HealthForward

## Stakeholder Information (1/3)

[Background](#)
[Principles](#)
[Business](#)
[Government](#)
[About](#)

## Stakeholder Analysis

Use this analysis as a tool to better understand the behavior, intentions, inter-relations and interests of the individuals and groups that are affecting the technology mediation. Stakeholders can be both organizations and people, but ultimately you must communicate with people. Identify the correct individual stakeholders within a stakeholder organization.

This tool will help you assess the feasibility and influence of the technology intervention, define organizational objectives and decisions, and develop strategies to manage important stakeholders.

---

### Risks and Benefits

Analyze the potential risks and benefits this stakeholder creates

☐ Environment
 ☒ Bi Regulations

**Motivations / Drivers**  
Reason for stakeholder investment in project success

- Support and follow EPA standards

**Stakeholder Expectations**  
What the stakeholder expects from the project

☒ Cognitive

**Bases**  
Stakeholder inclination and prejudice [View Examples](#)

☒ Progressive ☒ Political

**Anticipated Issues**  
Identify known concerns based on stakeholder information

**Power**

Negative	Keep Satisfied	Manage Closely
Positive	Monitor (Minimum Effort)	Keep Informed

**Interest**

Interest	Power
Negative <input type="radio"/> Neutral <input type="radio"/> Positive <input checked="" type="radio"/>	Low <input type="radio"/> Moderate <input type="radio"/> High <input checked="" type="radio"/>

**Stakeholder Map**  
Measure the power and interest of the stakeholder

3 of 3

Back
Fetch

Copyright © 2020 Thea Howard

## Risks and Benefits (3/3)

Background

Principles

Business

Governance

About

## Stakeholder Analysis

Use this analysis as a tool to better understand the behavior, intentions, inter-relations and interests of the individuals and groups that are affecting the technology mediation. Stakeholders can be both organizations and people, but ultimately you must communicate with people. Identify the correct individual stakeholders within a stakeholder organization. This tool will help you assess the feasibility and influence of the technology intervention, define organizational objectives and decisions, and develop strategies to manage important stakeholders.

### Project Information

Describe the project the stakeholder is involved with

IRIS

Name

Project title

Sun

Mon

Tue

Wed

Thu

Fri

Sat

Daily

Weekly

Bi-weekly

Monthly

Implementation

Frequency of stakeholder communication

See Ops

Quality Assurance

Responsible Party

Who the stakeholder communicates with

QA result review

Release confirmation for each phase

Activities

Meeting or impacting the stakeholder

OK the environmental QA results

Confirm release date and time to EU and Asia Pacific

Project Expectations

Stakeholder's predicted input

Milestone One

Complete and launch phase 1 and finalize phase 2 from July. Begin phase 3 customer research.

8/1

Milestone Two

Launch phase 2 and begin customer research on phase 3.

8/1

Milestone Three

Begin phase 3 development.

10/1

+ Add Milestones

End of 2021 Q1

Date Due

Implementation completion date

Launched P1

Phase 1 launched today. Phase 2 sent to QA.

8/18

P3 Research Done

Phase 3 customer research ready for analysis and stories.

9/12

P3 Dev Ready

Phase 3 development ready to begin.

not

+ Add Project Status

2 of 3

Back

Next

Copyright © 2020 Intelis Research

## Project Information (2/3)

Next Page

Click

## Results

Click

A↑

BackgroundPrinciplesBusinessesGovernmentAbout

## Stakeholder Analysis - Results

Congratulations, you have completed the Stakeholder Analysis for **Fiona Renee Lewis**. Download this information and use it as a resource in project reviews and decision-making meetings. This will help you better understand the behavior, intentions, inter-relations and interests of the individuals and groups that are affecting the project.

Now, you can exit or create another assessment.

**FIONA RENEE LEWIS**

Iris Project Manager

PERSONAL INFORMATION

**Affiliations:** Environmental Protection Agency (EPA)

**Knowledge Level (1-3):** Expert (3)

**Support Level (1-5):** Somewhat Supportive (4)

STAKEHOLDER ANALYSIS

9/30/20

UPDATES

**Milestone Two: 9 / 1**  
Launch phase 2 and begin customer research on phase 3

**Project Status: 9/12**  
Phase 3 development ready to begin

PROJECT INFORMATION

Involvement: Wednesday, weekly

**Responsible Party:** Dev Ops, Quality Assurance

**Activities:**  
-QA result review  
-Release confirmation for each ophase

**Project Expectations:**  
-OK the environmental QA results  
-Confirm release date and time to EU and Asia Pacific

**Date due:** End of 2021 Q1

RISKS AND BENEFITS

**Motivations / Drivers:** Environment, EU regulations

**Biases:** Cognitive

**Anticipated Issues:** Progressive, Political

**Stakeholder Expectations:**  
-Support and follow EPA standards

**Stakeholder Map:** Positive interest, High power

Recommendation: Manage this stakeholder closely

Create New AnalysisDownloadExit

Copyright © 2020 Shelbi Howard

Results provide a carded biography of the stakeholder with management recommendations for the entity.

The **Principle Alignment Assessment** is a novel tool developed to compare current business values to the Ten Principles of Ethical AI.

Data visualization updates as user completes form.

AI

BackgroundPrinciplesBusinessesGovernmentAbout

Principle Alignment Assessment

This tool helps quantify the alignment of entities current technologies and practices to the [Ethical AI Technology Principles](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

Prioritize Humanity

Promote Sustainability

Recognize Personal Agency

Protect Privacy

Provide Security

Ensure Transparency

Embrace Accountability

Equitable Opportunity

Democratize Decisions

Demonstrate Competence

Principle 1: Prioritize Humanity

Success metrics should be based on the improvement of the human condition with respect to the protection and promotion of international human rights and cultural norms.

Are you prioritizing humanity?

0 / 5

The entity has never considered this.

1 / 5

The entity does not care about this.

2 / 5

This is a low competency for the entity.

3 / 5

This is a neutral competency for the entity.

4 / 5

This is a high competency for the entity.

5 / 5

The entity is an expert in this.

Competence

1. Success indicators and business goals address holistic long-term improvements. Ethical success indicators go beyond baseline efficiency indicators of gross domestic product, consumption, and safety to encompass efficacy indicators focused on the customer, environment, and society.

0 / 5

2. Unintended negative consequences are identified and accounted for in business goals and human-focused indicators.

0 / 5

Impact

1. The lifespan of the technology application is understood and considered.

0 / 5

2. The entity makes an effort to understand the experience of users and indirect users with the intention of increasing their quality of life.

0 / 5

3. Interactions with the technology intervention avoid harmful dampening or amplification of human emotional experiences.

0 / 5

Culture

1. The entity's approach to building trust prioritizes the interests, experiences, and professional standards of the culture.

0 / 5

2. The entity's approach to building trust prioritizes societal norms of the region.

0 / 5

Regulation

1. The entity is compliant with national laws and government controls where the technology exists.

0 / 5

Total: 0 / 40

1 of 10

Cancel

Next

Copyright © 2020 Shelby Howard

Principle 1 Incomplete (1/10)

AI

BackgroundPrinciplesBusinessesGovernmentAbout

Principle Alignment Assessment

This tool helps quantify the alignment of entities current technologies and practices to the [Ethical AI Technology Principles](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

Promote Sustainability

Recognize Personal Agency

Prioritize Humanity

Protect Privacy

Provide Security

Ensure Transparency

Embrace Accountability

Equitable Opportunity

Democratize Decisions

Demonstrate Competence

Principle 1: Prioritize Humanity

Success metrics should be based on the improvement of the human condition with respect to the protection and promotion of international human rights and cultural norms.

Are you prioritizing humanity?

0 / 5

The entity has never considered this.

1 / 5

The entity does not care about this.

2 / 5

This is a low competency for the entity.

3 / 5

This is a neutral competency for the entity.

4 / 5

This is a high competency for the entity.

5 / 5

The entity is an expert in this.

Competence

1. Success indicators and business goals address holistic long-term improvements. Ethical success indicators go beyond baseline efficiency indicators of gross domestic product, consumption, and safety to encompass efficacy indicators focused on the customer, environment, and society.

3 / 5

2. Unintended negative consequences are identified and accounted for in business goals and human-focused indicators.

2 / 5

Impact

1. The lifespan of the technology application is understood and considered.

1 / 5

2. The entity makes an effort to understand the experience of users and indirect users with the intention of increasing their quality of life.

3 / 5

3. Interactions with the technology intervention avoid harmful dampening or amplification of human emotional experiences.

2 / 5

Culture

1. The entity's approach to building trust prioritizes the interests, experiences, and professional standards of the culture.

0 / 5

2. The entity's approach to building trust prioritizes societal norms of the region.

1 / 5

Regulation

1. The entity is compliant with national laws and government controls where the technology exists.

4 / 5

Total: 16 / 40

1 of 10

Cancel

Next

Copyright © 2020 Shelby Howard

Principle 1 Complete (1/10)

Next page



Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 2: Promote Sustainability](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 2: Promote Sustainability**  
Entities and individuals should work toward common goals of environmental well-being by adhering to internationally established indicators of societal flourishing. See [UN Sustainable Development Goals](#).

Are you promoting sustainability?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. The entity provides employees and partners with the means to flourish in work and personal life. 0/5

**Impact**

1. The technology intervention has a positive impact on the lives of the individuals that are directly and indirectly interacting with the technology intervention. 0/5

2. Long-term goals are in place for the eventual impact of the technology. 0/5

**Action**

1. The entity proactively minimizes the effect of the technology on the physical environment. 0/5

2. The entity proactively minimizes the effect of the technology on the social environment. 0/5

3. The entity works internally and with third parties to proactively improve its impact on the physical and social environments. 0/5

Total: 0 / 30 2 of 10

Back Next

Copyright © 2020 IBM Research

Principle 2 incomplete (2/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 3: Recognize Personal Agency](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 3: Recognize Personal Agency**  
Individuals should have control over the implementation and communication of personal data to maintain control and security over their identity.

Are you recognizing personal agency?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. Individuals are capable of assessing, sharing, and benefiting from their data. 0/5

**Accessibility**

1. Individuals understand the scope and impact of the data collected by the entity. 0/5

2. The entity has created policies and practices to make individual aware of the consequences of aggregating and meeting personal information. 0/5

3. Individuals have direct control over the data, specific, and false exchange of personal information at all times. 0/5

**Culture**

1. Individuals are empowered by their personal data policies and insights. 0/5

2. Individuals are in control of their digital identities, with respect to cultural norms. 0/5

Total: 0 / 30 2 of 10

Back Next

Copyright © 2020 IBM Research

Principle 3 incomplete (3/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 4: Protect Privacy](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 4: Protect Privacy**  
Individuals should have reasonable trust in the proper management of personal information. If security is breached, the entity is responsible for informing affected parties and safely fixing breaches.

Are you protecting privacy?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. Decisions regarding personal privacy information are readily available and accessible to individuals. 0/5

2. The entity is proactive about the protection of privacy issues in comparison to industry leaders. 0/5

**Regulation**

1. The entity implements third-party digital, physical, and regulatory protection settings. 0/5

**Culture**

1. The entity responds to security fears, safety, and efficiency when they are identified. 0/5

2. Individuals have reasonable trust in the security methods that the entity provides. 0/5

Total: 0 / 30 4 of 10

Back Next

Copyright © 2020 IBM Research

Principle 4 incomplete (4/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 5: Provide Security](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 5: Provide Security**  
Individuals should have reasonable trust in the proper management of personal information. If security is breached, the entity is responsible for informing affected parties and safely fixing breaches.

Are you providing security?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. The entity trains employees to recognize and eliminate security issues throughout the development and implementation of the technology. 0/5

2. The entity is proactive about the protection of security issues in comparison to industry leaders. 0/5

3. The entity maintains confidentiality with well-known, increasingly verifiable authorities. 0/5

4. The entity is willing to sacrifice short-term revenue to protect long-term trust or ongoing projects. 0/5

5. The entity implements third-party digital, physical, and regulatory protection settings. 0/5

**Culture**

1. The entity responds to security fears, safety, and efficiency when they are identified. 0/5

2. Individuals have reasonable trust in the security methods that the entity provides. 0/5

Total: 0 / 30 5 of 10

Back Next

Copyright © 2020 IBM Research

Principle 5 incomplete (5/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 6: Ensure Transparency](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 6: Ensure Transparency**  
Entities and individuals should demonstrate good faith through clear communication and accessibility to information. When problems arise, speed and response quality should strengthen user trust.

Are you ensuring transparency?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. Technical business decisions involving data, transactions, and security policies are available for individual review. 0/5

2. Individuals are provided evidence of effectiveness and efficiency metrics resulting from the technology. 0/5

**Accessibility**

1. The entity provides resources for all decisions made involving the technology and its customers. 0/5

2. Individuals are notified when issues arise involving personal information. 0/5

3. Issues are dealt with quickly and efficiently when they do arise. 0/5

Total: 0 / 30 6 of 10

Back Next

Copyright © 2020 IBM Research

Principle 6 incomplete (6/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 2: Promote Sustainability](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 2: Promote Sustainability**  
Entities and individuals should work toward common goals of environmental well-being by adhering to internationally established indicators of societal flourishing. See [UN Sustainable Development Goals](#).

Are you promoting sustainability?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. The entity provides employees and partners with the means to flourish in work and personal life. 0/5

**Impact**

1. The technology intervention has a positive impact on the lives of the individuals that are directly and indirectly interacting with the technology intervention. 0/5

2. Long-term goals are in place for the eventual impact of the technology. 0/5

**Action**

1. The entity proactively minimizes the effect of the technology on the physical environment. 0/5

2. The entity proactively minimizes the effect of the technology on the social environment. 0/5

3. The entity works internally and with third parties to proactively improve its impact on the physical and social environments. 0/5

Total: 9 / 30 2 of 10

Back Next

Copyright © 2020 IBM Research

Principle 2 complete (2/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 3: Recognize Personal Agency](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 3: Recognize Personal Agency**  
Individuals should have control over the implementation and communication of personal data to maintain control and security over their identity.

Are you recognizing personal agency?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. Individuals are capable of assessing, sharing, and benefiting from their data. 0/5

**Accessibility**

1. Individuals understand the scope and impact of the data collected by the entity. 0/5

2. The entity has created policies and practices to make individual aware of the consequences of aggregating and meeting personal information. 0/5

3. Individuals have direct control over the data, specific, and false exchange of personal information at all times. 0/5

**Culture**

1. Individuals are empowered by their personal data policies and insights. 0/5

2. Individuals are in control of their digital identities, with respect to cultural norms. 0/5

Total: 7 / 30 3 of 10

Back Next

Copyright © 2020 IBM Research

Principle 3 complete (3/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 4: Protect Privacy](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 4: Protect Privacy**  
Individuals should have reasonable trust in the proper management of personal information. If security is breached, the entity is responsible for informing affected parties and safely fixing breaches.

Are you protecting privacy?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. Decisions regarding personal privacy information are readily available and accessible to individuals. 0/5

2. The entity is proactive about the protection of privacy issues in comparison to industry leaders. 0/5

**Regulation**

1. The entity implements third-party digital, physical, and regulatory protection settings. 0/5

**Culture**

1. The entity responds to security fears, safety, and efficiency when they are identified. 0/5

2. Individuals have reasonable trust in the security methods that the entity provides. 0/5

Total: 18 / 30 4 of 10

Back Next

Copyright © 2020 IBM Research

Principle 4 complete (4/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 5: Provide Security](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 5: Provide Security**  
Individuals should have reasonable trust in the proper management of personal information. If security is breached, the entity is responsible for informing affected parties and safely fixing breaches.

