

Understanding Big Data & AI

Ivan Portilla 

Portilla@gmail.com

AI Technical Architect
7/12/23



Robo del siglo

<https://www.bbc.com/news/world-latin-america-28783027>

Entity Resolution

Entity Resolution at Degrees of Difficulty

<https://senzing.com>

Exactly
Same

Bob Jones
123455

Bob Jones
123455

Fuzzy

Bob Jones
123455

Robert T Jonnes
000123455

Incompatible
Features

Bob Jones
123455

Why are enterprises struggling to capture the value of AI?

Data

- Data resides in silos & is difficult to access
- Unstructured and external data isn't considered

Governance

- If the data isn't secure, self-service isn't a reality
- Understanding data lineage and getting to a system of truth is a challenge

Skills

- Data Science skills are in low supply and high demand
- Nurturing new data professionals is challenging

Tools & Infrastructure

- Need an environment that enables a “fail fast” approach
- Discrete tools present barriers to productivity

Agenda

Big Data & AI

AI Essentials Framework

Demo

Agenda

Big Data & AI

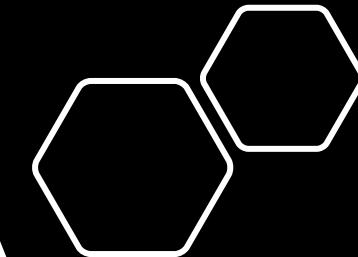
AI Essentials Framework

Demo

“Every day, we create 2.5 quintillion bytes of data”



Estimates are that less than 0.5% of data is ever analyzed!

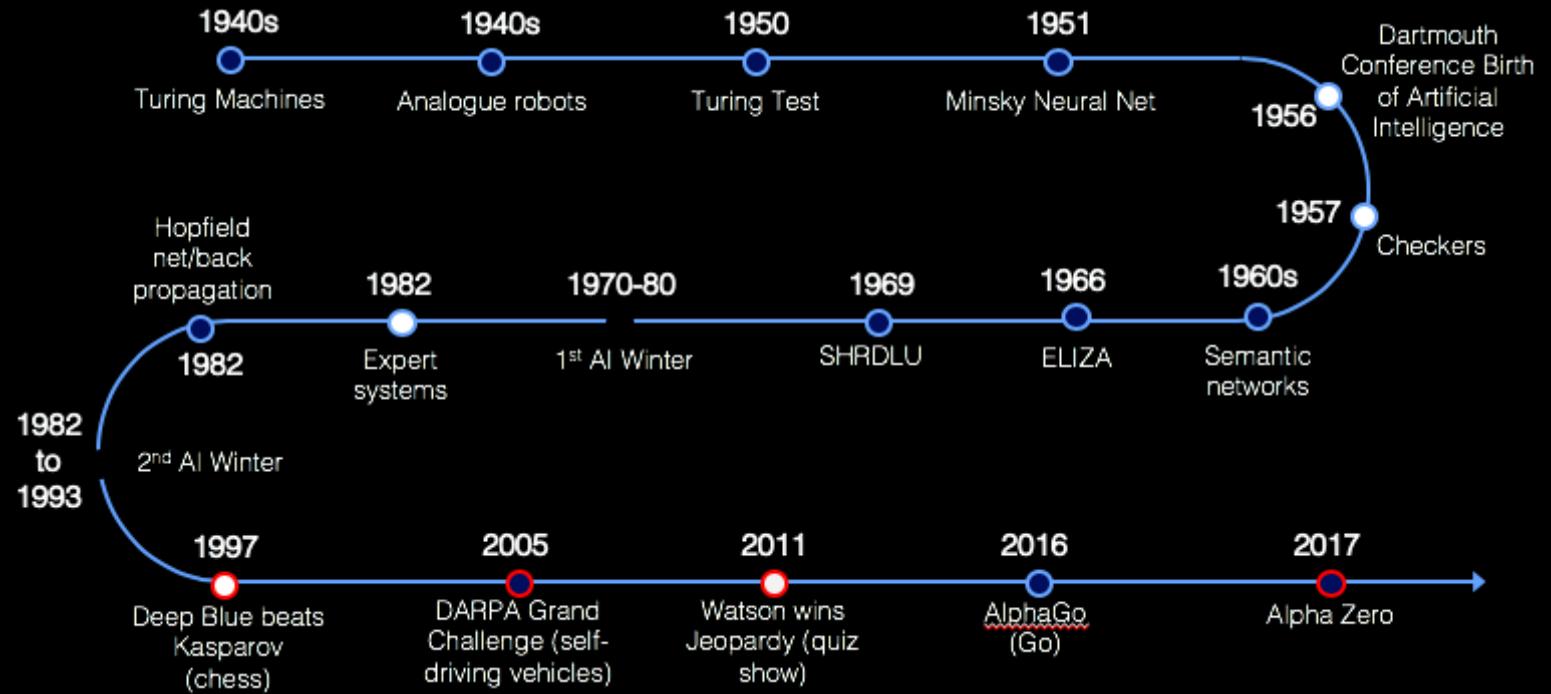


Understanding Big Data

Term	Factor	# RAMACs	# IPODs
1 Giga (GB)	10^9	200	
1 Tera (TB)	10^{12}	200K	200
1 Peta (PB)	10^{15}	200M	200K
1 Exa (EB)	10^{18}		200M
1 Zetta (ZB)	10^{21}		
?	10^{24}		
?	10^{100}		



A brief history of AI



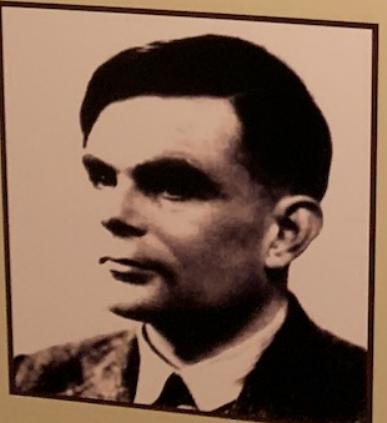
ALAN TURING'S OFFICE

Here at his desk in Hut 8, Turing took the lead on breaking naval Enigma ciphers – something few thought could ever be done. His mathematical skills also enabled him to break other ciphers, including the complex Lorenz cipher where he used a method that became known as Turingy. Together with his fellow Codebreaker Gordon Welchman, he developed the Bombe machine to help speed up the codebreaking process.

IN THEIR WORDS

If anyone was indispensable to Hut 8 it was Turing. The pioneer work always tends to be forgotten when experience and routine later make everything seem easy, and many of us in Hut 8 felt that the magnitude of Turing's contribution was never fully realised by the outside world.

