

---

# What is Autonomy

## Part 1 of 2

Kaan Eren, James, Olritch



# Autonomy as Defined by ChatGPT

“Autonomy in AI ethics means protecting your right to make informed, voluntary decisions without manipulation by automated systems. AI should support your choices, not steer or coerce you through hidden algorithms, biased outputs, or opaque data use. Respecting autonomy requires transparency, consent, and meaningful human control.”



# Philosophical Autonomy

Personal freedom, self determination, or agency  
aka being able to give rule to yourself

- Not being able to give rules for yourself is not freedom
- Rules imposed by others can reduce user autonomy
- The rules can be made to a point that it doesn't hurt another person's freedom



# Physical Autonomy

Not being physically strained from completing an action or task

- Tied to a chair
- Locked in a room
- Chained to a wall



# Mental Autonomy

The minimal level of accurate and true information about one's existence and surroundings

- Cannot be free without knowledge
- Lack of knowledge leads to manipulation of your knowledge and informational freedom



# Psychological Reasoning

Proper functioning of one's mind

- A lack of mind function diminishes freedom
- Drugs and alcohol reduce mental functionality
- Irrational choices like eating 40 pounds of bananas



# Is running over someone being drunk considered murder?



**No**

A person has a lack of a properly functioning mind (psychological reasoning) while drunk, creating irrational decisions without autonomy



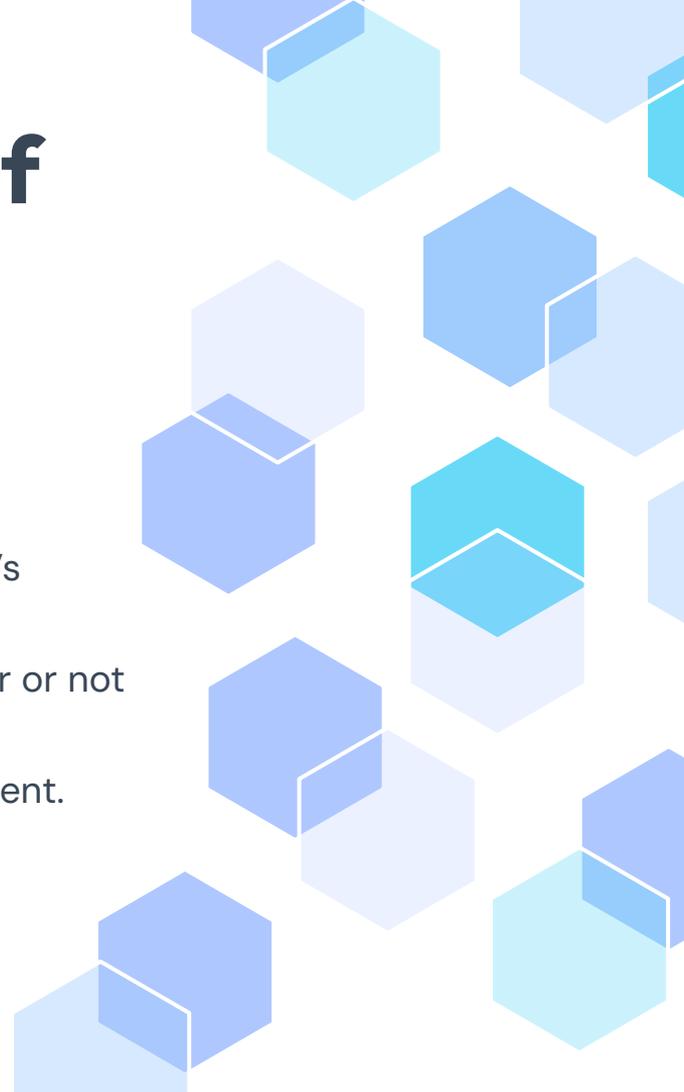
**Yes**

The person had control whether or not to consume that alcohol and that decision or rule they made hurt another person's freedom (philosophical)

# Example of Dimensions of Autonomy

Five dimensions of autonomy in health and fitness apps

- Degree of physical control within the app
- Degree of customization
- Degree of reliability of information and how it affects the user's decisions.
- Degree of the user understanding the app's goals and whether or not the app hinders/promotes this understanding
- Whether or not the app promotes some form of moral statement.



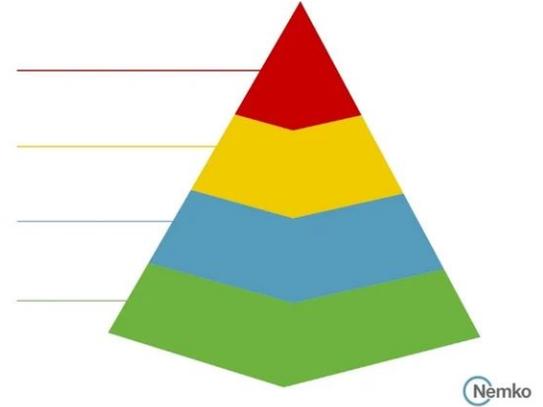
# EU AI Act

The EU AI Act divides AI system into three different levels

- Unacceptable Risk (Social scoring live with FR)
- High/Limited Risk (medical devices, cars/Chatbots)
- Minimal Risk (Recommenders like Netflix)