Are you providing security?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. The entity trains employees to recognize and eliminate security issues throughout the development and implementation of the technology. 0/5

2. The entity is proactive about the protection of security issues in comparison to industry leaders. 0/5

3. The entity maintains confidentiality with well-known, increasingly verifiable authorities. 0/5

4. The entity is willing to sacrifice short-term revenue to protect long-term trust or ongoing projects. 0/5

5. The entity implements third-party digital, physical, and regulatory protection settings. 0/5

**Culture**

1. The entity responds to security fears, safety, and efficiency when they are identified. 0/5

2. Individuals have reasonable trust in the security methods that the entity provides. 0/5

Total: 22 / 30 5 of 10

Back Next

Copyright © 2020 IBM Research

Principle 5 complete (5/10)

Principle Alignment Assessment

The user has selected the alignment of various current technologies and practices to the [Principle 6: Ensure Transparency](#). After assessing the technology intervention using these statements, entities are better positioned to identify and address strengths and weaknesses in a technology ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 6: Ensure Transparency**  
Entities and individuals should demonstrate good faith through clear communication and accessibility to information. When problems arise, speed and response quality should strengthen user trust.

Are you ensuring transparency?

1. The entity has never considered this. 0/5  
2. The entity does not care about this. 0/5  
3. This is a low competency for the entity. 0/5  
4. This is a neutral competency for the entity. 0/5  
5. This is a high competency for the entity. 0/5  
6. The entity is an expert in this. 0/5

**Competence**

1. The entity provides resources for all decisions made involving the technology and its customers. 0/5

2. Individuals are notified when issues arise involving personal information. 0/5

3. Issues are dealt with quickly and efficiently when they do arise. 0/5

Total: 12 / 30 6 of 10

Back Next

Copyright © 2020 IBM Research

Principle 6 complete (6/10)

Next  
Page

Click

Next  
page

Principle Alignment Assessment

This tool helps quantify the alignment of various current technologies and practices to the [Global Technology Principles](#). After assessing the technology intervention using these dimensions, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 7: Ensure Transparency**

Entities and systems should demonstrate open and honest communication and accessibility to information. When problems arise, speed and expertise qualify product design and development.

Are you ensuring transparency?

1. The entity has never considered this. 0/10  
2. The entity does not care about this. 0/10  
3. This is a low competency for the entity. 0/10  
4. This is a neutral competency for the entity. 0/10  
5. This is a high competency for the entity. 0/10  
6. The entity is an expert in this. 0/10

**Accessibility**

1. Technical business decisions involving data, transactions, and security (limited and available for individual access). 0/10  
2. Individuals are provided evidence of effectiveness and efficiency metrics resulting from the technology. 0/10

**Competence**

1. The entity provides resources for all decisions made involving the technology and its customers. 0/10  
2. Individuals are notified about issues when involving personal information. 0/10  
3. Issues are dealt with quickly and efficiently when they do arise. 0/10

Total: 0 / 25 0 of 10 Back Next

Principle 7 incomplete (7/10)

Principle Alignment Assessment

This tool helps quantify the alignment of various current technologies and practices to the [Global Technology Principles](#). After assessing the technology intervention using these dimensions, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 8: Equitable Opportunity**

Technology systems should be useful and accessible to people with diverse abilities throughout its lifespan and applications, see [Universal Design Principles](#). Entities are responsible for ensuring equitable and enabling inclusive teams to create equitably inclusive technology systems.

Are you providing equitable opportunity?

1. The entity has never considered this. 0/10  
2. The entity does not care about this. 0/10  
3. This is a low competency for the entity. 0/10  
4. This is a neutral competency for the entity. 0/10  
5. This is a high competency for the entity. 0/10  
6. The entity is an expert in this. 0/10

**Accessibility**

1. Relevant user limitations have been considered throughout development and distribution of the technology. 0/10

**Competence**

1. The technology intervention is simple and intuitive to use relevant to the application. 0/10  
2. The design accommodates for a range of individual abilities and user preferences. 0/10  
3. Important information is perceivable for individuals in varying conditions and with differing sensory abilities. 0/10  
4. The design requires low physical input that can be sustained throughout the individual's experience and the technology's lifespan. 0/10

Total: 0 / 25 0 of 10 Back Next

Principle 8 incomplete (8/10)

Principle Alignment Assessment

This tool helps quantify the alignment of various current technologies and practices to the [Global Technology Principles](#). After assessing the technology intervention using these dimensions, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 9: Democratic Decisions**

Technology decisions should be consultative, transparent and external groups that represent product's diverse user base throughout the development and implementation process. This includes governing bodies, users, business leaders, technology, and third-party partners.

Are you making inclusive decisions?

1. The entity has never considered this. 0/10  
2. The entity does not care about this. 0/10  
3. This is a low competency for the entity. 0/10  
4. This is a neutral competency for the entity. 0/10  
5. This is a high competency for the entity. 0/10  
6. The entity is an expert in this. 0/10

**Accessibility**

1. The entity allocates time for high-level discussions about technology requirements that include the customer, third parties affected by the technology, subject matter experts, and human advisors. 0/10  
2. The technology intervention's stakeholders, teams, and roles accurately represent the diverse population of its customers. 0/10  
3. Multiple responses are represented, including different industries, economic backgrounds, educational experiences, genders, and ethnic backgrounds. 0/10

**Competence**

1. The entity hires employees to recognize and address personal biases throughout the development and implementation of the technology. 0/10  
2. The entity provides necessary resources to help employees recognize ethical dilemmas, evaluate alternatives, and make and test ethical technology decisions that maintain public trust. 0/10  
3. Internal decisions are justified, recorded, and accessible for clear communication throughout the development process. 0/10

**Culture**

1. The entity is open to new skills and welcomes through active hiring and recruitment of relevant potential employees. 0/10

Total: 0 / 25 0 of 10 Cancel Next

Principle 9 incomplete (9/10)

Principle Alignment Assessment

This tool helps quantify the alignment of various current technologies and practices to the [Global Technology Principles](#). After assessing the technology intervention using these dimensions, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 10: Demonstrate Competence**

Entities responsible for technology creation and implementation should operate with the required knowledge and skill to make effective, responsible decisions around the ethical implementation of the technology.

Are you demonstrating competence?

1. The entity has never considered this. 0/10  
2. The entity does not care about this. 0/10  
3. This is a low competency for the entity. 0/10  
4. This is a neutral competency for the entity. 0/10  
5. This is a high competency for the entity. 0/10  
6. The entity is an expert in this. 0/10

**Competence**

1. The entity is a subject matter expert in the technology that is being implemented. 0/10  
2. The entity is aware of the social, environmental, and economic risks associated with the design, implementation, and future of the technology. 0/10  
3. The entity has clear standards for medication failure and is prepared to respond in case of failure. 0/10  
4. The entity adheres to safe and effective operation standards for the technology. 0/10  
5. The entity has specified the knowledge and skills that employees are required to learn and adhere to for the safe and ethical operation of the technology intervention. 0/10  
6. Sufficient resources are provided for employees to continue education and industry relevant to maintain future development and management of the technology. 0/10

Total: 0 / 25 0 of 10 Back Next

Principle 10 incomplete (10/10)

Principle Alignment Assessment

This tool helps quantify the alignment of various current technologies and practices to the [Global Technology Principles](#). After assessing the technology intervention using these dimensions, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 7: Ensure Transparency**

Entities and systems should demonstrate open and honest communication and accessibility to information. When problems arise, speed and expertise qualify product design and development.

Are you ensuring transparency?

1. The entity has never considered this. 0/10  
2. The entity does not care about this. 0/10  
3. This is a low competency for the entity. 0/10  
4. This is a neutral competency for the entity. 0/10  
5. This is a high competency for the entity. 0/10  
6. The entity is an expert in this. 0/10

**Accessibility**

1. Technical business decisions involving data, transactions, and security (limited and available for individual access). 0/10  
2. Individuals are provided evidence of effectiveness and efficiency metrics resulting from the technology. 0/10

**Competence**

1. The technology entity assumes responsibility when issues arise. 0/10  
2. Individuals are notified about issues when involving personal information. 0/10  
3. Issues are dealt with quickly and efficiently when they do arise. 0/10

**Culture**

1. The entity is open and honest about the medications implemented by the technology and issues with the intervention, and its potential effects on individuals. 0/10  
2. Individuals are supported by the entity when faced with issues. 0/10

Total: 7 / 25 7 of 10 Back Next

Principle 7 complete (7/10)

Next  
Page

Click

Principle Alignment Assessment

This tool helps quantify the alignment of various current technologies and practices to the [Global Technology Principles](#). After assessing the technology intervention using these dimensions, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 8: Equitable Opportunity**

Technology systems should be useful and accessible to people with diverse abilities throughout its lifespan and applications, see [Universal Design Principles](#). Entities are responsible for ensuring equitable and enabling inclusive teams to create equitably inclusive technology systems.

Are you providing equitable opportunity?

1. The entity has never considered this. 0/10  
2. The entity does not care about this. 0/10  
3. This is a low competency for the entity. 0/10  
4. This is a neutral competency for the entity. 0/10  
5. This is a high competency for the entity. 0/10  
6. The entity is an expert in this. 0/10

**Accessibility**

1. Relevant user limitations have been considered throughout development and distribution of the technology. 0/10

**Competence**

1. The technology intervention is simple and intuitive to use relevant to the application. 0/10  
2. The design accommodates for a range of individual abilities and user preferences. 0/10  
3. Important information is perceivable for individuals in varying conditions and with differing sensory abilities. 0/10  
4. The design requires low physical input that can be sustained throughout the individual's experience and the technology's lifespan. 0/10

Total: 15 / 25 8 of 10 Back Next

Principle 8 complete (8/10)

Principle Alignment Assessment

This tool helps quantify the alignment of various current technologies and practices to the [Global Technology Principles](#). After assessing the technology intervention using these dimensions, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 9: Democratic Decisions**

Technology decisions should be consultative, transparent and external groups that represent product's diverse user base throughout the development and implementation process. This includes governing bodies, users, business leaders, technology, and third-party partners.

Are you making inclusive decisions?

1. The entity has never considered this. 0/10  
2. The entity does not care about this. 0/10  
3. This is a low competency for the entity. 0/10  
4. This is a neutral competency for the entity. 0/10  
5. This is a high competency for the entity. 0/10  
6. The entity is an expert in this. 0/10

**Accessibility**

1. The entity allocates time for high-level discussions about technology requirements that include the customer, third parties affected by the technology, subject matter experts, and human advisors. 0/10  
2. The technology intervention's stakeholders, teams, and roles accurately represent the diverse population of its customers. 0/10  
3. Multiple responses are represented, including different industries, economic backgrounds, educational experiences, genders, and ethnic backgrounds. 0/10

**Competence**

1. The entity hires employees to recognize and address personal biases throughout the development and implementation of the technology. 0/10  
2. The entity provides necessary resources to help employees recognize ethical dilemmas, evaluate alternatives, and make and test ethical technology decisions that maintain public trust. 0/10  
3. Internal decisions are justified, recorded, and accessible for clear communication throughout the development process. 0/10

**Culture**

1. The entity is open to new skills and welcomes through active hiring and recruitment of relevant potential employees. 0/10

Total: 7 / 25 9 of 10 Cancel Next

Principle 9 complete (9/10)

Principle Alignment Assessment

This tool helps quantify the alignment of various current technologies and practices to the [Global Technology Principles](#). After assessing the technology intervention using these dimensions, entities are better positioned to identify and address strengths and weaknesses in a technology's ethical competencies.

This assessment is best completed with a diverse team of stakeholders that represent the entity's varying interests and priorities.

**Principle 10: Demonstrate Competence**

Entities responsible for technology creation and implementation should operate with the required knowledge and skill to make effective, responsible decisions around the ethical implementation of the technology.

Are you demonstrating competence?

1. The entity has never considered this. 0/10  
2. The entity does not care about this. 0/10  
3. This is a low competency for the entity. 0/10  
4. This is a neutral competency for the entity. 0/10  
5. This is a high competency for the entity. 0/10  
6. The entity is an expert in this. 0/10

**Competence**

1. The entity is a subject matter expert in the technology that is being implemented. 0/10  
2. The entity is aware of the social, environmental, and economic risks associated with the design, implementation, and future of the technology. 0/10  
3. The entity has clear standards for medication failure and is prepared to respond in case of failure. 0/10  
4. The entity adheres to safe and effective operation standards for the technology. 0/10  
5. The entity has specified the knowledge and skills that employees are required to learn and adhere to for the safe and ethical operation of the technology intervention. 0/10  
6. Sufficient resources are provided for employees to continue education and industry relevant to maintain future development and management of the technology. 0/10

Total: 19 / 25 10 of 10 Back Next

Principle 10 complete (10/10)

Results

Click



Results show strengths and weaknesses in businesses' ethical AI values and instruct them on ways to improve.



Results

Click

Share

Click

Download Button

Click

Click

Exit Button



Results (expanded)



10 Principles

Meta Principle: The benefits of the technology intervention never overcome the benefits for humanity. Ultimately, the technology intervention creates overall improvement in humanity and society.

Ethical AI technology development and implementation will:

1. Reprioritize humanity

Success metrics should be based on the improvement of the human condition with respect to the protection and promotion of international human rights and cultural norms.

Are you prioritizing humanity?

1-5 ranking based on number of "yes" answers.

1. Competence: Does the entity identify clear success indicators for the technology that go beyond gross domestic product, consumption, and safety?

2. Competence: Do success indicators and business goals address holistic long-term improvements to the customer, environment, and society?

3. Competence: Are unintended negative consequences identified and accounted for in business goals and human-focused indicators?

4. Impact: Is the entire lifespan of the technology understood and considered?

5. Impact: Is the entity aware of all of the individuals affected by the technology, including those who are not direct customers?

6. Impact: Does the technology increase the quality of life for all individuals affected by it?

7. Impact: Do interactions with the technology avoid harmful dampening or amplification of human emotional experiences?

8. Culture: Does the entity's approach to building trust prioritize the interests, experiences, and professional standards of the region?

9. Culture: Does the entity's approach to building trust prioritize societal norms and government controls of the region?

10. Regulation: Does the entity respect the international laws where the technology exists?


Rating: \_\_\_\_\_

2. Promote sustainability

Entities and individuals should work toward common goals of environmental well-being by adhering to internationally established indicators of societal flourishing, see [UN Sustainable Development Goals](#).

Are you promoting sustainability?

1-5 ranking based on number of "yes" answers.



The Principles

1. Do these 10 principles make sense to you?  
-Want it to not be profit driven — need it to be mandated by unbiased party to trust this  
-What's in it for me (business strategist)? Ultimately better my bottom line as a company  
-Promise: help me increase my bottom line in a highly regulated industry  
-Goal is it make it an incentive, not punishment in regulated industry  
-Being on-board allows me to get the benefits from my compliance  
-They make sense if I'm already bought in  
-Sell the reason to believe before offering the principles

2. Are there any that you think should be removed, de-prioritized, or changed?  
This comes off as being very progressive  
-neutralize the language, stay politically correct

1. Prioritizing humanity - very lofty, might write off based on language

3. Do all of the descriptions below the principles make sense?  
(see notes below)


The Ranking

For each principle, there is a ranking system of 5 yes/no questions that will indicate an Ethical Rating of 1-5 (5 is best, 1 is worst) of the technology.