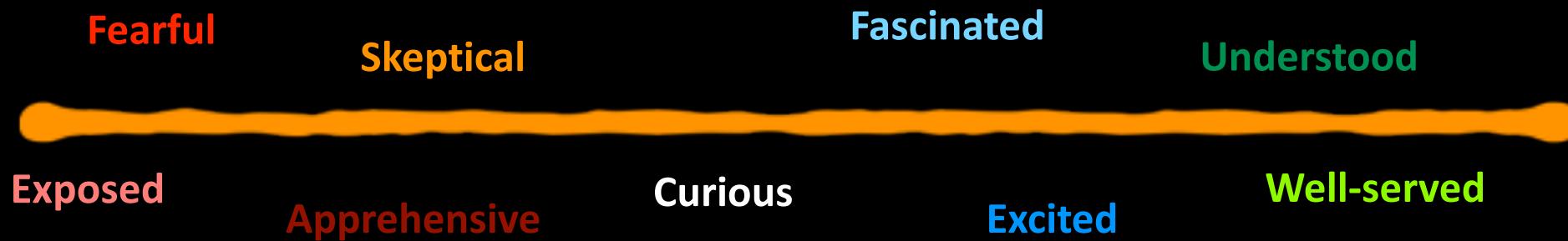
Hugh Alexander, Codebreaker, Hut 8



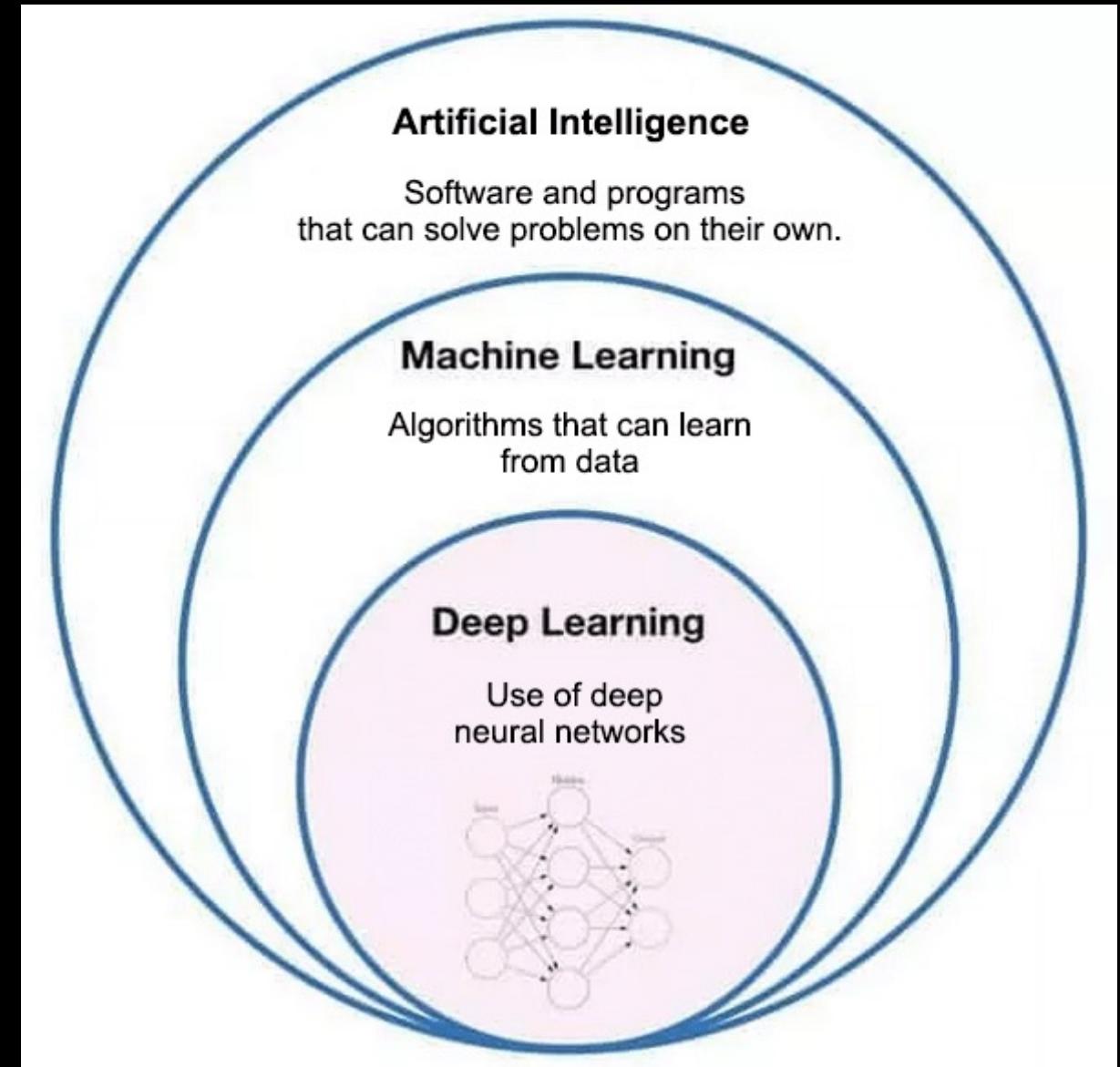
What is AI?

AI Perception

What do people feel about AI?



What we know



Why Machine Learning?

Traditional Programming



AI-Machine Learning



ML provides systems the ability to **automatically** learn from **experience**
(without being explicitly programmed)

Intelligence demonstrated
by CU Students

Exercise One

PB & J Recipe

Exercise 1

Program how to make a Peanut Butter & Jelly sandwich

1. Get ingredients (PB, Jelly, Bread)
2. Get cooking utensils (plate, knives)
3. Get two slices of bread
4. Open PB, Jelly jars
5. With a knife spread 2 Tablespoons of peanut butter on one piece of bread
6. With a different knife, spread 2 Tablespoons of jelly on the other slice of bread
7. Put the slices together
8. Toddler adaptation: cut off crusts before serving.
9. Place on plate & serve



Exercise Two

A photograph of a highway scene. In the foreground, the side of a white truck is visible, showing its trailer and some red reflective markings. On the left side of the road, there is a yellow diamond-shaped road sign with a black silhouette of a car on a winding road. The road has a solid yellow line on the left and dashed white lines in the center. In the distance, another vehicle is approaching. The background consists of a dense forest of green trees under a clear sky.

Passing truck in
highway

Passing truck in highway

Don't Follow Closely Behind Prior to Passing

- As you prepare to pass, keep a healthy distance between your car and the truck in front of you. At minimum, 30 feet should separate the two vehicles. Signal clearly to indicate your next move.

Pass in the Left Lane

- The left lane is referred to as the 'passing lane' for a reason; it's the safest place to pass any vehicle, but particularly trucks and buses. It's frustrating to be stuck behind a slow truck in the left lane, but resist the urge to pass on the right side.

Don't Linger in the Truck's Blind Spot

- Trucks have huge blind spots on both sides. Pass promptly while still abiding by the speed limit. Don't hang out in the blind spot area, which, in the left (passing) lane, constitutes the truck's entire front half.

Recognition

- With GPS navigation, HiRes cameras, Laser-radar cameras: Location, speed of trucks, road conditions, emergency lane
- With Image Recognition: Turn signal (!lamp), paper or debris on the road

Insights

- Data is fed into multiple CPUs + collected data + rules
- Manufacturer knows who has the right of the way, when trucks can pass + MM miles driven have feedback on different driving style

Action (Cognitive Decision)

- Assist (alerts driver)
- Act (Pass or Not pass, Measure decision effects, feedback loop, and share w other cars)

Exercise Three

How is AI being used today?

