Week 11 - Part 1: Intro to Reinforcement Learning

- What is Reinforcement Learning (RL)?
  - Learning through experience/data to make good decisions under uncertainty
- Decision making is an essential part of inteligence

* This part includes some slides from Emma Brunskill

Week 11 - Part 1: Intro to Reinforcement Learning

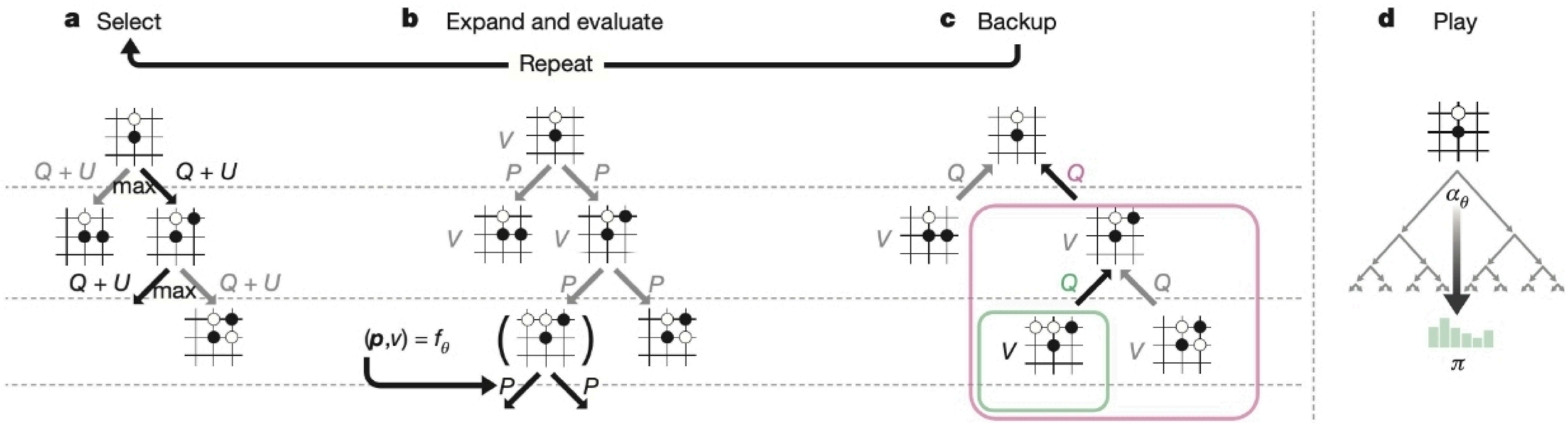
- So far in the course, we have studied techniques to identify things
- Also in real life, we saw a lot of progress on what is called "Perceptual machine learning", e.g., to perceive faces, cats and dogs, ...
  - e.g., to perceive faces, cats and dogs, digits, ...
  - perceptual machine learning tries to identify something.
- In reality, what we are trying to do is to make decision based on our perception/information we receive.
  - So, it's critical to think about how to make "good" decisions, when it comes to intelligence.

How to Make Good Decision from limited Experience/Data

- These sort of questions, particularly when faced by uncertainty, had been studied in depth at least since 1950.
- RL builds strongly from theory and ideas starting in the 1950s with Richard Bellman
- So, why should we study RL?
  - Because, understanding how to make good decisions from limited experience when faced by uncertainty is essential for any (artificial) intelligent entity
  - Also, because it's cool. It's practical.