1. Does the ranking system help you better understand the principle it refers to?  
yes

2. Could the system of ranking be improved? (Does 1-5 overall or 1-5 on each work better for you?)  
Better to ask them to order them in terms of priority — tool to create the strategic plan  
-yes/no doesn't help them grow  
-Understand where you are and lets them know the org is adaptable and flexible to working with me — not all or nothing

3. Which questions are the most important to include? The list will be narrowed to 5 questions per principle.  
a. 1, 2, 3, 4, 5, 10  
Note: Be mindful of tribalism - aka too progressive language  
b. 2, 3, 4, 5, 6  
Note: what does flourish mean?  
Note: "Net positive impact" is too vague



10 Principles

Meta Principle: The benefits of the technology intervention never overcome the benefits for humanity. Ultimately, the technology intervention creates overall improvement in humanity and society.

Ethical AI technology development and implementation will:

1. Reprioritize humanity

Success metrics should be based on the improvement of the human condition with respect to the protection and promotion of international human rights and cultural norms.

Are you prioritizing humanity?

1-5 ranking based on number of "yes" answers.

1. Competence: Does the entity identify clear success indicators for the technology that go beyond gross domestic product, consumption, and safety?

2. Competence: Do success indicators and business goals address holistic long-term improvements to the customer, environment, and society?

3. Competence: Are unintended negative consequences identified and accounted for in business goals and human-focused indicators?

4. Impact: Is the entire lifespan of the technology understood and considered?

5. Impact: Is the entity aware of all of the individuals affected by the technology, including those who are not direct customers?

6. Impact: Does the technology increase the quality of life for all individuals affected by it?

7. Impact: Do interactions with the technology avoid harmful dampening or amplification of human emotional experiences?

8. Culture: Does the entity's approach to building trust prioritize the interests, experiences, and professional standards of the region?

9. Culture: Does the entity's approach to building trust prioritize societal norms and government controls of the region?

10. Regulation: Does the entity respect the international laws where the technology exists?

Rating: \_\_\_\_\_


2. Promote sustainability

Entities and individuals should work toward common goals of environmental well-being by adhering to internationally established indicators of societal flourishing, see [UN Sustainable Development Goals](#).

Are you promoting sustainability?

1-5 ranking based on number of "yes" answers.

1. Competence: Does the entity provide employees responsible for the technology with the means to flourish in work and personal life?



The Principles

1. Do these 10 principles make sense to you? Yes

2. Are there any that you think should be removed, de-prioritized, or changed? No

3. Do all of the descriptions below the principles make sense? Yes

The Ranking


For each principle, there is a ranking system of 5 yes/no questions that will indicate an Ethical Rating of 1-5 (5 is best, 1 is worst) of the technology.

1. Does the ranking system help you better understand the principle it refers to? The ranking system somewhat confused me and I had to re-read everything a couple of times. It seems like you answer 5 questions and the aggregate number of "yes" answers gives you a score vs. answering each individual question as a "yes" on a 1 through 5 scale. The latter may allow for some more nuance vs. a binary "yes/no" response to fairly complex questions within an organization.

2. Which questions are the most important to include? The list will be narrowed to 5 questions per principle.  
a. 1, 1, 3, 6, 8, 10  
b. 2, 2, 3, 5, 1  
c. 3, 2, 3, 5, 6, 1  
d. 4, 1, 5, 6, 7, 8  
e. 5, 1, 2, 3, 5, 6  
f. 6, 3, 4, 5, 2, 1  
g. 7, 1, 2, 5, 6, 7  
h. 8, 2, 3, 4, 5, 1  
i. 9, 2, 3, 5, 1, 4  
j. 10, 2, 1, 5, 6, 4

3. Are there any questions you would add to the principles that are important to creating ethical AI technologies? No

4. Is this ranking system beneficial to you or your customers? Perhaps look at an enhanced ranking system that takes into account degrees of implementation. For instance, I could answer "yes" on couple of these questions even though they are partially true or if I don't have a comprehensive grasp of the principal. It could make my organization look at competent as a premiere organization with a much greater grasp on the subject matter.



10 Principles

Meta Principle: The benefits of the technology intervention never overcome the benefits for humanity. Ultimately, the technology intervention creates overall improvement in humanity and society.

Ethical AI technology development and implementation will:

1. Reprioritize humanity

Success metrics should be based on the improvement of the human condition with respect to the protection and promotion of international human rights and cultural norms.

Are you prioritizing humanity?

1-5 ranking based on number of "yes" answers.

1. Competence: Does the entity identify clear success indicators for the technology that go beyond gross domestic product, consumption, and safety?

2. Competence: Do success indicators and business goals address holistic long-term improvements to the customer, environment, and society?

3. Competence: Are unintended negative consequences identified and accounted for in business goals and human-focused indicators?

4. Impact: Is the entire lifespan of the technology understood and considered?

5. Impact: Is the entity aware of all of the individuals affected by the technology, including those who are not direct customers?

6. Impact: Does the technology increase the quality of life for all individuals affected by it?

7. Impact: Do interactions with the technology avoid harmful dampening or amplification of human emotional experiences?

8. Culture: Does the entity's approach to building trust prioritize the interests, experiences, and professional standards of the region?

9. Culture: Does the entity's approach to building trust prioritize societal norms and government controls of the region?

10. Regulation: Does the entity respect the international laws where the technology exists?

Rating: \_\_\_\_\_

2. Promote sustainability

Entities and individuals should work toward common goals of environmental well-being by adhering to internationally established indicators of societal flourishing, see [UN Sustainable Development Goals](#).


Are you promoting sustainability?

1-5 ranking based on number of "yes" answers.

1. Competence: Does the entity provide employees responsible for the technology with the means to flourish in work and personal life?

2. Impact: Does the technology have a net positive impact on the lives of the individuals using it?

3. Impact: Does the technology have a net positive impact on the lives of the individuals affected by it but not directly using it?



The Principle Alignment Assessment was validated through user interviews with stakeholders and industry experts to determine the most strategically insightful questionnaire for AI technology entities.

The **Moral Assessment** is a novel tool combining Verbeek’s Technology Mediation Theory and a consequentialist framework of moral reasoning to determine if the proposed technology is ethical or not.

BackgroundPrinciplesBusinessesGovernmentAbout

Moral Assessment - Results

Congratulations, you have completed a moral assessment of **Atlas Technology's Iris**. Download this summary to help guide ethical planning and decision-making as the entity develops these technology mediations.

Iris will learn routines, provide insights, and automate home management for members of a household to create a world where everyone is safe and connected.

Atlas Technologies will ensure that remote home management, control, and safety is a reality by centralizing software, connecting remotely, and prioritizing cybersecurity, so that the impact of Iris on all people living and working in an Iris connected home is ethical.

Summary

Background

Impact

Mediations

Outcomes

Share on Facebook

Share on Twitter

Download

Exit

Copyright © 2020 Shelbi Howard

Results provide a vision for the technology’s impact and a clear roadmap for the technology to meet entity goals.

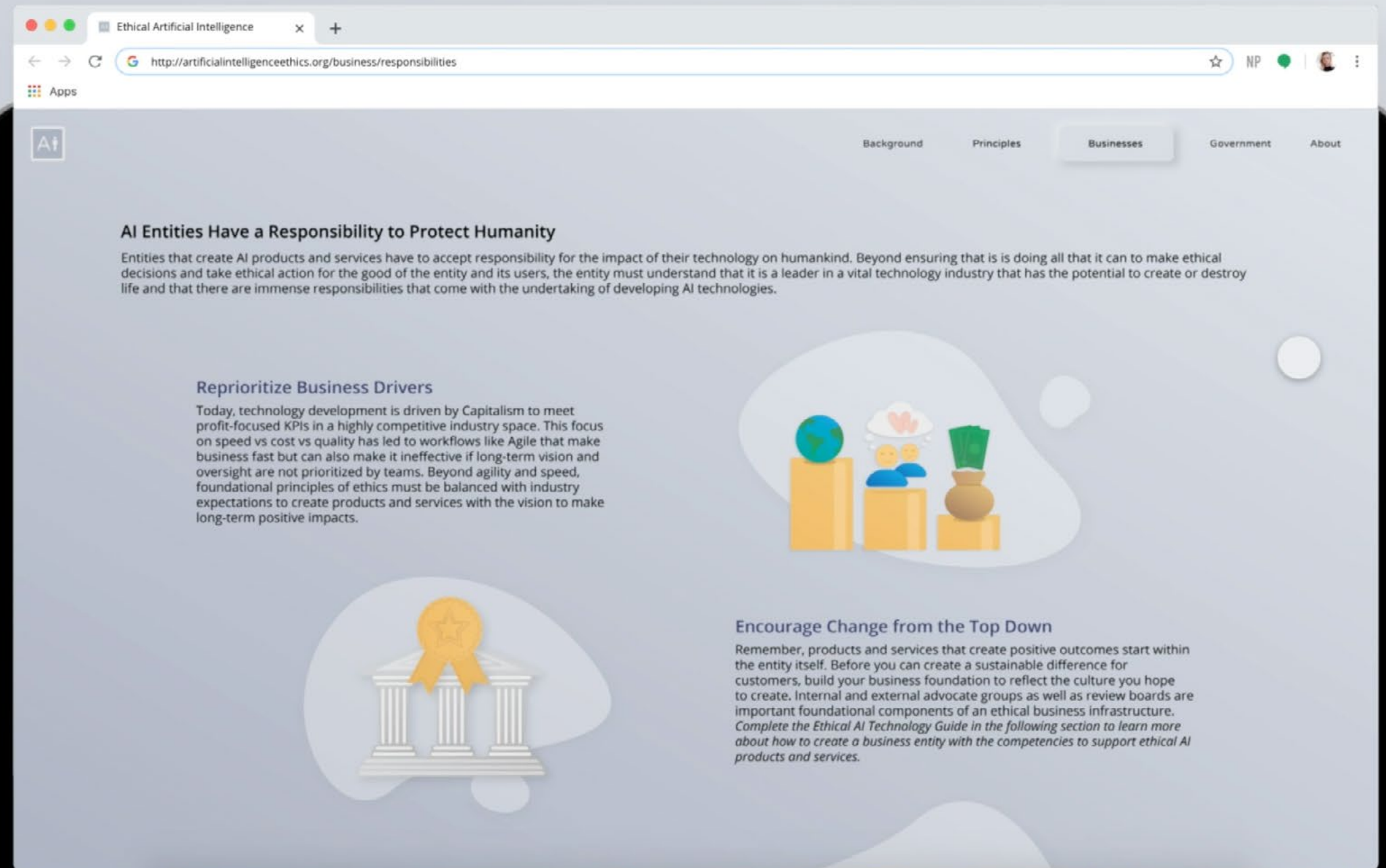
The diagram illustrates the four-step process of the Moral Assessment tool, showing the progression from background information to final results. Each step is represented by a screenshot of the tool's interface, with arrows indicating the flow and 'Click' labels pointing to the 'Next' buttons.

- Background (1/4):** The first screen shows the 'Background' tab selected. It includes sections for 'Describe the Stakeholders' (Individuals, Technology, Society) and 'Briefly describe the technology'. The 'Next' button is highlighted.
- Impact (2/4):** The second screen shows the 'Impact' tab selected. It includes sections for 'What is the good you're trying to maximize?' and 'How will this good be maximized?'. The 'Next' button is highlighted.
- Mediations (3/4):** The third screen shows the 'Mediations' tab selected. It includes sections for 'Describe the factors contributing to the technology mediations' and 'Describe the mediations implemented by the technology'. The 'Next' button is highlighted.
- Outcomes (4/4):** The fourth screen shows the 'Outcomes' tab selected. It includes sections for 'Outcomes' (Utopia, Dystopia) and 'Based on the mediations previously described, what are the potential outcomes these will have?'. The 'Next' button is highlighted.

A 'Next Page' box is positioned between the Impact and Mediations screens, and a 'Results' box is at the end of the flow.



AI entities' **responsibilities** provide insight into the implications of AI technology for humans and their ethical imperative.





Home button



Background

Principles

Businesses

Government

About

Mediating Tech  
Business Strategy  
Ethical AI Guide  
Responsibilities



Sticky Nav Bar

### AI Entities Have a Responsibility to Protect Humanity

Entities that create AI products and services have to accept responsibility for the impact of their technology on humankind. Beyond ensuring that it is doing all that it can to make ethical decisions and take ethical action for the good of the entity and its users, the entity must understand that it is a leader in a vital technology industry that has the potential to create or destroy life and that there are immense responsibilities that come with the undertaking of developing AI technologies.

### Reprioritize Business Drivers

Today, technology development is driven by Capitalism to meet profit-focused KPIs in a highly competitive industry space. This focus on speed vs cost vs quality has led to workflows like Agile that make business fast but can also make it ineffective if long-term vision and oversight are not prioritized by teams. Beyond agility and speed, foundational principles of ethics must be balanced with industry expectations to create products and services with the vision to make long-term positive impacts.



Microanimation 1

Microanimation 2



Animate



### Encourage Change from the Top Down

Remember, products and services that create positive outcomes start within the entity itself. Before you can create a sustainable difference for customers, build your business foundation to reflect the culture you hope to create. Internal and external advocate groups as well as review boards are important foundational components of an ethical business infrastructure. Complete the Ethical AI Technology Guide in the following section to learn more about how to create a business entity with the competencies to support ethical AI products and services.

### Protect the Health, Safety, and Rights of Humanity

AI technology will change the course of humanity. It is up to the leaders of AI to build an infrastructure that benefits humankind by ensuring ethical implementation of mediating AI technologies and prioritizing humans in the products and services they create. Learn more about how to do this with the following tools to help entities learn about, design for, and provide continued support of ethical AI technology businesses, products, and communities.



Microanimation 3

"Much has been written about AI's potential to reflect both the best and the worst of humanity. We have seen AI providing conversation and comfort to the lonely; we have also seen AI engaging in racial discrimination... As leaders, it is incumbent on all of us to make sure we are building a world in which every individual has an opportunity to thrive."

— Andrew Ng, Co-founder and lead of Google Brain.

Next: About the Project

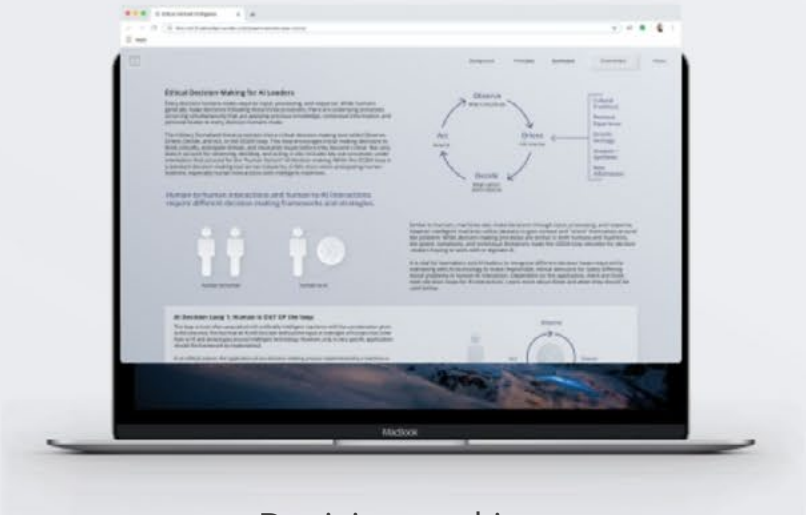


Next Button

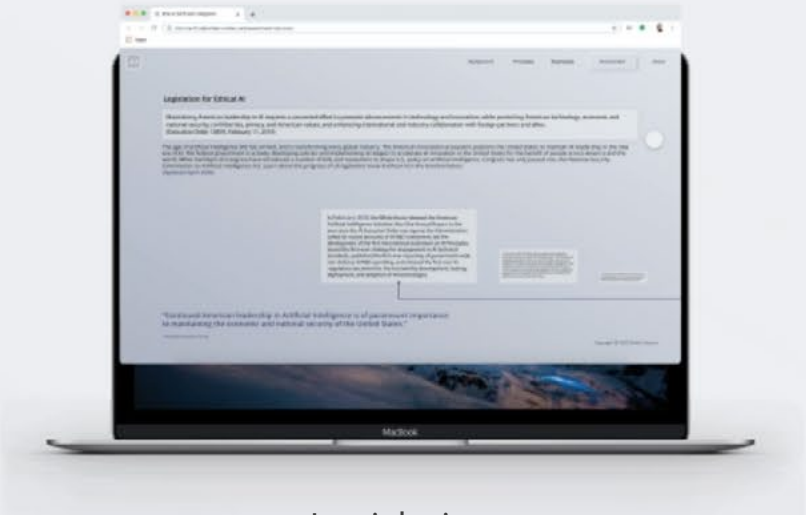
Appear at  
page bottom

Wireframes show the  
**Responsibilities** experience for businesses.