AI is Everywhere

AUTOMATION

CONVERSATION

OPTIMIZATION

RECOMMENDING
AND INSIGHTS

PERSONALIZING

VISUAL RECOGNITION



Automating actions, processes, and alerts

Understanding text and answering questions in natural language

Analyzing data for patterns and outliers and deriving insights

Understanding needs and recommending solutions

Personalizing experiences

Recognizing objects in images and videos

What Machine Learning Can Do

A simple way to think about supervised learning.

INPUT A	RESPONSE B	APPLICATION
Picture	Are there human faces? (0 or 1)	Photo tagging
Loan application	Will they repay the loan? (0 or 1)	Loan approvals
Ad plus user information	Will user click on ad? (0 or 1)	Targeted online ads
Audio clip	Transcript of audio clip	Speech recognition
English sentence	French sentence	Language translation
Sensors from hard disk, plane engine, etc.	Is it about to fail?	Preventive maintenance
Car camera and other sensors	Position of other cars	Self-driving cars

SOURCE ANDREW NG

© HBR.ORG

Future of AI Example (Semantic Analysis)

The Wining Shot

The soccer game was nearly over. The two teams were tied, one to one.

Alice kicked the ball. Oops! She kicked it the wrong way.

"What a bad kick," Alice thought.

But wait!. Her teammate John jumped to stop the ball. The ball went off John into the goal!.

Alice's team won the game. Hooray! It was not a bad kick, after all.

Who won the game?

Why?

Future of AI Example

The Wining Shot

The soccer game was nearly over. **The two teams were tied, one to one.**

Alice kicked the ball. Oops! She kicked it the wrong way.

"What a bad kick," Alice thought.

But wait!. Her teammate John jumped to stop the ball. **The ball went off John into the goal!!.**

Alice's team won the game. Hooray! It was not a bad kick, after all.

Who won the game?

Answer: Alice's team
Correct

Explanation: Alice's team won the game because Alice's team had more points than the other team

✓ **Correct**

ChatGPT

JE

read this: The Wining Shot

The soccer game was nearly over. The two teams were tied, one to one.

Alice kicked the ball. Oops! She kicked it the wrong way.

"What a bad kick," Alice thought.

But wait!. Her teammate John jumped to stop the ball. The ball went off John into the goal!.

Alice's team won the game. Hooray! It was not a bad kick, after all.

ChatGPT

JE who won the game

 Alice's team won the game, even though Alice initially thought she had made a bad kick. John's quick thinking and intervention helped the team to score the winning goal, and they were able to secure victory in the game.

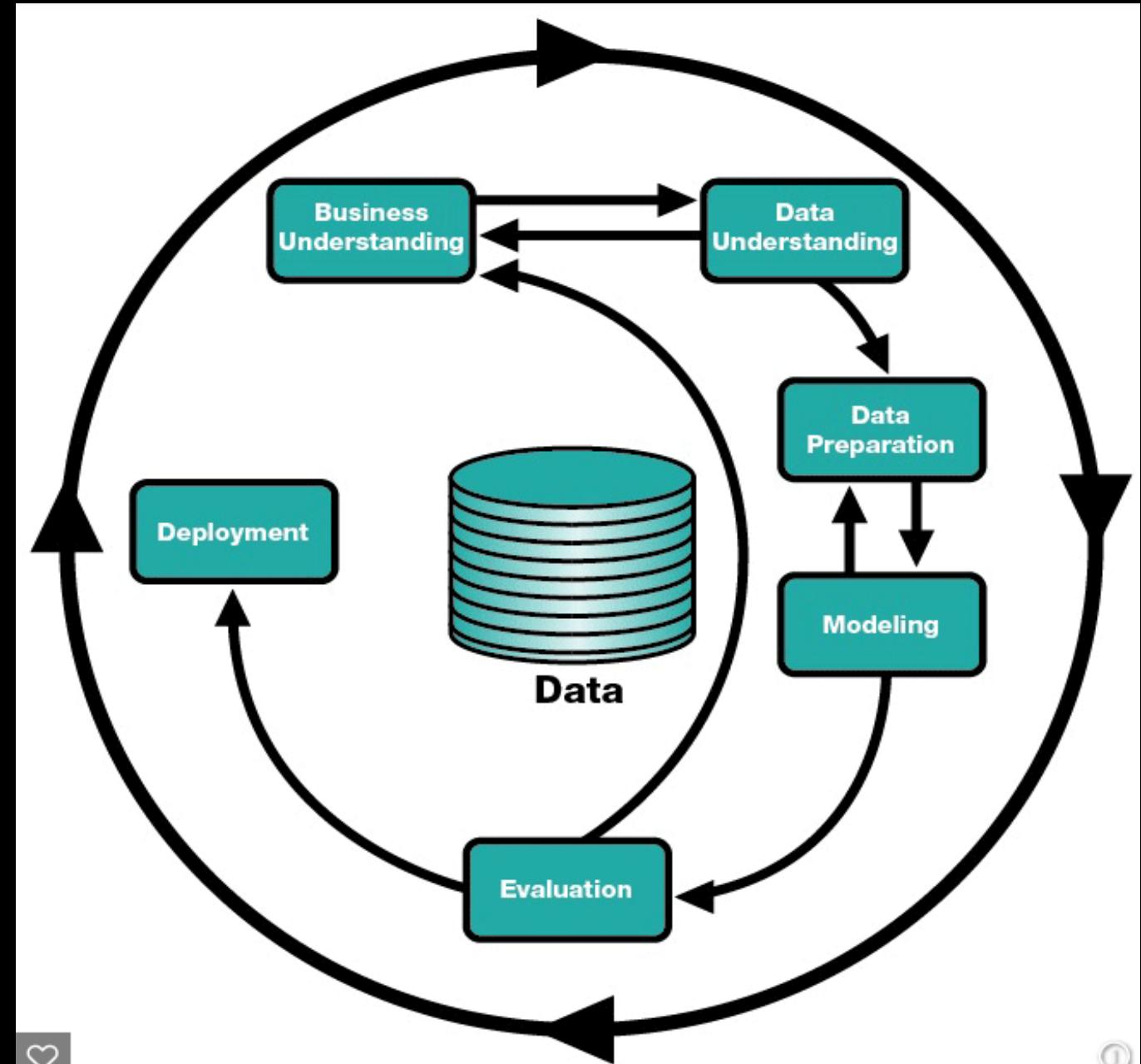
JE why?

 Alice's team won the game because John was able to use his skills and reflexes to intercept the ball, even though Alice's kick was not aimed in the right direction. John's intervention caused the ball to go into the goal, which resulted in a score for Alice's team. Since this happened near the end of the game and there was no time left for the other team to catch up, Alice's team won the game with a score of 2-1.

CRISP-DM

Solution Development Method Approach

CrossIndustry Standard Process for Data Mining (CRISP-DM)



AI Project Roles

Drives governance
policy effectiveness
Data Steward

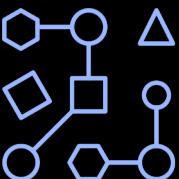
Builds data pipelines that power
dashboards

Data Engineer



Prepares data to tease out
the insights the biz is looking

Data Scientist



Business Analyst

Works with data to apply insights
to the business strategy



App Developer

Makes insights immediately
actionable



CRISP-DM Exercise

Business Understanding

Business Understanding

- ✓ **What** the business expects to gain from the data analysis?
- ✓ **how** the results of the analysis will be used?