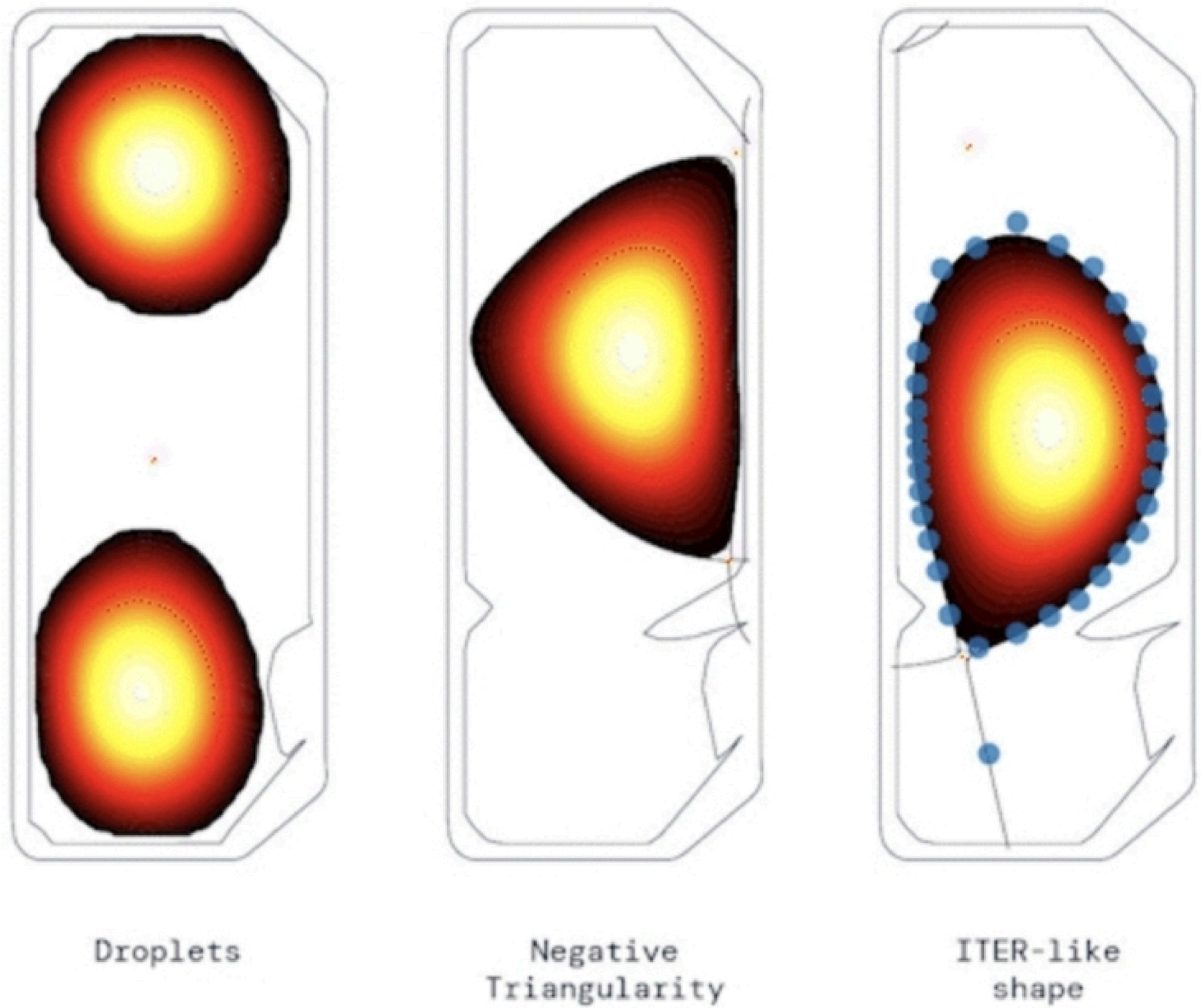
Some impressive successes in the last decade.

- Board game Go.
  - An extremely hard board game.
  - In 2016, a team called "deepmind", by combining RL and Monte_carlo_tree_search, they built an agent that could defeat the world champion.