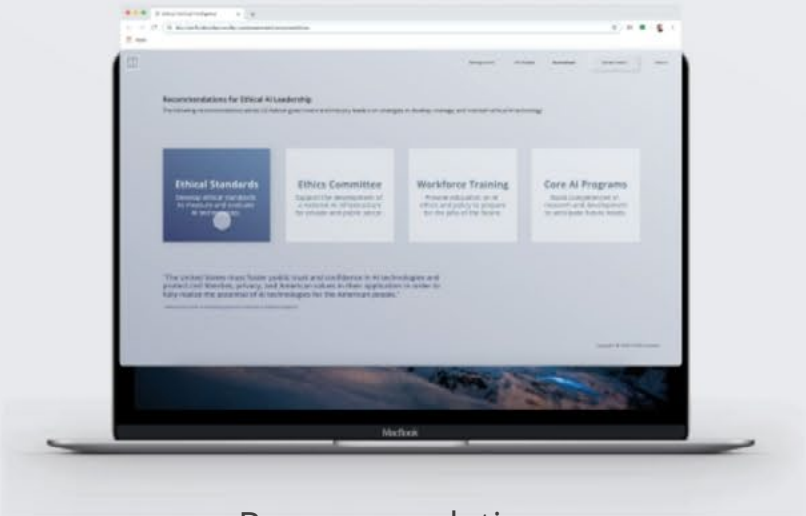
Governments are given the resources to make educated decisions for AI technology legislation.



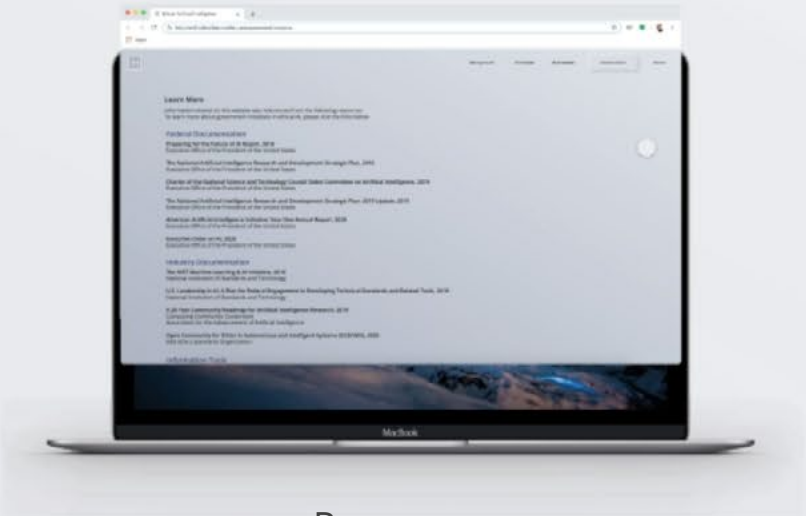
Decision-making



Legislation



Recommendations



Resources

Legislators must understand how **decision-making** needs differ in AI technologies to determine the best trajectory for each industry's needs.

Ethical Artificial Intelligence

http://artificialintelligenceethics.org/government/decision-making

Apps

BackgroundPrinciplesBusinessesGovernmentAbout

### Ethical Decision-Making for AI Leaders

Every decision humans make requires input, processing, and response. While humans generally make decisions following these three processes, there are underlying processes occurring simultaneously that are applying previous knowledge, contextual information, and personal biases to every decision humans make.

The military formalized these processes into a critical decision-making tool called Observe, Orient, Decide, and Act, or the OODA loop. This loop encourages those making decisions to think critically, anticipate threats, and neutralize issues before they become critical. Not only does it account for observing, deciding, and acting, it also includes key sub-processes under orientation that account for the "human factors" of decision-making. While the OODA loop is a standard decision-making tool across industries, it falls short when anticipating human-machine, especially human interactions with intelligent machines.

#### Human-to-human interactions and human-to-AI interactions require different decision-making frameworks and strategies.

human-to-human

human-to-AI

```
graph TD; Observe[Observe  
What is around you] --> Orient[Orient  
Add meaning]; Orient --> Decide[Decide  
Weigh options about response]; Decide --> Act[Act  
Respond]; Act --> Observe; Context[Cultural Traditions  
Previous Experience  
Genetic Heritage  
Analysis + Synthesis  
New Information] --> Orient
```

Similar to humans, machines also make decisions through input, processing, and response, however intelligent machines utilize datasets to gain context and "orient" themselves around the problem. While decision-making processes are similar in both humans and machines, the speed, complexity, and contextual limitations make the OODA loop obsolete for decision-makers hoping to work with or legislate AI.

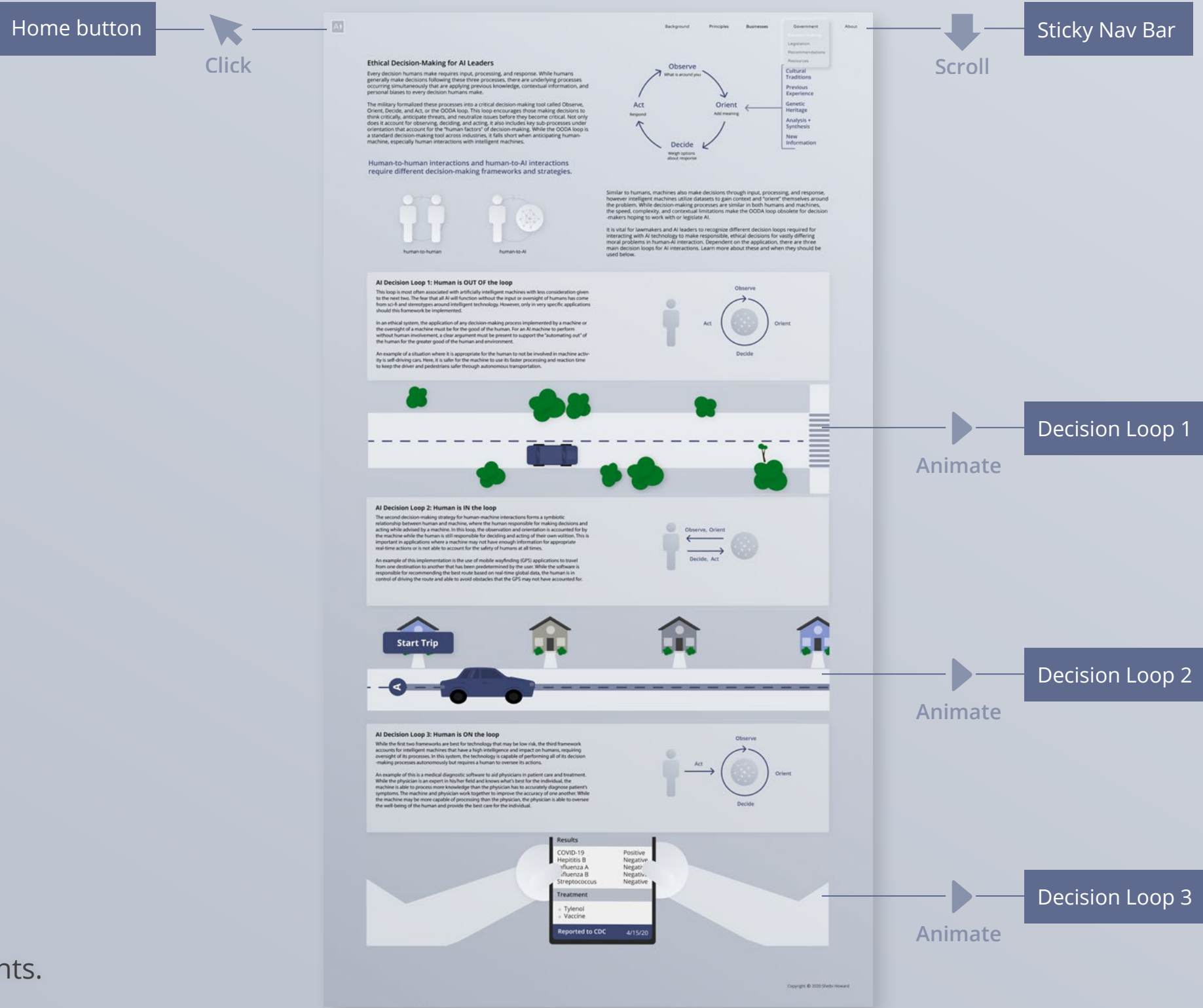
It is vital for lawmakers and AI leaders to recognize different decision loops required for interacting with AI technology to make responsible, ethical decisions for vastly differing moral problems in human-AI interaction. Dependent on the application, there are three main decision loops for AI interactions. Learn more about these and when they should be used below.

#### AI Decision Loop 1: Human is OUT OF the loop

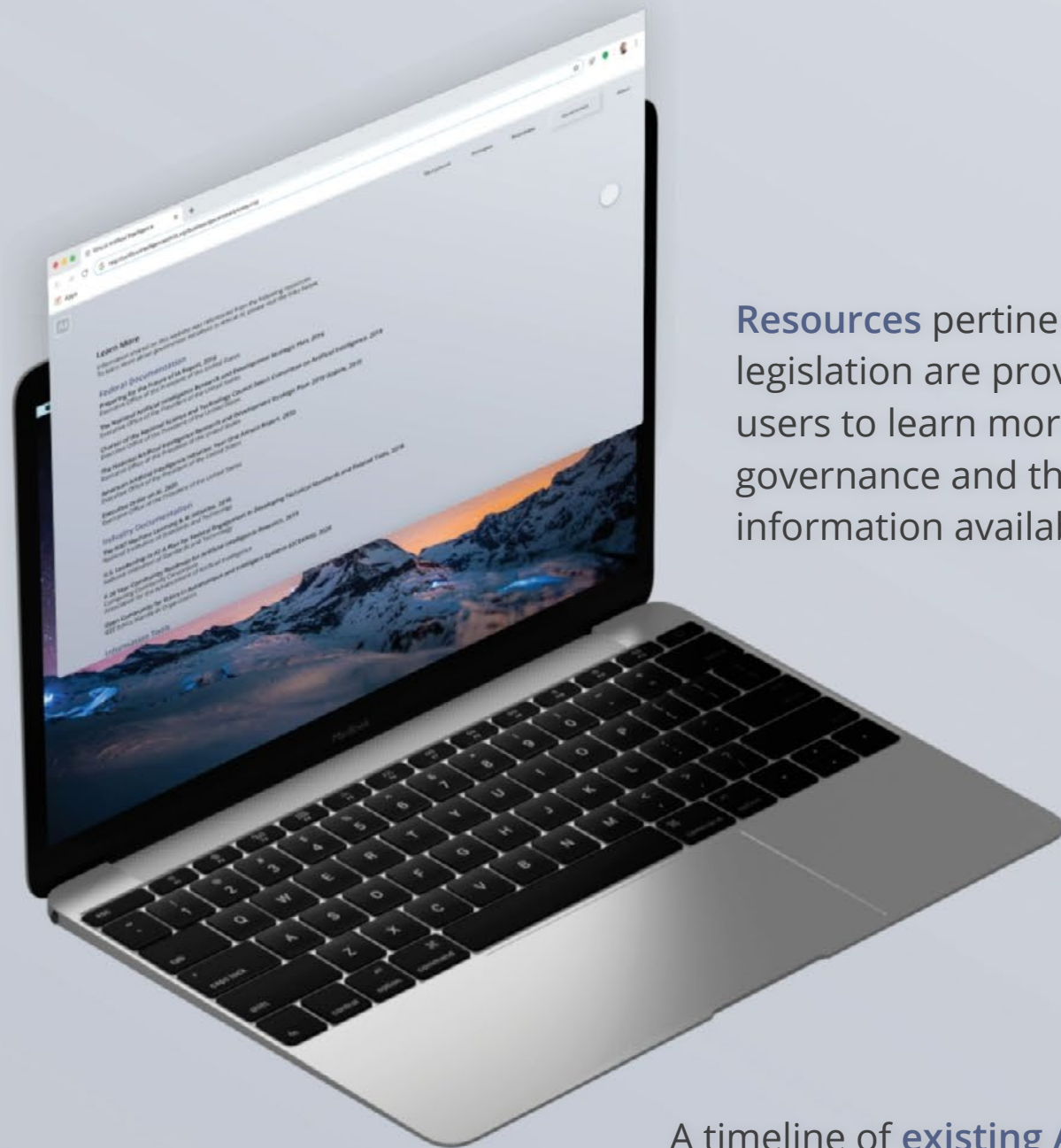
This loop is most often associated with artificially intelligent machines with less consideration given to the next two. The fear that all AI will function without the input or oversight of humans has come from sci-fi and stereotypes around intelligent technology. However, only in very specific applications should this framework be implemented.

```
graph TD; Observe[Observe];
```





Wireframes show the **decision-making** experience for governments.