The focus is on the Business Problem, not the solution (underlying technology).

The focus needs to stay business

Task List

1. As Is Scenario & pain points: Gather background information about the current business situation
2. To Be Scenario: Document specific business objectives decided upon by key decision makers
3. Agree upon success criteria: How can we measure success / sign off

Defining Business Objectives

- Describe the problem to be solved
- Specify all business questions as precisely as possible
- Determine any other business requirements (not losing an existing customer while increasing cross sell opportunities)
- Specify expected benefits in business terms (reducing churn among high-value customers by 10%)

Business Understanding

- **Business Success Criteria**
- When can you say the task is complete?
- There are two types of success criteria:
- **Objective** – The criteria quantifiable and can be measured
- **Subjective** – The criteria is less concrete and harder to measure (“Discover clusters of effective treatments”)
- Did you connect each business objective to a success criteria?

Business Understanding

Case Study

- Maria, the VP of marketing at Macy's, is feeling the pressure from growing competition from web retailers (Amazon, eBay)
- Customer acquisition is **expensive**, she has decided that she needs to **cultivate** existing customer relationships in order to **maximise** the value of the company's **current** customers

She commissions a study with the following objectives

- Improve **cross-sales** by making **better recommendations**
- Increase **customer loyalty** with a more **personalised service**

She identifies the following success criteria

- Cross-sales increase by **10%**
- Customers **Spend** more time (+10 minutes) and **see** more pages (+5 pages) on the site per visit.

Business Understanding

- **Case Study**
- Maria tasks her business analyst Jean to undertake the study.
- Jean **expands** the business objectives by exploring the data that the company has available
- Use historical information about previous purchases to generate a model that links "related" items. When users look at an item description, provide links to other items in the related group (**market basket analysis**).
- Use Web logs to determine what different customers are trying to find, and then redesign the site to highlight these items. Each different customer "type" will see a different main page for the site (**profiling**).
- Use Web logs to try to predict where a person is going next, given where he or she came from and has been on your site (**sequence analysis**).