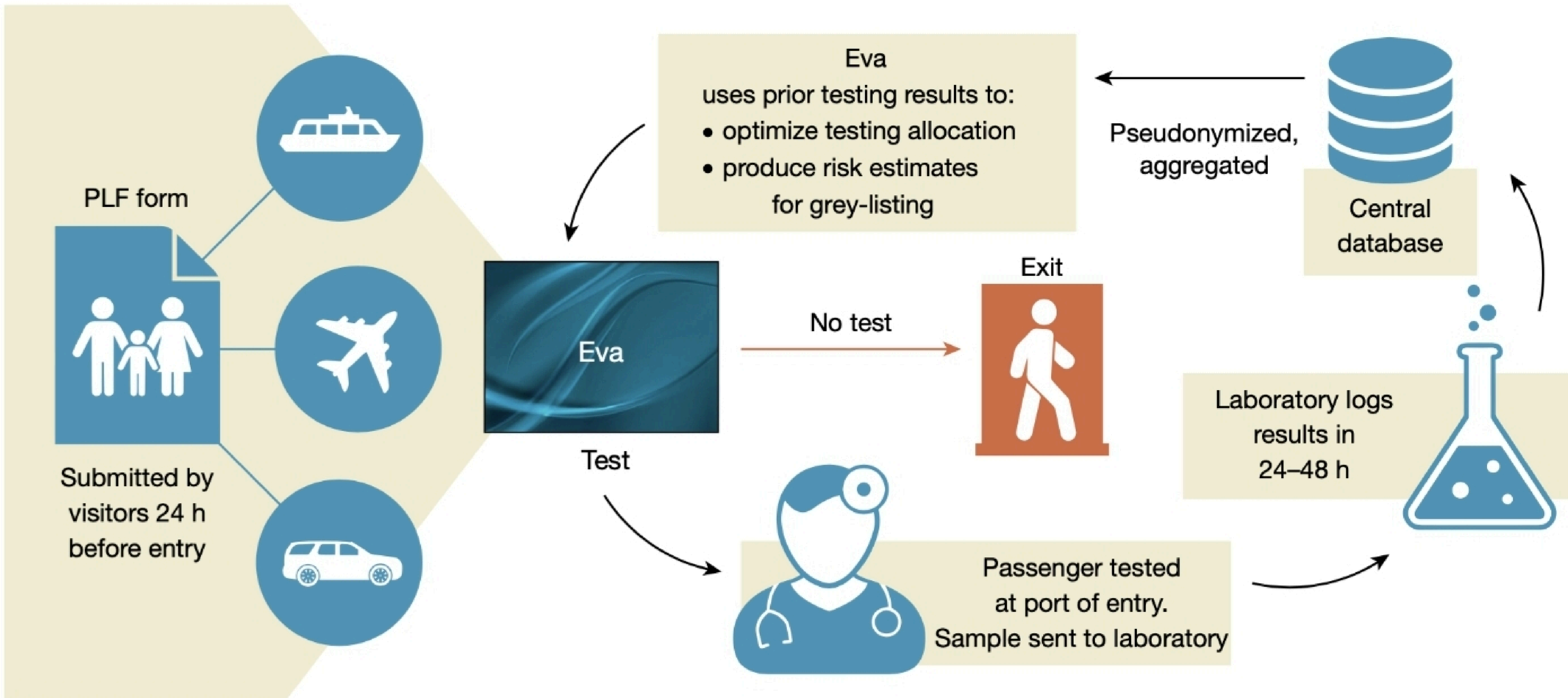
Some impressive successes in the last decade.