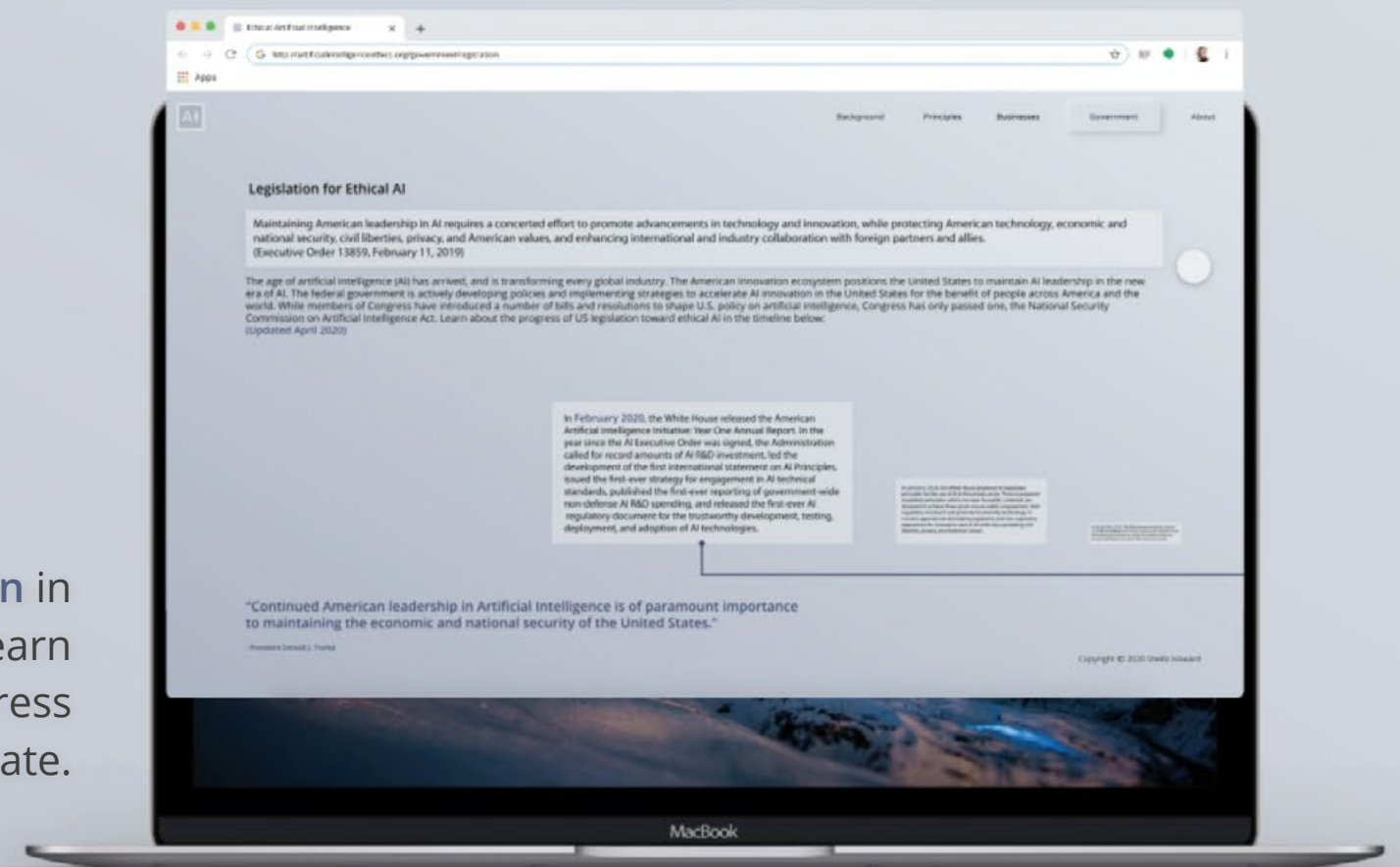


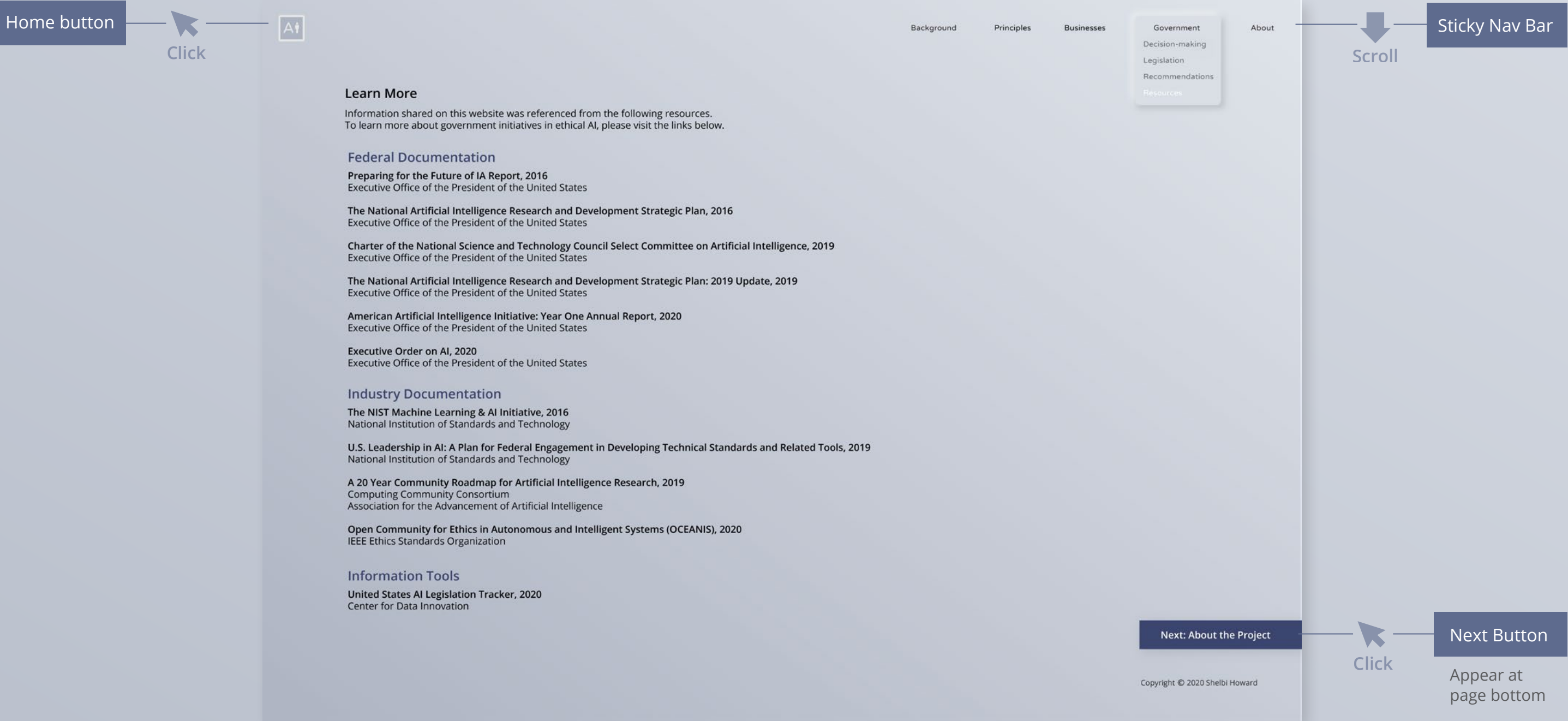
Resources pertinent to AI legislation are provided for users to learn more about AI governance and the origin of information available here.

A timeline of **existing AI legislation** in the US is available for users to learn about policy decisions and progress made in AI governance to date.

“In this era of profound digital transformation, it’s important to remember that business, as well as government, has a role to play in creating shared prosperity — not just prosperity. After all, the same technologies that can be used to concentrate wealth and power can also be used to distribute it more widely and empower more people.”

— Erik Brynjolfsson, Director of the MIT Initiative on the Digital Economy





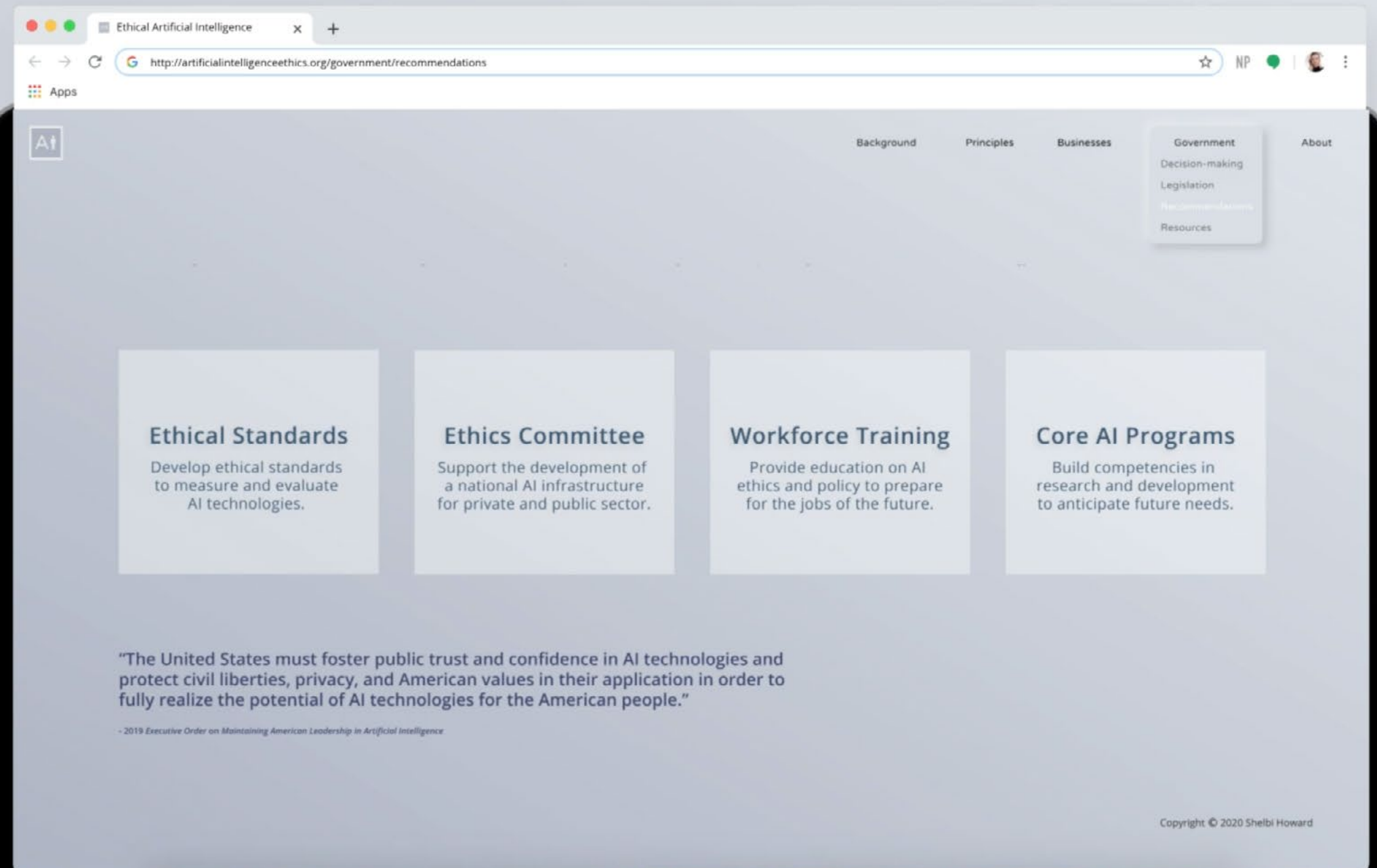
Wireframes show the **resources** experience for governments.

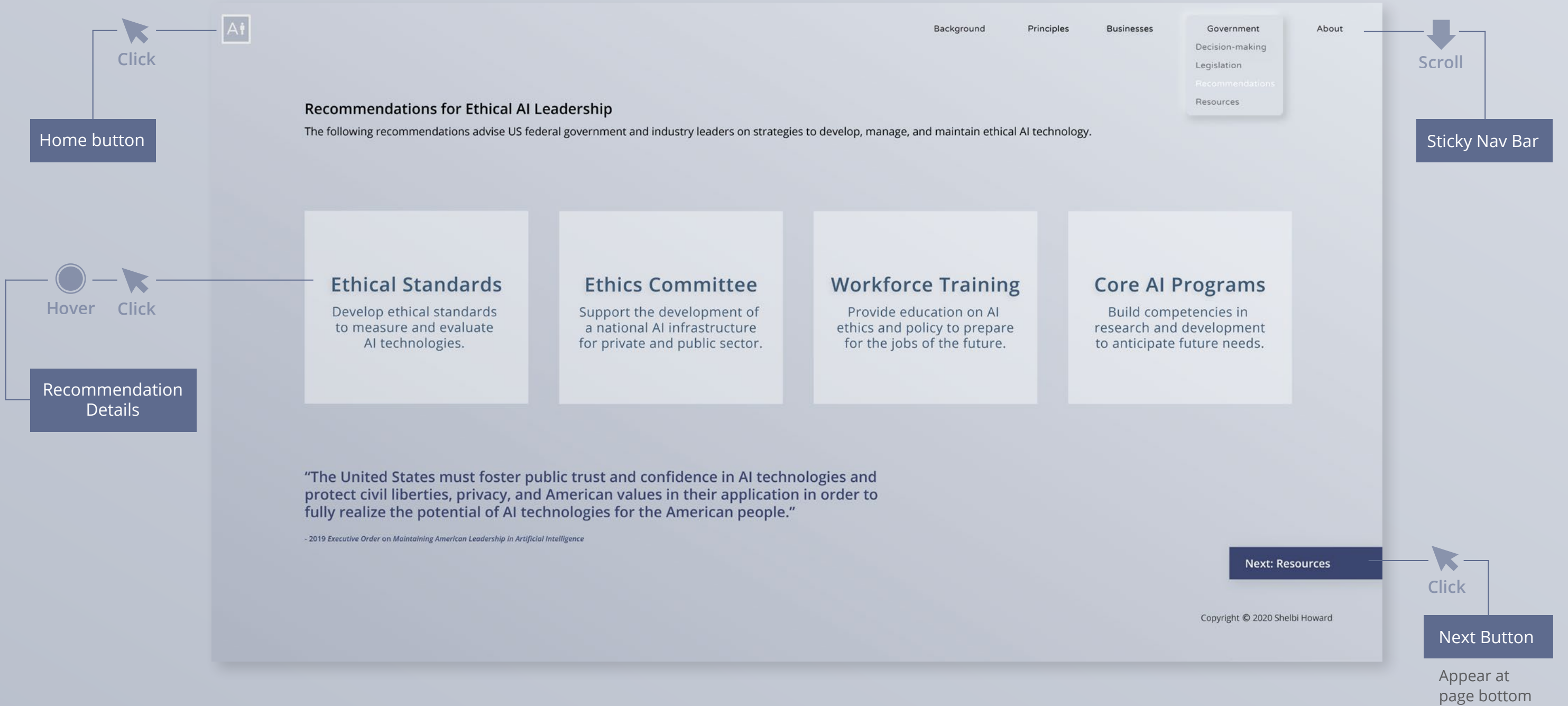


Wireframes show the **AI Legislation** experience for governments.



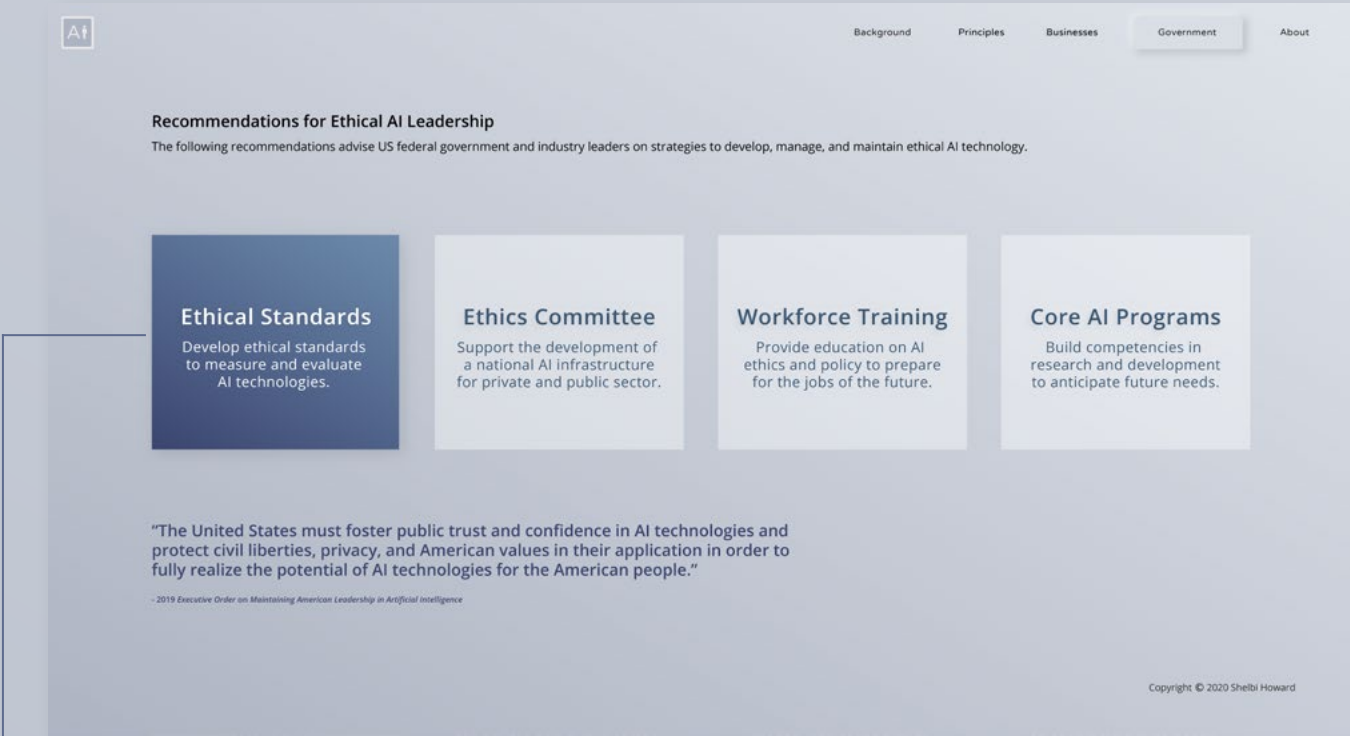
Legislative **recommendations** aggregate executive decisions and current AI trajectories to provide a future action plan for US policy-makers.