Business Understanding

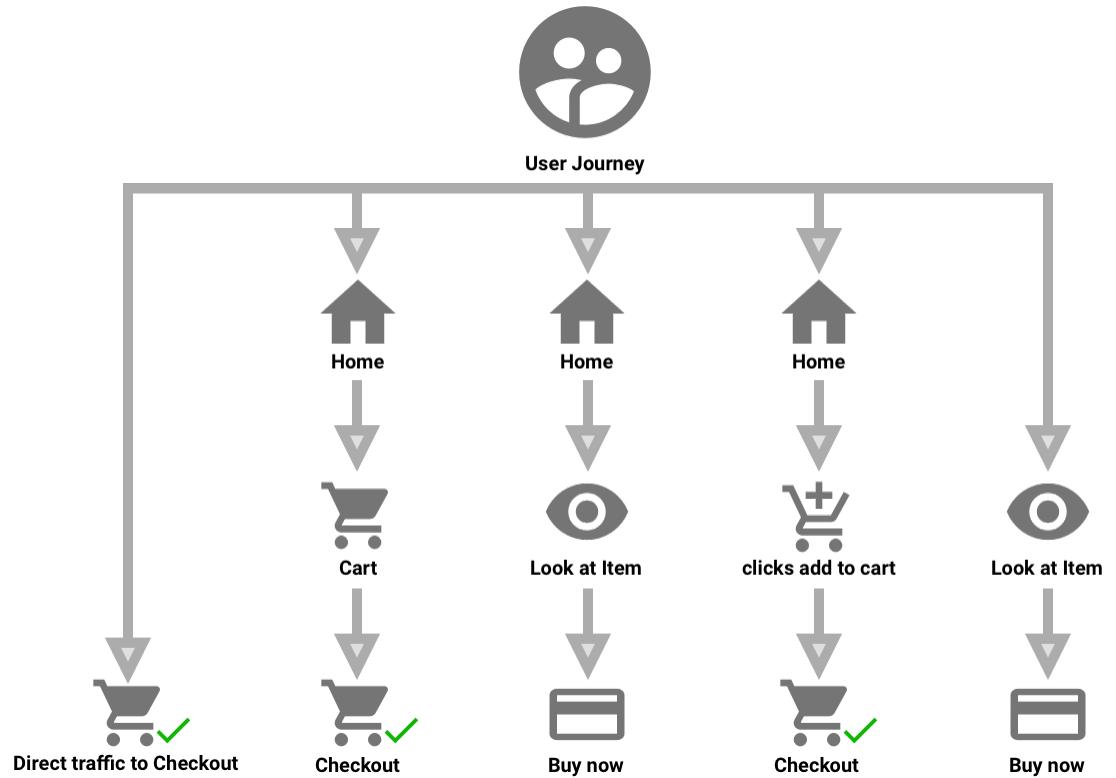
- **As-is scenario**

- Customer

1. Clicks on sales items page
2. Selects item, add it to cart
3. Leaves site

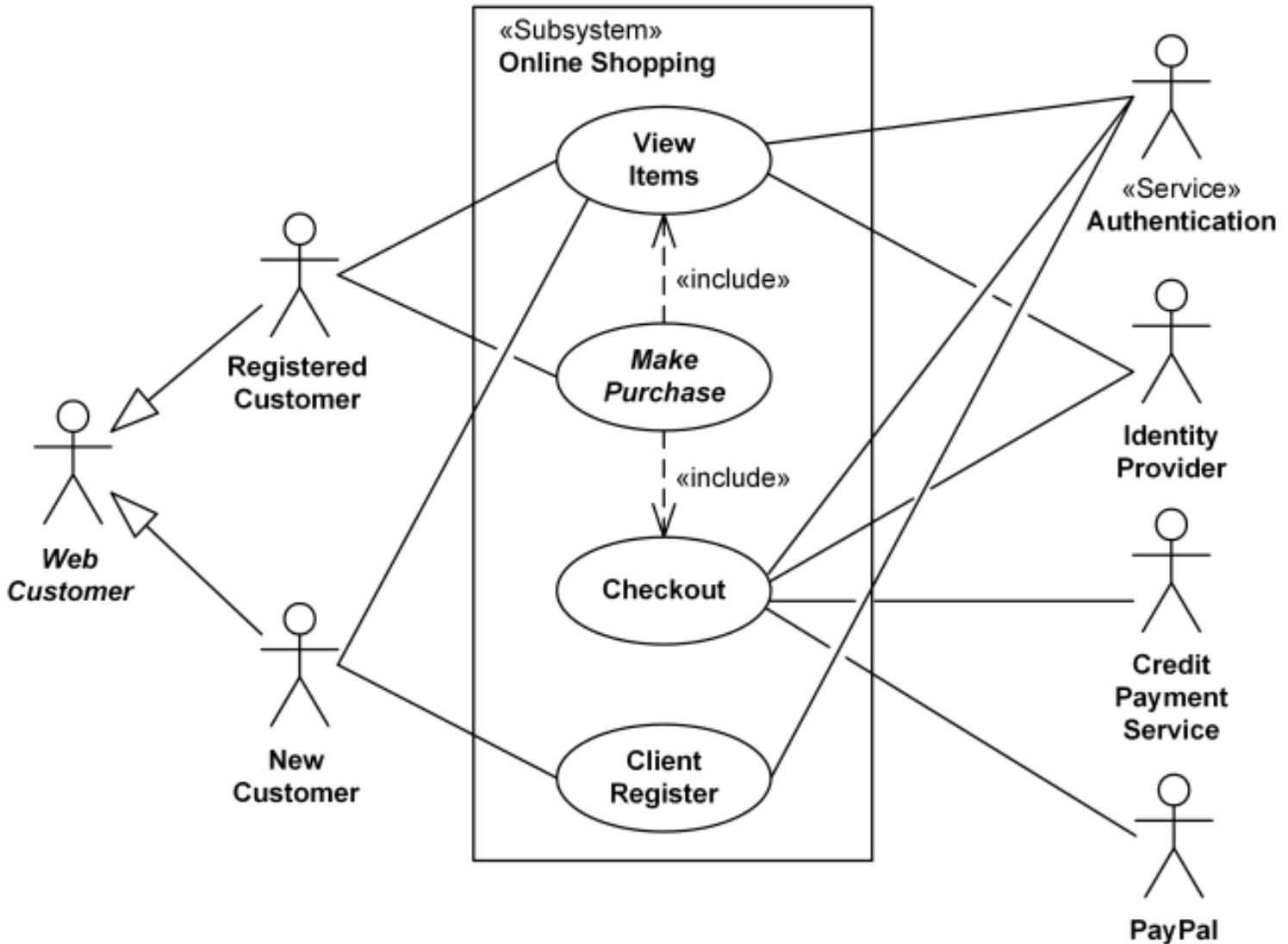
- **Pain points**

- Short web visits
- Large number of abandoned shopping carts



Business Understanding

- To-be scenario
- Customer
 - 1. Customer logins
 - 2. Clicks on sales items page
 - 3. Sees personalized offers
 - 4. Selects item, adds it to cart
 - 5. Completes purchase



Exercise

- **Case Study - A**
- Karan, the VP of marketing at Netflix, is feeling the pressure from growing competition from stream companies (Hulu, Disney+)
- Customer acquisition is **expensive**, he has decided that he needs to **cultivate** existing customer relationships in order to **maximise** the value of the Netflix **current subscribers**
- **Case Study - B**
- Kavitha, the VP of marketing at Twitter, is feeling the pressure from growing competition from social companies (Instagram, Threads)
- Customer acquisition is **expensive**, she has decided that she needs to **cultivate** existing customer relationships in order to **maximise** the value of the Netflix **current users**

Agenda

Big Data & AI

AI Essentials Framework

Demo

AI Essentials Framework Intro

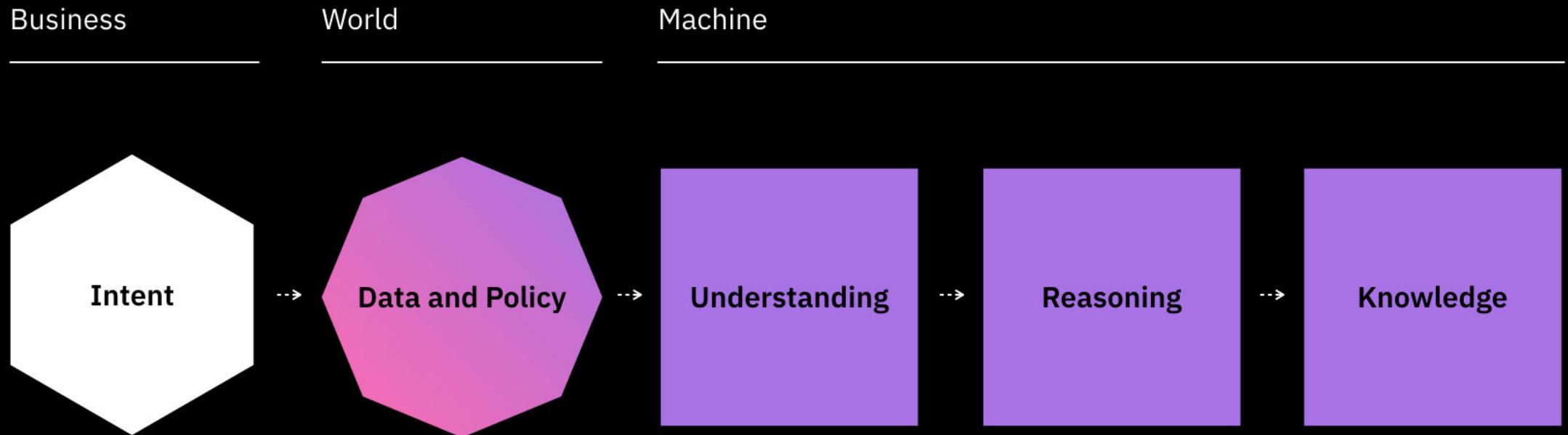
Adam video

Artificial
Intelligence
Design:

The purpose planning or
intention behind simulated
human thought process

AI Essentials Framework

https://www.ibm.com/design/thinking/page/courses/AI_Essentials/topic/Introduction/01/01/01/7



AI Essentials Framework

What are you trying to accomplish?

When in doubt, always design for people first and what they value.

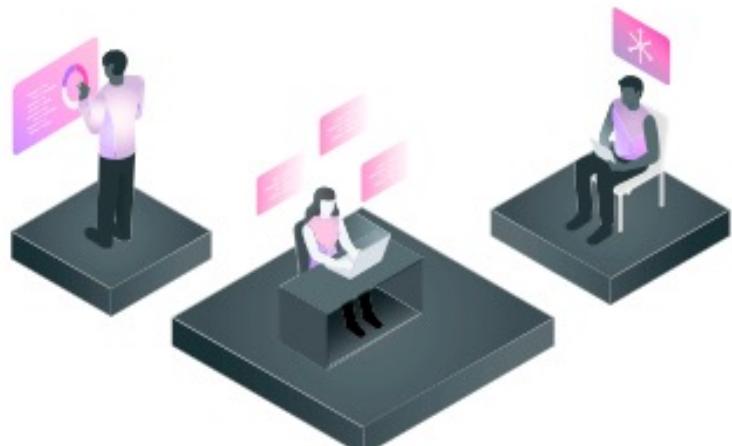
You're here because you have some level of interest in designing an AI experience.

artificial intelligence design

the purpose, planning, or intent behind simulated human thought processes

Why do we use AI?

To achieve higher quality outcomes faster than humanly possible



Add a sticky note

Describe the project you want to pursue.

Let's get started

Why do you think AI makes sense for this project?

Think about whom you need to lean on in order to design an AI experience. These people could be peers, stakeholders, executives, users, subject matter experts, etc.

Add a sticky note

Name:

Job role:

How could this person help?

Add a sticky note

Are any roles missing on your team?

What other expertise do you need? Fill this in as you continue through the AI Essentials Framework.

AI Essentials Framework

AI Essentials Framework

Assumptions and Questions

What research will you need to address the assumptions and questions?

Print more as needed

Add a sticky note

Place an assumption or question which you're uncertain about and is high-risk to your project here.