- Plasma Control for Fusion Science



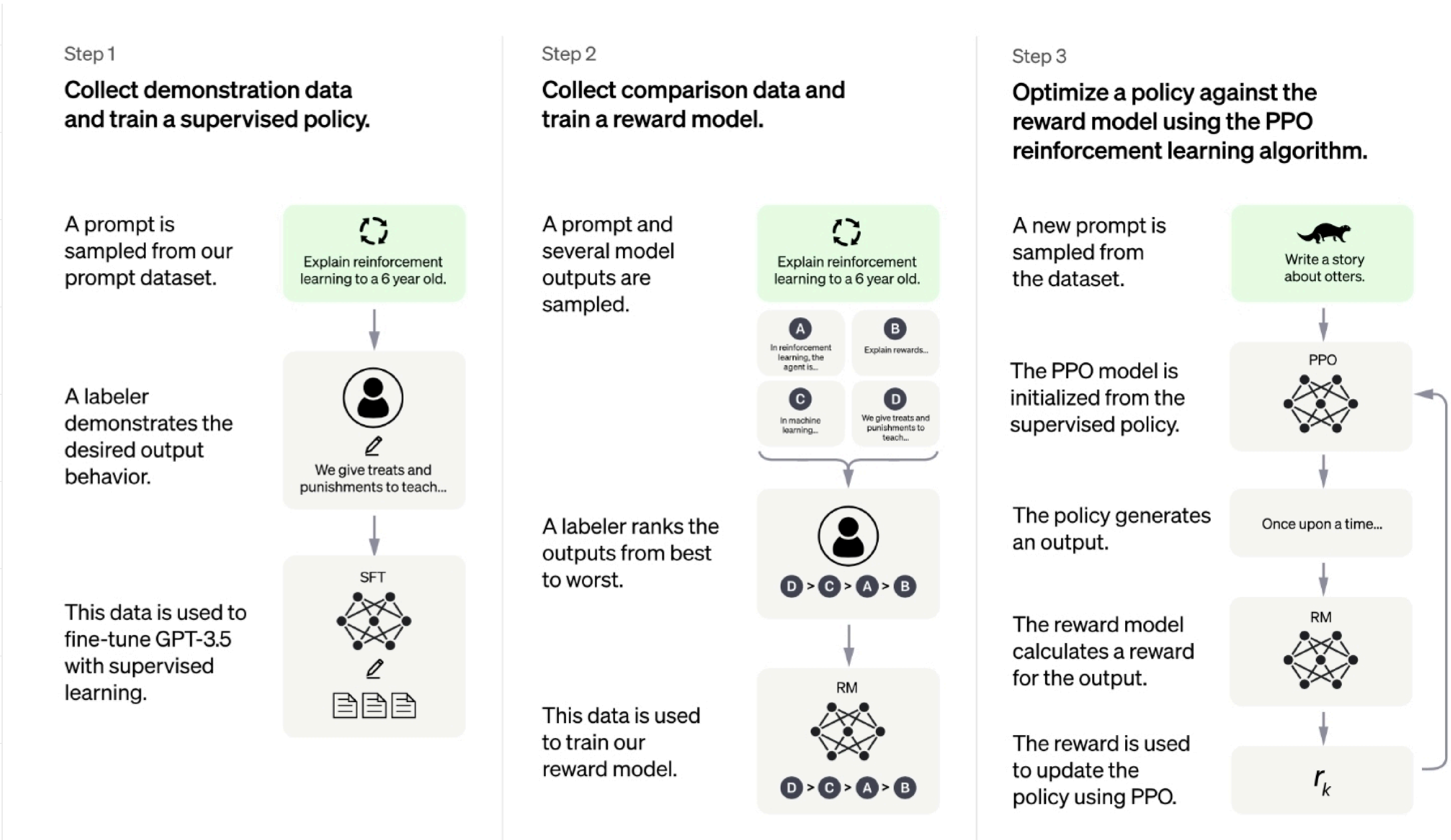
Some impressive successes in the last decade.

- Efficient and targeted Covid-19 testing
  - Bastani et al. Nature 2021, "Efficient and targeted Covid-19 border testing via reinforcement learning"



Some impressive successes in the last decade.

- ChatGPT



RL Generally Involves . . .