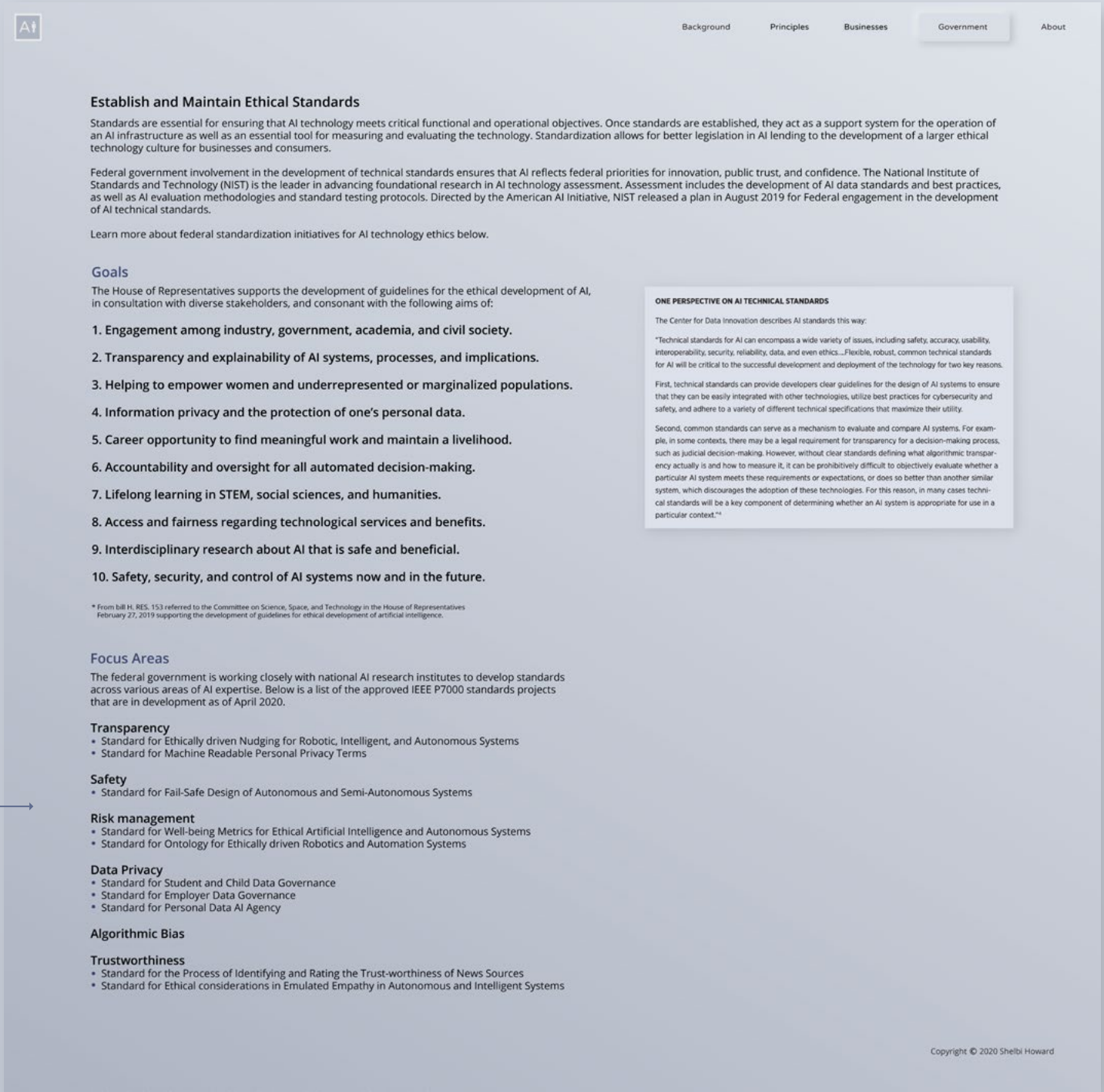


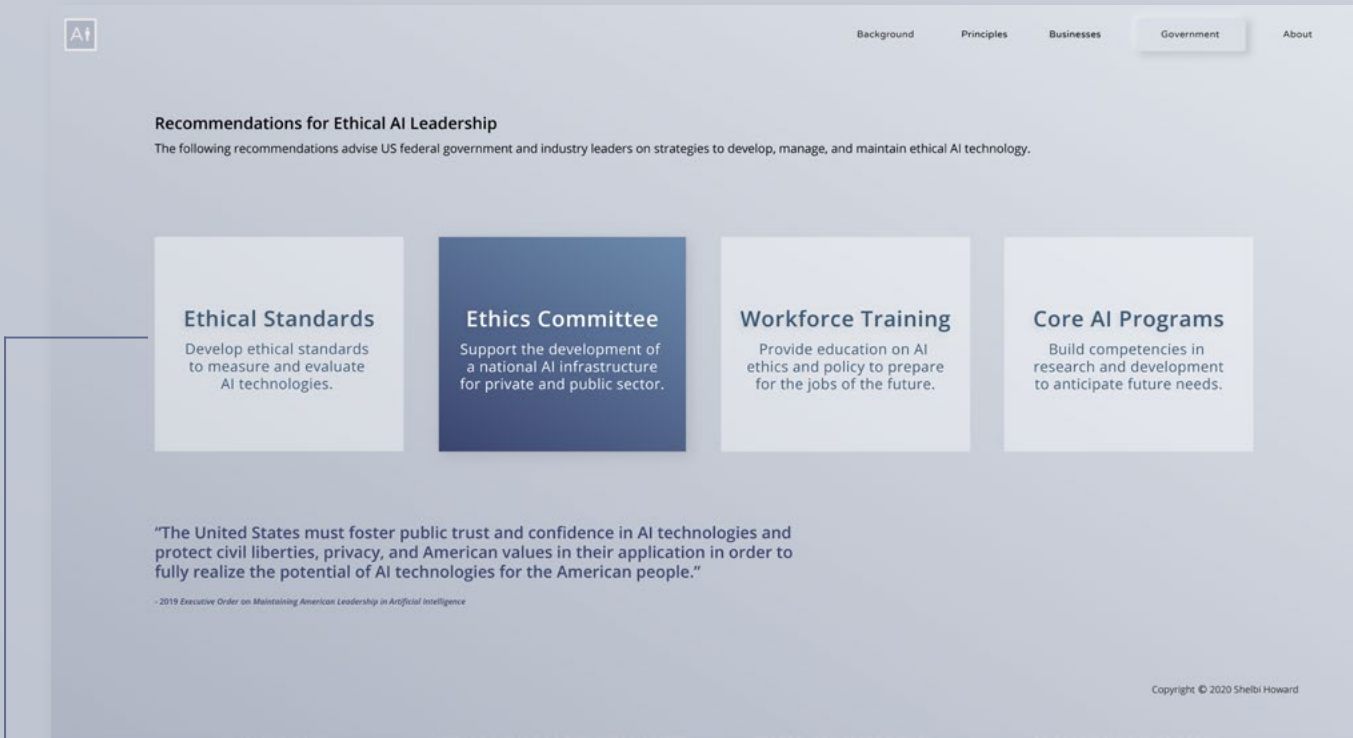
Wireframes show the **recommendations** experience for governments.





Wireframes show the **recommendations** experience for governments.





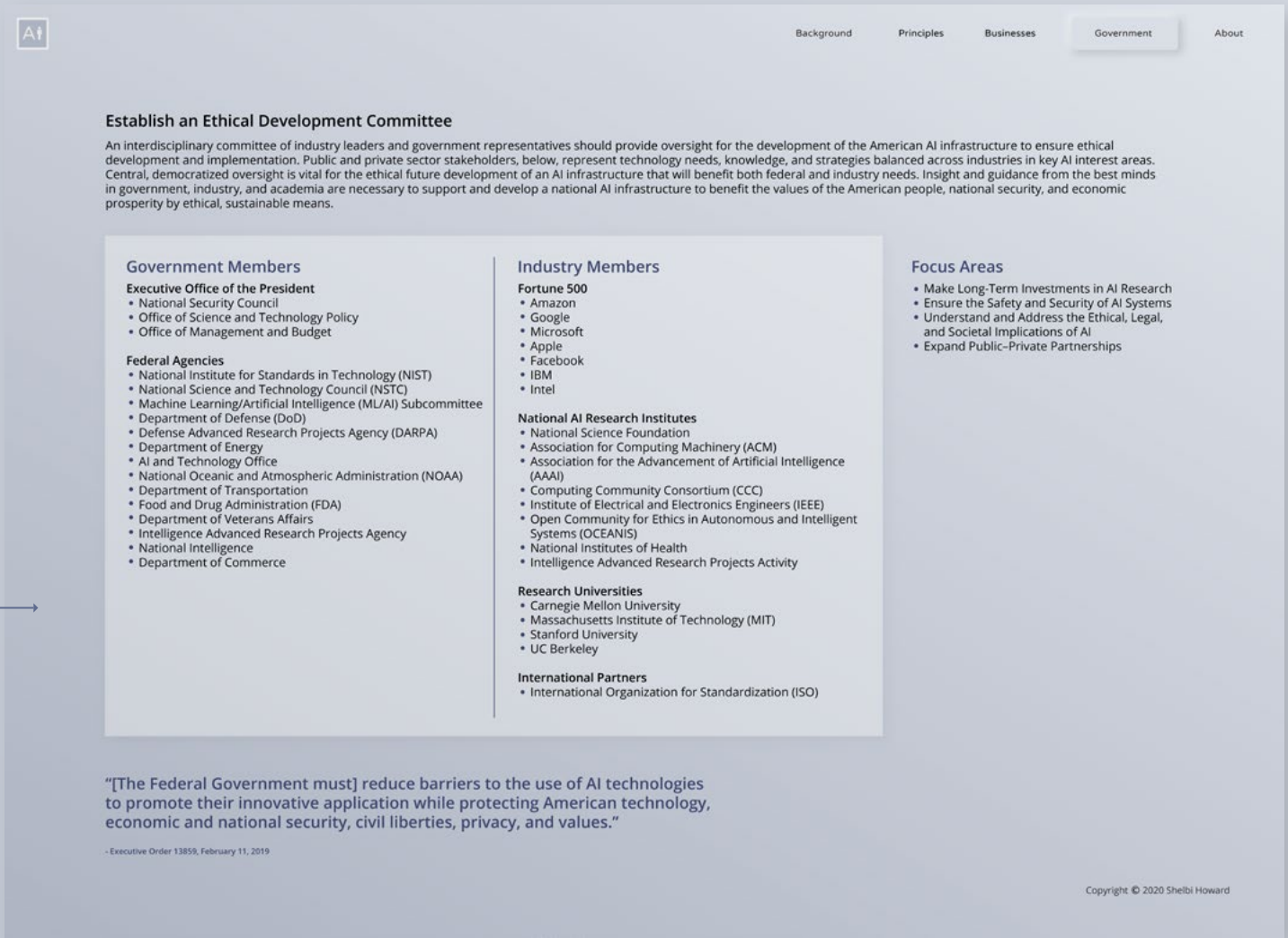
Link to Recommendation Details



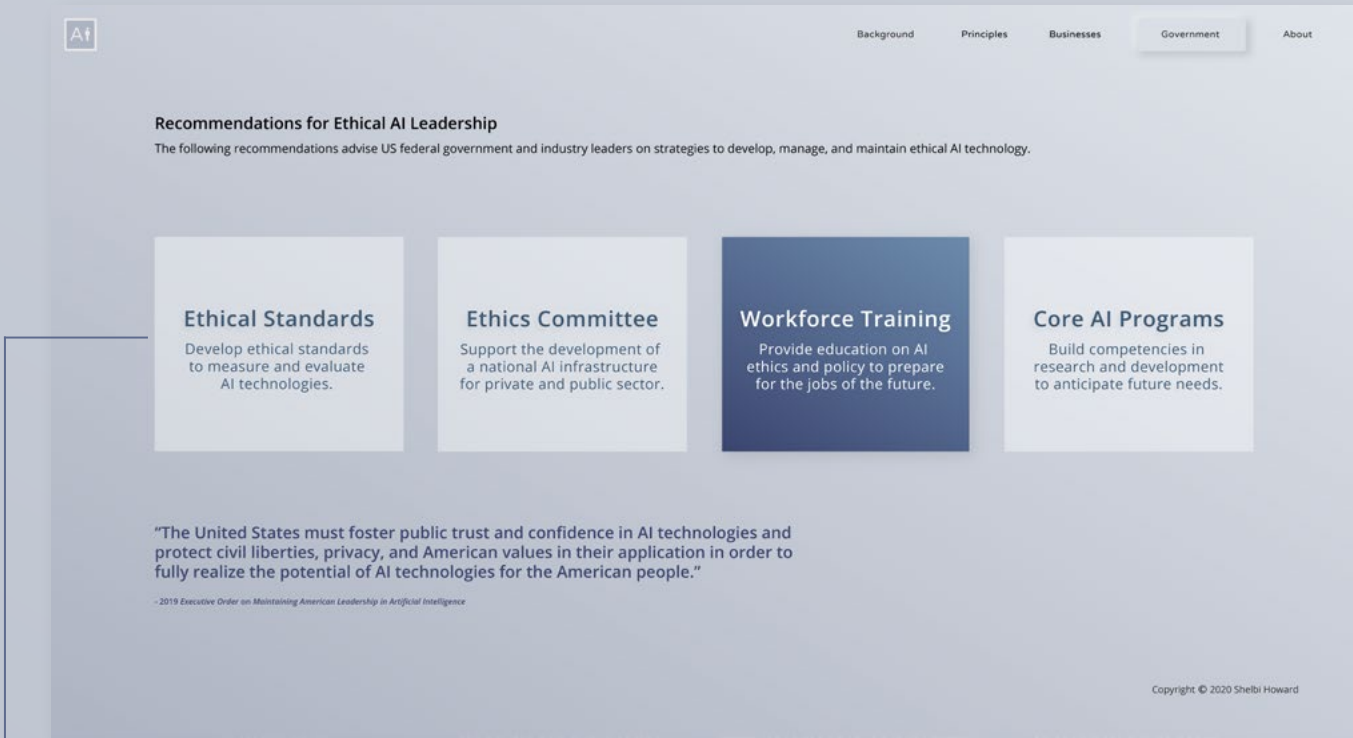
Click



Hover



Wireframes show the **recommendations** experience for governments.