How will you validate your assumption or uncover an answer to your question?

How will you validate your assumption or uncover an answer to your question?

How will you validate your assumption or uncover an answer to your question?

Empathy Map

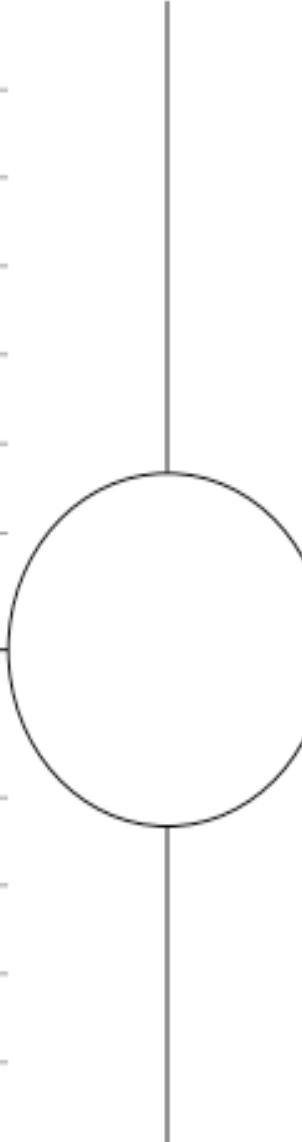
Empathy Maps help synthesize your team's collective knowledge about your user persona, bringing you closer to a common understanding of who they are. Reference the Enterprise Design Thinking Toolkit to learn how to run through the Empathy Map activity.

Says

Thinks

Does

Feels



AI Essentials Framework

As-is pain points

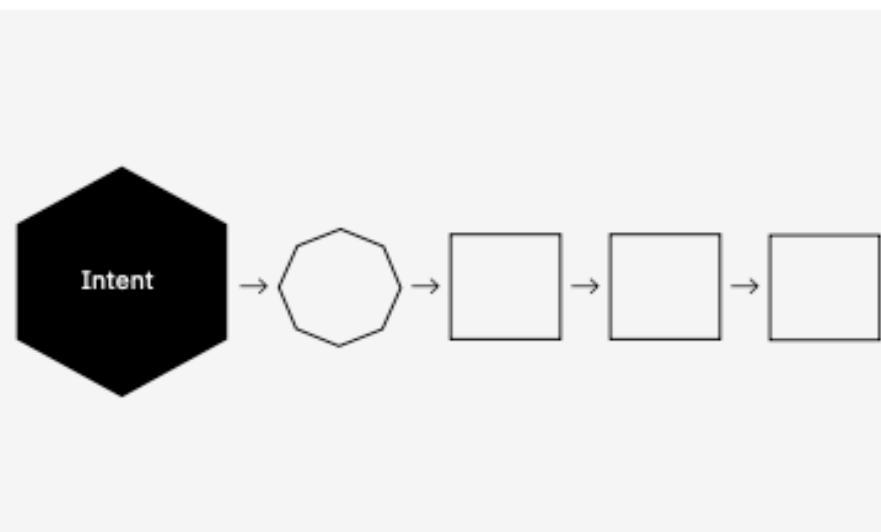
Reference the Enterprise Design Thinking Toolkit to learn how to develop an As-is Scenario Map. As you conduct research and gain insight into your users' experiences, document common pain points you uncover here.

Add a sticky note

Write down a pain point in your users' experience.

Define your intent

Align your team on the intersection between your business and user intents. This will help you determine why you would or wouldn't use AI in your effort.



What are you trying to accomplish for your users? Think about both the needs of your business and the needs of your users. What does success for both look like?

6 core AI intents

With AI, we enable our users to:

Accelerate research and discovery

Conduct rigorous, domain-specific research faster by using machine learning and AI to comb through your data and extract the information you find most important.



Enrich your interactions

Understand and communicate with customers and employees using natural language, responding to their needs with tailored dialogue and personalized experiences.



Anticipate and preempt disruptions

Monitor your systems and equipment at all times to identify and address potential issues before they become larger, more expensive problems.



Recommend with confidence

Make more confident, targeted recommendations using AI to evaluate a broad set of information based on an understanding of the parameters that are important to you.



Scale expertise and learning

Collect know-how from experts and combine it with the latest information from your industry to create a deep source of tribal knowledge that all employees can access on-demand.



Detect liabilities and mitigate risk

Use AI's understanding of the written word to identify risks to your company, particularly in terms of regulatory compliance.

What are you trying to accomplish for your users?

Accelerated Research & Discovery

Less time looking for information, more time acting on it

Enrich Interactions

Train AI to reduce response time, more transaction, more productive

Recommend w/ Confidence

Teach AI thoroughly, make better decisions, give tailored advice, better client relations

Scale Expertise & Learning

Offer deep instructional knowledge to everyone in the organization

Anticipate & Preempt Disruptions

Continuously monitor, mitigate problems before disruption, identify issues

Detect Liabilities & Mitigate Risk

Identify compliance issues quickly, protect business & people

Align on your intent

•AI

Essentials
Framework

Add a sticky note

Write down the core AI intent you chose to help guide the AI intent for your project.

Add a sticky note

Write down the AI intent you came up with during the intent activity.

The intent activity will get you from a core AI intent to an AI intent you can move forward with.

Add a sticky note

Add another core AI intent if it makes sense for your project.

How would your users benefit from the intent above?

How would your business benefit from the intent above?

Come up with big ideas

Reference the Enterprise Design Thinking Toolkit to learn how to come up with big ideas. Record a few of your ideas here. Then, document the idea you want to move forward with on the next page.

Add a sticky note

Title your big idea and provide a written description.

•AI

Essentials
Framework

Use a thick permanent marker to illustrate your idea on a second sticky note.

Your big idea

Place your best idea here. You will refine this idea later.

•AI

Essentials
Framework

Add a sticky note

Write down the title and description of your best idea here.

Add a sticky note

Place an illustration of your big idea here.