- Optimization
- Delayed Consequences
- Exploration
- Generalization

Optimization

- To find the "best" way to make decision
  - Our decisions must yield to best outcomes or at least very good outcomes
- Best outcomes? How can we specify if an outcome is the best outcome?
  - How to compare outcomes?
    - We do it with an explicit notion of decision utility

Delayed Consequences

- Decisions now can impact things much later
  - ↳ e.g., Consider saving for retirement
- Delayed Consequences introduces two challenges
  - ↳ When planning:

    Even when we know how the world works, decisions involves reasoning about not just immediate benefit of a decision, but also its long-term consequences.

  - ↳ When learning:

Delayed Consequences

- Decisions now can impact things much later
  - ↳ e.g., Consider saving for retirement
- Delayed Consequences introduces two challenges
  - ↳ When planning
  - ↳ When learning:

  When learning, we don't know how the world works.
  We want to learn it through direct world experience.
  But, temporal credit assignment is hard.

  - You take some action now, and later on you receive a good/bad outcome. How do you figure out which of your actions caused that good or bad later result.

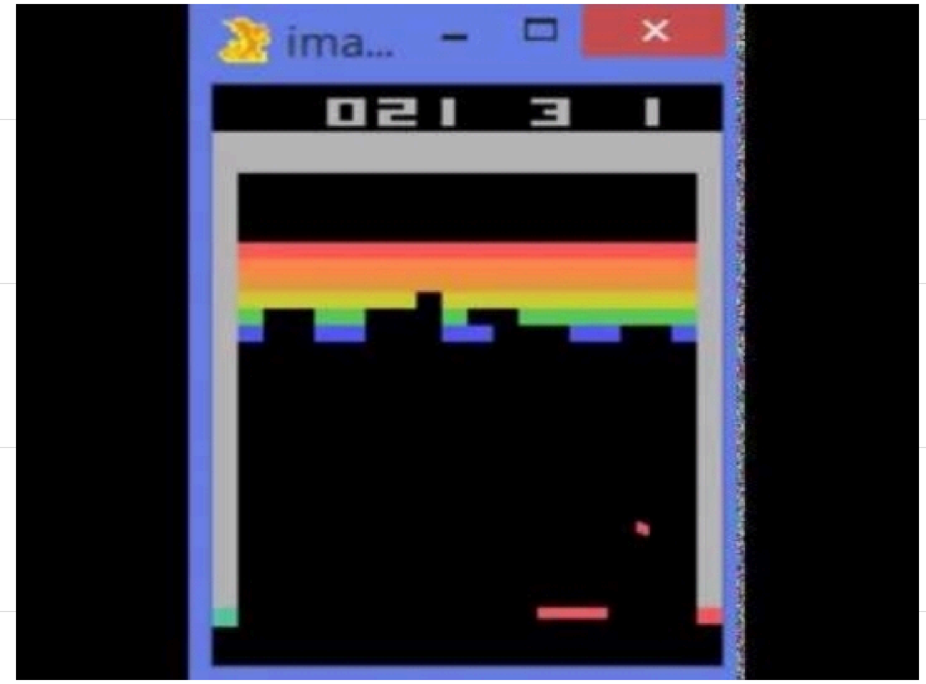
Exploration

- We learn from direct experience from interacting with environment
- you only learn about what you try out.
  - Don't know what would have happened for other decisions.

- That's why it's important to sometimes explore alternative actions cause it may give you valuable information.

Generalization

- Good decisions are learnt from past experience.
  - ↳ We need a mapping from possible states to decisions
- Why not just preprogram a decision Policy / mapping?
  - ↳ Because the number of possible states of the environment can be huge.
  - ↳ From a small set of states that we have seen we must learn a mapping that generalizes well to the states that we have not seen.



Atari Game

RL vs. AI planning vs. (Un)Supervised Learning

	Supervised Learning	Unsupervised Learning	AI Planning	RL
Optimization (over actions)			✓	✓
Learns from experience/data	✓	✓		✓
Generalization	✓	✓		✓
Delayed Consequences			✓	✓
Exploration				✓