Link to Recommendation Details

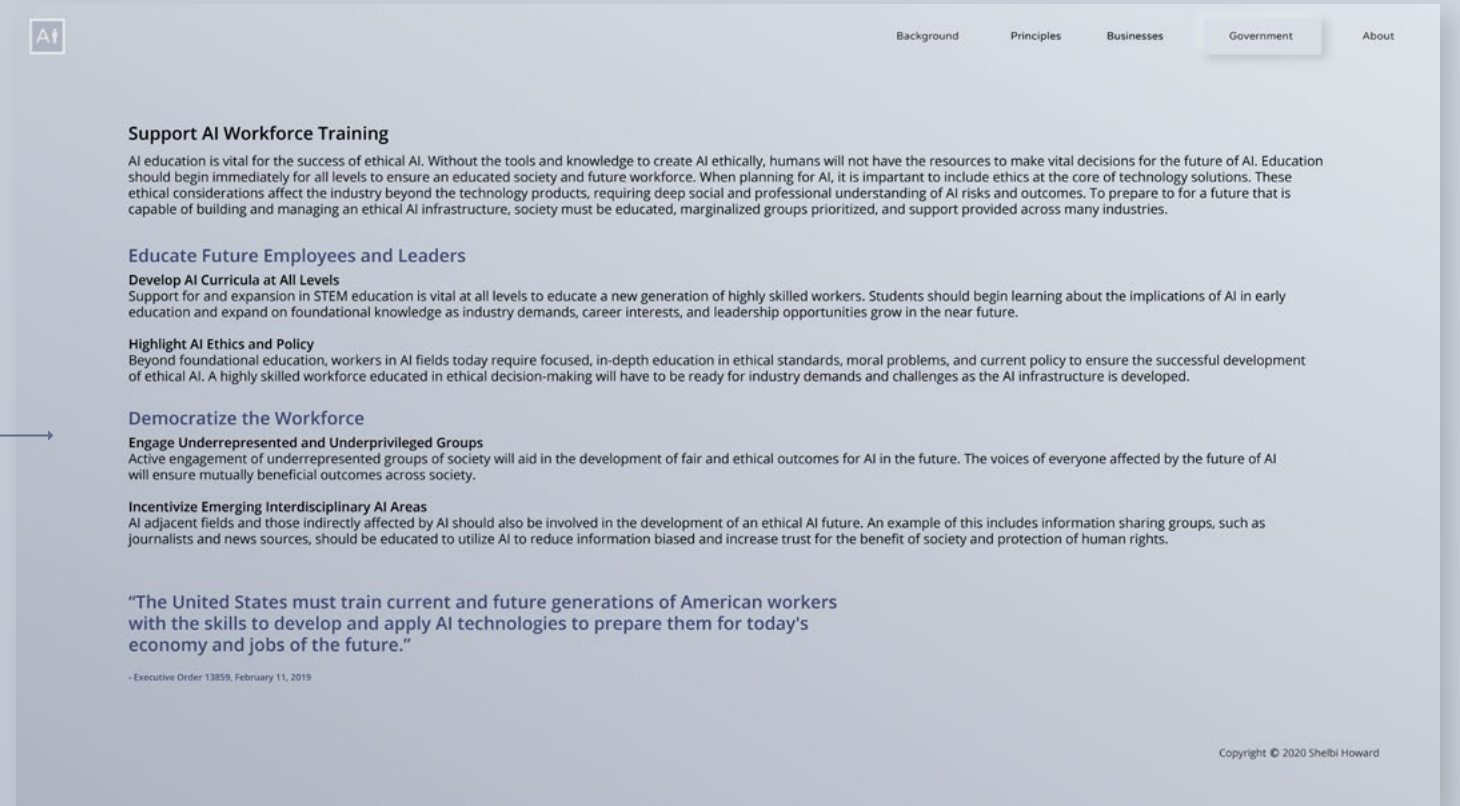


Click

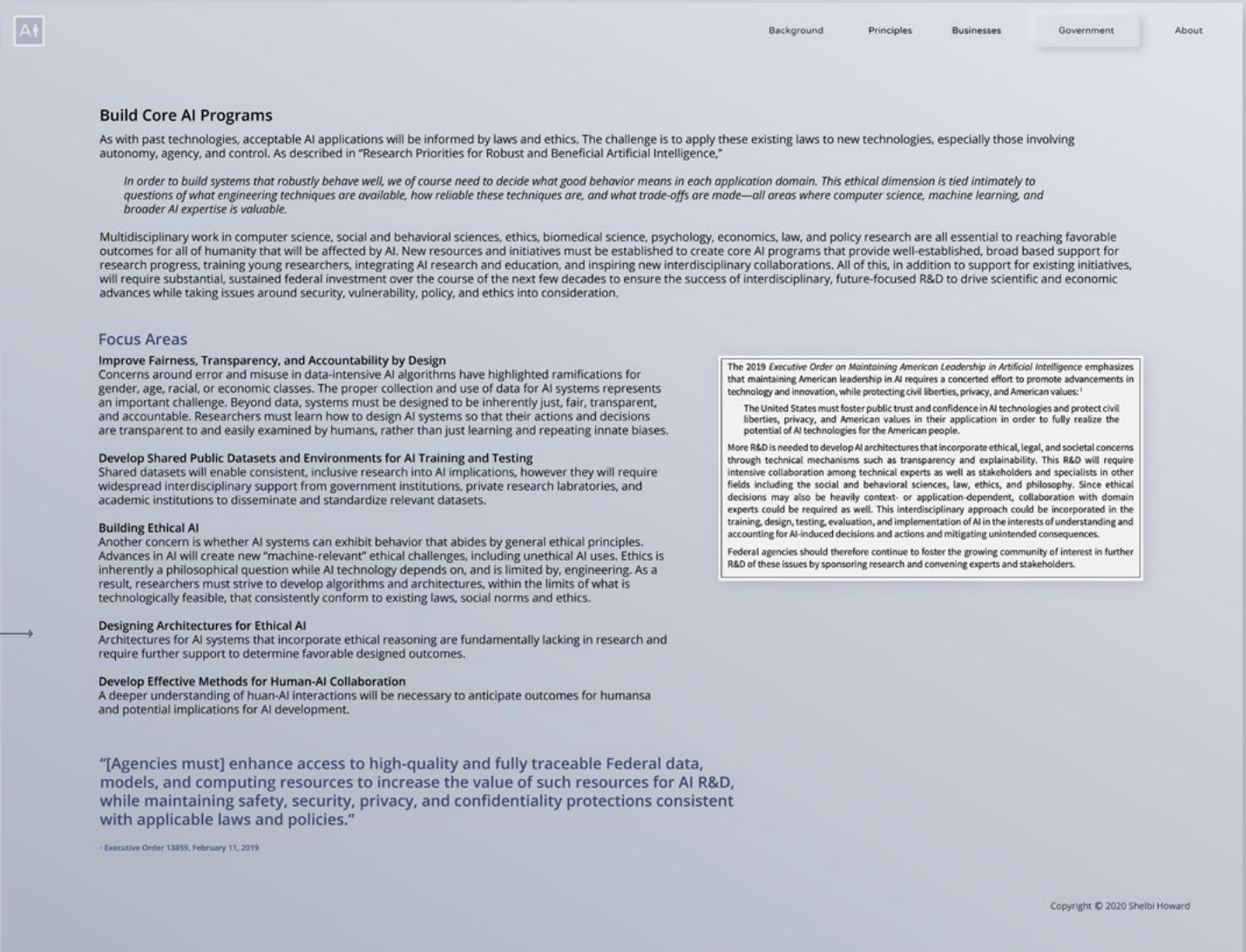
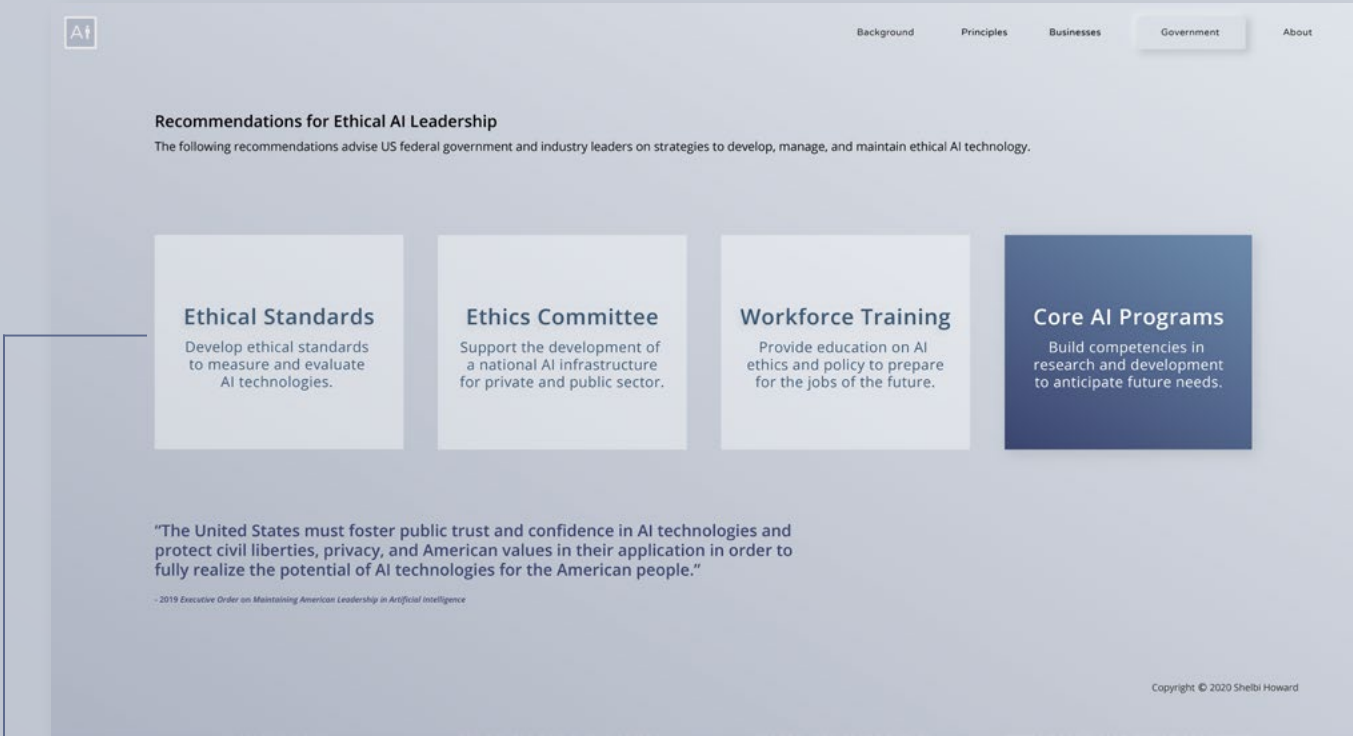


Hover

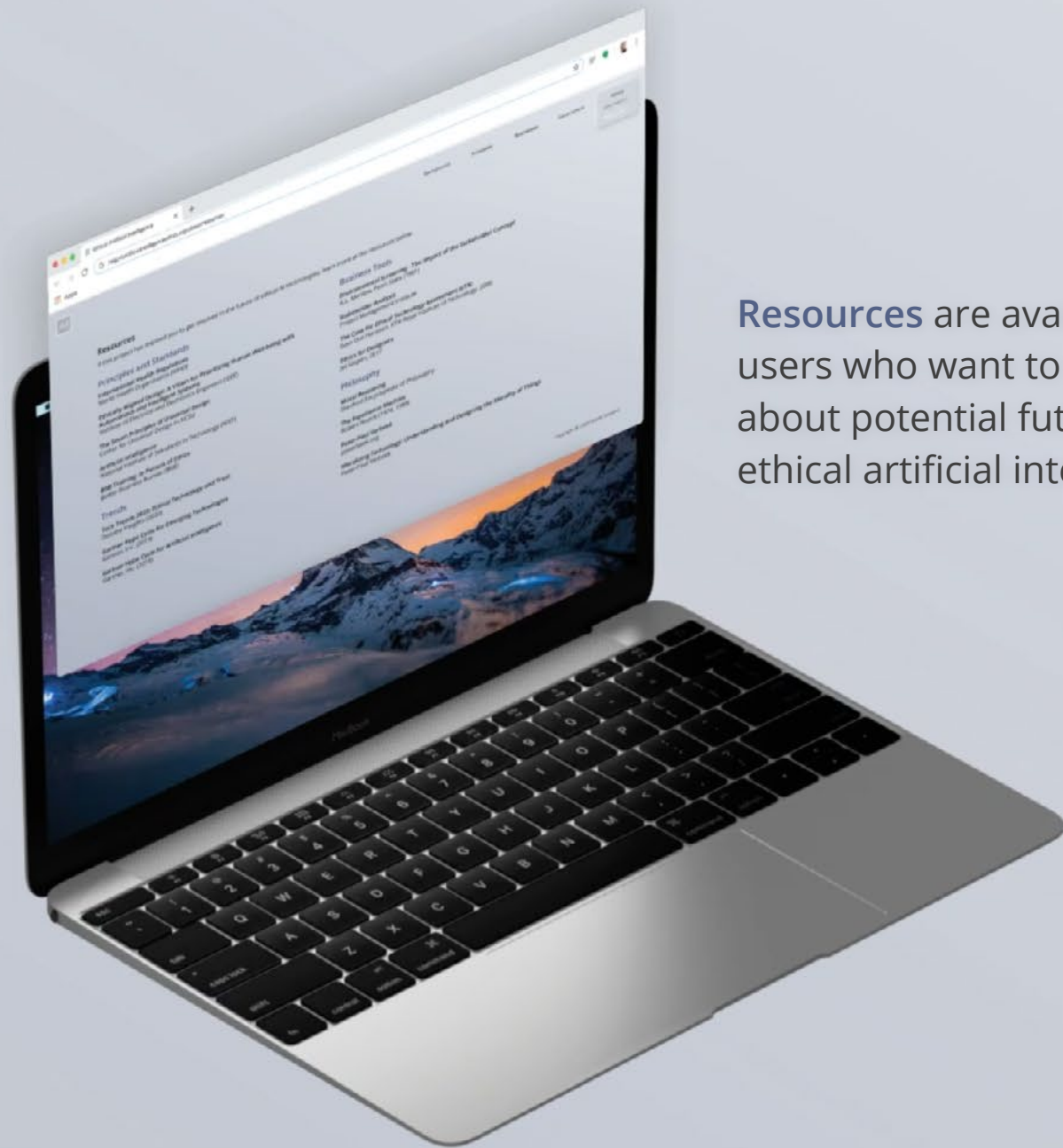
Wireframes show the **recommendations** experience for governments.