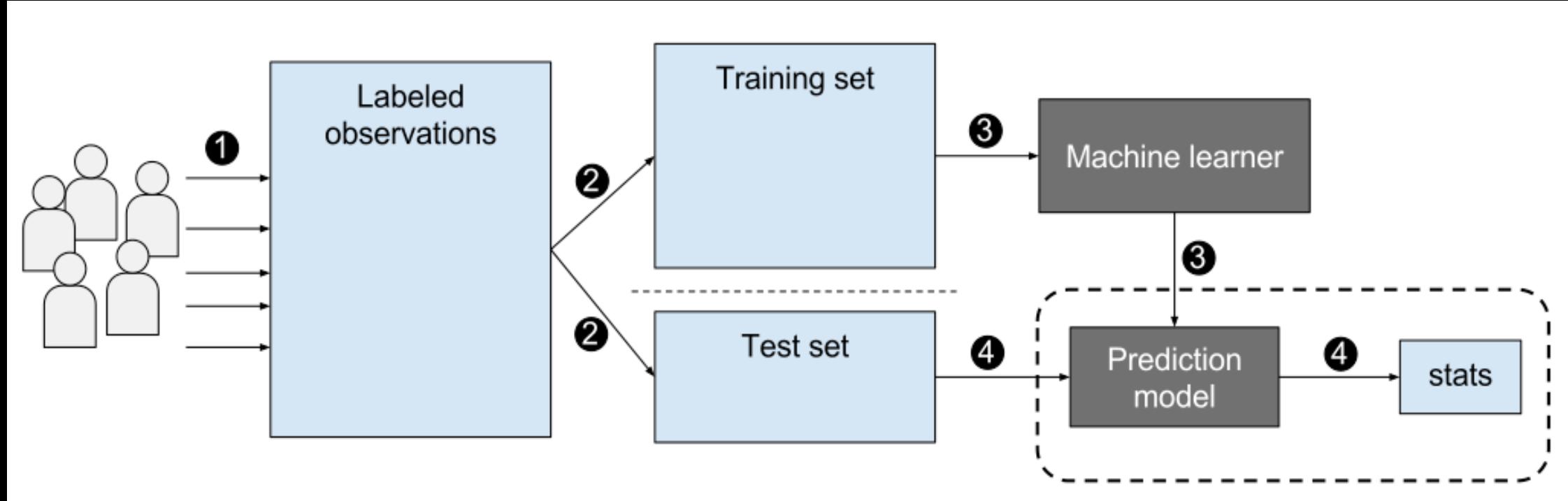
Agenda

Big Data & AI

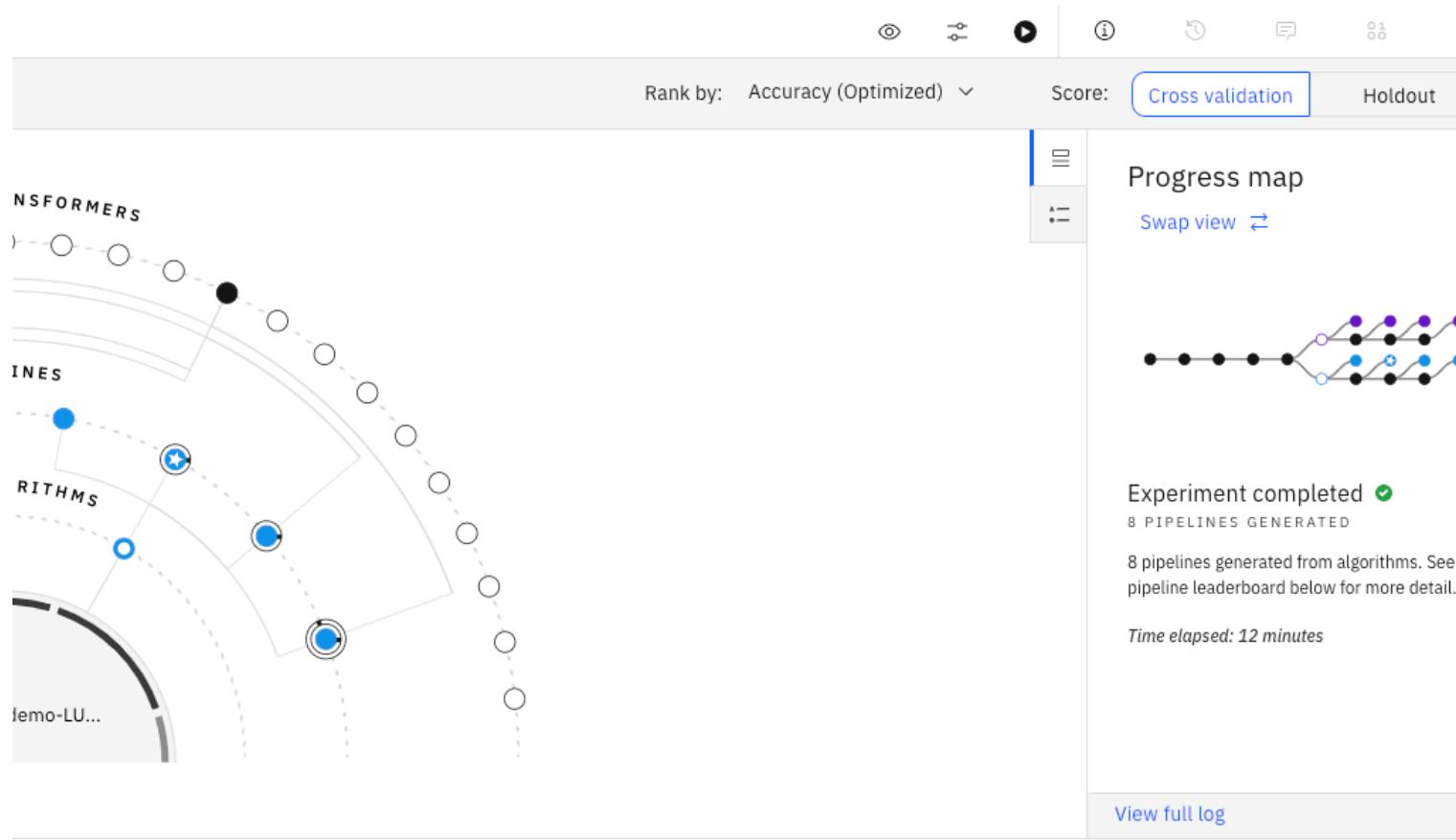
AI Essentials Framework

Demo

Supervised Machine Learning



Supervised machine learning algorithms can apply what has been learned in the past to new data using labeled examples to predict future events



Accuracy (Optimized)	Enhancements	Build time	Save as
0.975	HPO-1	00:01:00	Save as
0.974	HPO-1 FE	00:05:31	