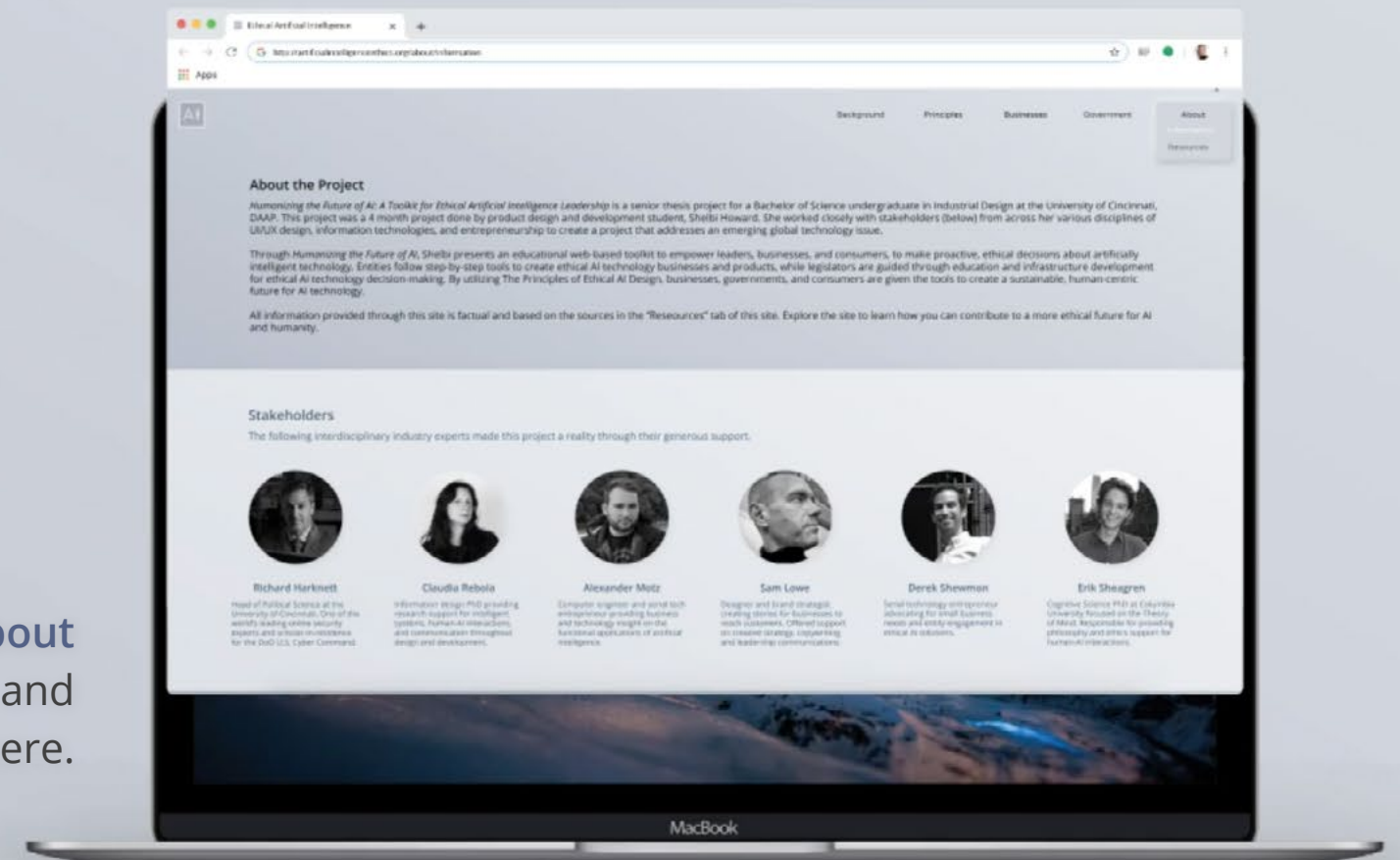




Wireframes show the recommendations experience for governments.



Resources are available for users who want to learn more about potential futures for ethical artificial intelligence.



Users access information about the project, stakeholders, and future considerations here.



Home button



Background Principles Businesses Government About Information Resources



Sticky Nav Bar

#### About the Project

*Humanizing the Future of AI: A Toolkit for Ethical Artificial Intelligence Leadership* is a senior thesis project for a Bachelor of Science undergraduate in Industrial Design at the University of Cincinnati, DAAP. This project was a 4 month project done by product design and development student, Shelbi Howard. She worked closely with stakeholders (below) from across her various disciplines of UI/UX design, information technologies, and entrepreneurship to create a project that addresses an emerging global technology issue.

Through *Humanizing the Future of AI*, Shelbi presents an educational web-based toolkit to empower leaders, businesses, and consumers, to make proactive, ethical decisions about artificially intelligent technology. Entities follow step-by-step tools to create ethical AI technology businesses and products, while legislators are guided through education and infrastructure development for ethical AI technology decision-making. By utilizing The Principles of Ethical AI Design, businesses, governments, and consumers are given the tools to create a sustainable, human-centric future for AI technology.

All information provided through this site is factual and based on the sources in the "Resources" tab of this site. Explore the site to learn how you can contribute to a more ethical future for AI and humanity.

#### Stakeholders

The following interdisciplinary industry experts made this project a reality through their generous support.



Richard Harknett

Head of Political Science at the University of Cincinnati. One of the world's leading online security experts and scholar-in-residence for the DoD U.S. Cyber Command.



Claudia Rebola

Information design PhD providing research support for intelligent systems, human-AI interactions, and communication throughout design and development.



Alexander Motz

Computer engineer and serial tech entrepreneur providing business and technology insight on the functional applications of artificial intelligence.



Sam Lowe

Designer and brand strategist creating stories for businesses to reach customers. Offered support on creative strategy, copywriting, and leadership communications.



Derek Shewmon

Serial technology entrepreneur advocating for small business needs and entity engagement in ethical AI solutions.



Erik Sheagren

Cognitive Science PhD at Columbia University focused on the Theory of Mind. Responsible for providing philosophy and ethics support for human-AI interactions.

#### Future Considerations

##### Build an Ethical Culture

The education provided through this website is only the beginning of creating a successful solution for ethical AI technology. As technologies and their entities are pressured to meet increasingly strict federal and international governance guidelines, consumer cultures will form in response to the shift. It will be essential to the ethical development of technologies to recognize the consumer's opinions, responses, and ultimately culture as an active driver in the future of AI technologies' success. Let's build a culture that encourages ethical behavior and holds one another accountable for our actions with and through AI technologies.



Microanimation 1

Microanimation 2



Animate



##### Automate Ethical AI Guidance

Beyond the initial purpose of this site to provide educational tools for businesses, government, and consumers, resources such as these could evolve into more prescriptive solutions as technologies become more intelligent. This website has been envisioned as a three-phase tool to continue providing guidance for these stakeholders in the future. First, the current solution, serves as a minimum viable product acting as a catalyst for change in ethical AI education and conversation.

Next, a web tool would be added to provide businesses with a place to submit business and technology information to governance groups for approval and Ethical AI Certification. This would help legitimize the Ethical AI effort by materializing the philosophical principles into solidified requirements. The process for review and acceptance would be managed by a governing party, similar to patent applications today.

Finally, as AI becomes more applicable, this application process would be automated to provide immediate results to entities about how to make technologies more ethical and if their technologies qualify for an Ethical AI Certification.



Shelbi Howard

User Advocate • Digital Designer • Web Developer

#### About the Designer

*The best solutions lie at the center of problem solving, technology innovation, and an entrepreneurial mindset.*

I am a user experience designer and developer with a passion for creating thoughtful solutions to complex problems. I use design to not only communicate problems, but also to better analyze and understand the importance of them. I will be graduating from the University of Cincinnati, DAAP in May 2020 and am currently seeking full-time employment.

In addition to design, I have minors in information technology, entrepreneurship, and fashion design. I hope to use my leadership, communication, and design thinking skills to help others analyze, communicate, and act on vital problems with immediate implications for humanity.

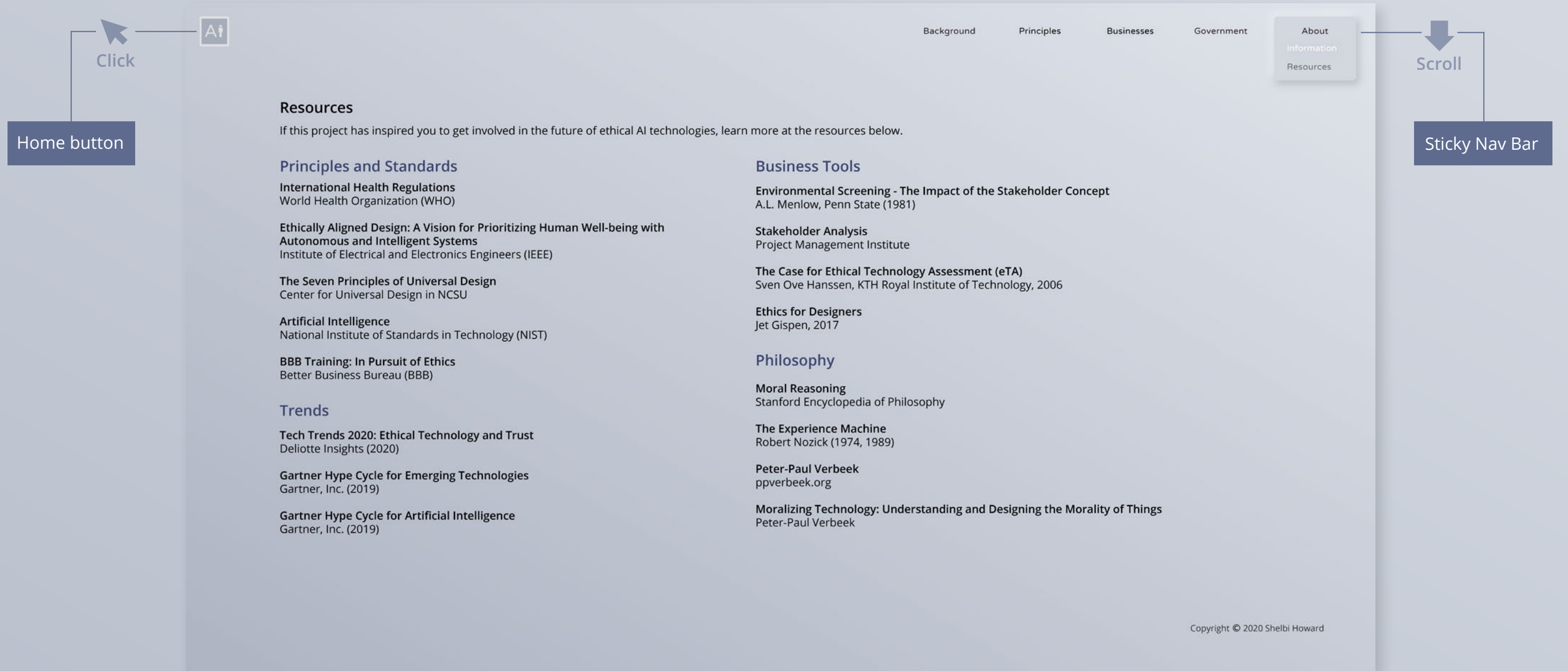
Love to talk about tech? Let's chat!



Copyright © 2020 Shelbi Howard

Wireframe shows **information** about the project.





Wireframe shows **resources** about the project.



Design is responsible for providing human advocacy.

**Technology will progress regardless of design decisions.**

Humans have to choose a proactive or reactive relationship.



**“AI will be the best or worst thing ever for humanity”**

- Elon Musk

# Future Considerations



## Build an Ethical Culture

This website is only the beginning of creating a successful solution. As technologies and their entities are pressured to meet governance guidelines, consumer cultures will form in response to the shift.

It will be essential to the ethical development of technologies to recognize the consumer's opinions, responses, and ultimately culture as an active driver in the future of AI technologies' success. Let's build a culture that encourages ethical behavior and holds one another accountable for our actions with and through AI technologies.

# Future Considerations

## Automate Ethical AI Guidance

This website has been envisioned as a three-phase tool to continue providing guidance for these stakeholders in the future.

1. The current solution serves as a minimum viable product, acting as a catalyst for change in ethical AI education and conversation.
2. A web-based business application would provide a place to submit technology information to governance groups for approval and Ethical Certification. The process for review and acceptance would be managed by a governing party, similar to patent applications today.
3. The application process would be automated with AI to provide immediate results to entities about how to make technologies more ethical and if their technologies qualify for an Ethical AI Certification.





To the experts, advocates, and leaders:

Thank you for your endless support that made this project a reality. Without your passion and intelligence I could not have explored AI solutions with the voracity and depth I hoped for. Your enthusiasm for learning and many hours of feedback have been invaluable to the success of this capstone.

# Advisors



**Richard Harknett**

Head of Political Science at the University of Cincinnati. One of the world's leading online security experts and scholar-in-residence for the DoD U.S. Cyber Command.



**Claudia Rebola**

Information design PhD providing research support for intelligent systems, human-AI interactions, and communication throughout design and development.



**Alexander Motz**

Computer engineer and serial tech entrepreneur providing business and technology insight on the functional applications of artificial intelligence.



**Sam Lowe**

Designer and brand strategist creating stories for businesses to reach customers. Offered support on creative strategy, copywriting, and leadership communications.



**Derek Shewmon**

Serial technology entrepreneur advocating for small business needs and entity engagement in ethical AI solutions.



**Erik Sheagren**

Cognitive Science PhD at Columbia University focused on the Theory of Mind. Responsible for providing philosophy and ethics support for human-AI interactions.

# Let's chat

Curious about technology and its implications?

Looking for a passionate UX designer?

Let me help you design a better future.



## Shelbi Howard

User Advocate • Digital Designer • Web Developer



[shelbihoward.com](https://shelbihoward.com)



[howardsb@mail.uc.edu](mailto:howardsb@mail.uc.edu)



[www.linkedin.com/in/shelbihoward/](https://www.linkedin.com/in/shelbihoward/)





# Business

## Principles and Standards

**International Health Regulations**  
World Health Organization (WHO)

**Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems**  
Institute of Electrical and Electronics Engineers (IEEE)

**The Seven Principles of Universal Design**  
Center for Universal Design in NCSU

**Artificial Intelligence**  
National Institute of Standards in Technology (NIST)

**BBB Training: In Pursuit of Ethics**  
Better Business Bureau (BBB)

## Trends

**Tech Trends 2020: Ethical Technology and Trust**  
Deloitte Insights (2020)

**Gartner Hype Cycle for Emerging Technologies**  
Gartner, Inc. (2019)

**Gartner Hype Cycle for Artificial Intelligence**  
Gartner, Inc. (2019)

## Business Tools

**Environmental Screening - The Impact of the Stakeholder Concept**  
A.L. Menlow, Penn State (1981)

**Stakeholder Analysis**  
Project Management Institute

**The Case for Ethical Technology Assessment (eTA)**  
Sven Ove Hanssen, KTH Royal Institute of Technology, 2006

**Ethics for Designers**  
Jet Gispen, 2017

## Philosophy

**Moral Reasoning**  
Stanford Encyclopedia of Philosophy

**The Experience Machine**  
Robert Nozick (1974, 1989)

**Peter-Paul Verbeek**  
ppverbeek.org

**Moralizing Technology: Understanding and Designing the Morality of Things**  
Peter-Paul Verbeek

# Government

## Federal Documentation

**Preparing for the Future of IA Report, 2016**  
Executive Office of the President of the United States

**The National Artificial Intelligence Research and Development Strategic Plan, 2016**  
Executive Office of the President of the United States

**Charter of the National Science and Technology Council Select Committee on Artificial Intelligence, 2016**  
Executive Office of the President of the United States

**The National Artificial Intelligence Research and Development Strategic Plan: 2016-2020**  
Executive Office of the President of the United States

**American Artificial Intelligence Initiative: Year One Annual Report, 2020**  
Executive Office of the President of the United States

**Executive Order on AI, 2020**  
Executive Office of the President of the United States

## Industry Documentation

**The NIST Machine Learning & AI Initiative, 2016**  
National Institution of Standards and Technology

**U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards, 2016**  
National Institution of Standards and Technology

**A 20 Year Community Roadmap for Artificial Intelligence Research, 2019**  
Computing Community Consortium  
Association for the Advancement of Artificial Intelligence

**Open Community for Ethics in Autonomous and Intelligent Systems (OCEANIS), 2019**  
IEEE Ethics Standards Organization

## Information Tools

**United States AI Legislation Tracker, 2020**  
Center for Data Innovation