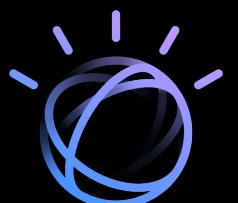
Supervised Machine Learning Demo



Closing Thoughts

Industry Principles for Trust and Transparency

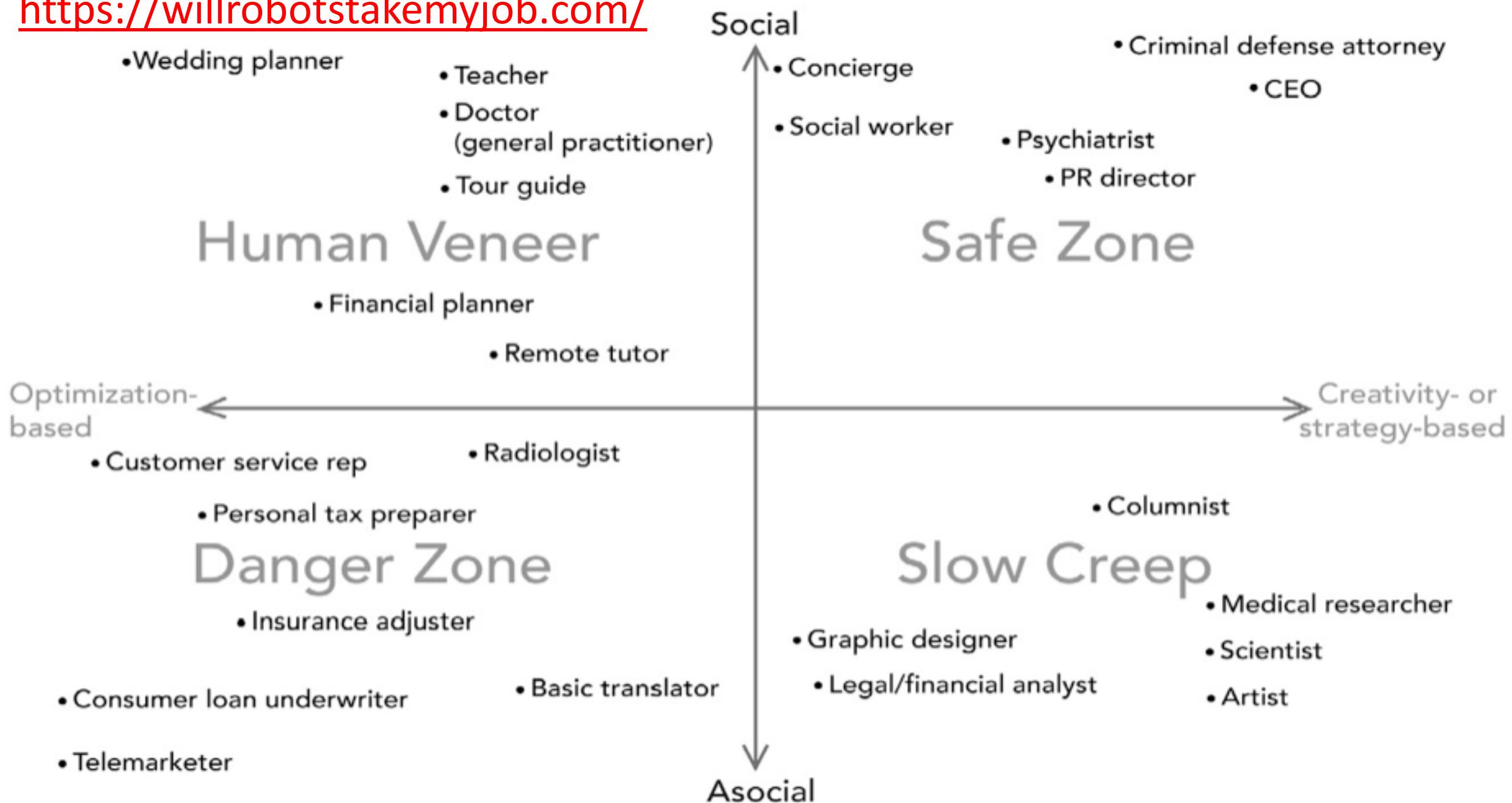
<https://www.ibm.com/downloads/cas/J45XZOAR>

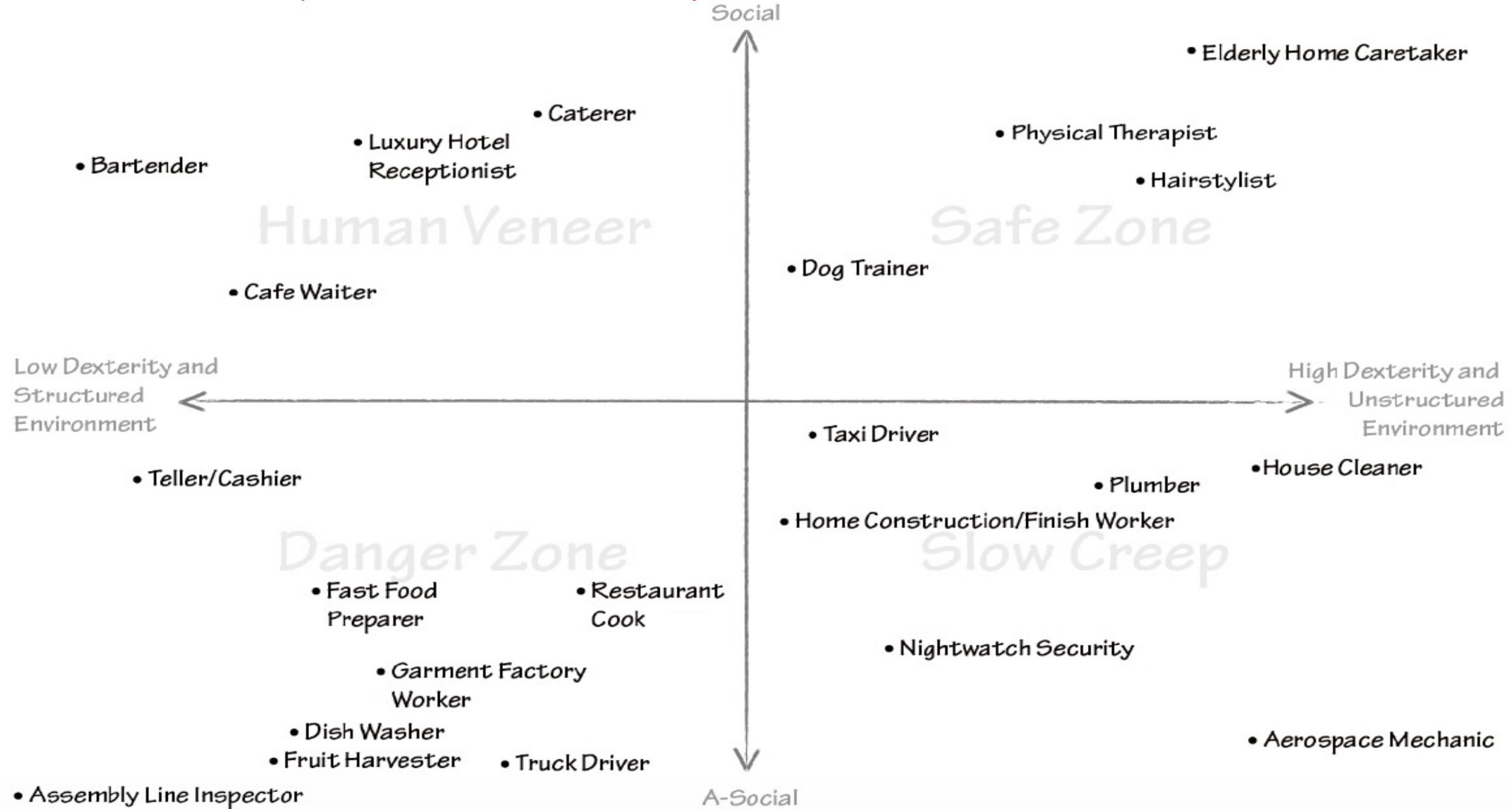


1.The purpose of AI is to augment
human intelligence

2.Data and insights belong to
their creator

3.AI systems must be transparent
and explainable





Takeaways AI lecture

Citizen AI:

Develop critical thinking skills, have your voice heard

<https://www.accenture.com/cz-en/insight-explainable-citizen-ai>

Big Data Advantage:

When giving away your personal data, choose wisely

<https://www.slideshare.net/RobertoVII/ai-and-big-data-for-business-and-people-advantage>

Intelligent Automation:

Educate yourself on AI, thrive in your profession

https://1.dam.s81c.com/m/3de136737e51fb20/original/IBM-Automation-Whitepaper_Final.pdf

Q&A

Ivan Portilla

Portilla@gmail.com