## EU AI Act: Risk Levels

- **PROHIBITED AI SYSTEMS**  
Prohibited
- **HIGH RISK AI SYSTEMS**  
Must undergo a conformity assessment
- **LOW RISK AI SYSTEMS**  
Must adhere to transparency requirements
- **NO RISK AI SYSTEMS**  
No obligations



# Can chatbots present high risk?



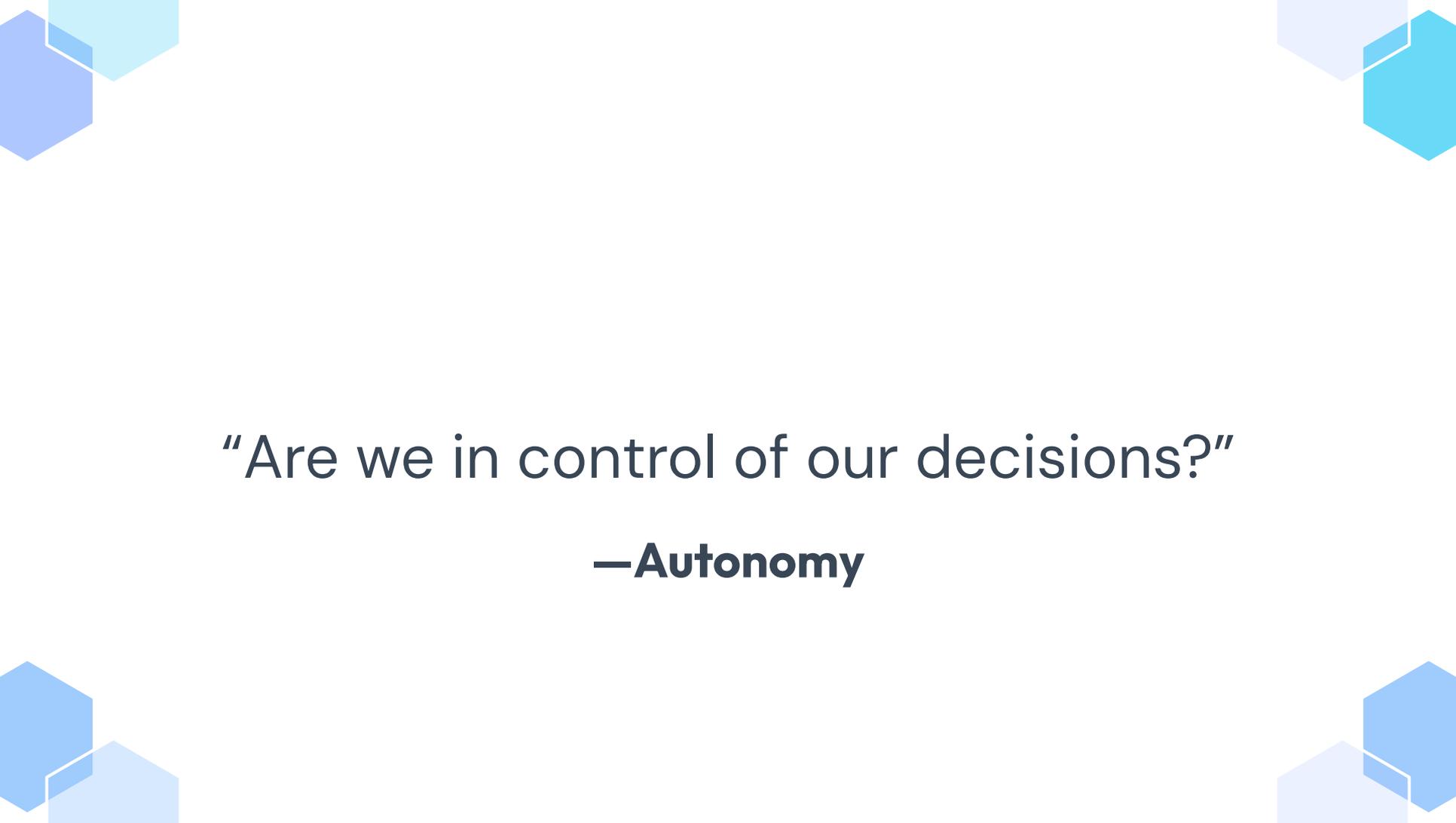
## No

If the user interprets the results of the chatbot while having the bare minimum knowledge about the field(mental), the decision was made autonomously



## Yes

If the user doesn't think about the results and trusts the chatbot, the chatbot made the decision for them (philosophical), **preventing** user autonomy

The slide features a white background with decorative hexagonal shapes in the corners. The top-left and bottom-left corners have overlapping hexagons in shades of blue and light blue. The top-right and bottom-right corners have overlapping hexagons in shades of light blue and blue.

“Are we in control of our decisions?”

—**Autonomy**

# Algorithm Design

Companies write algorithms using the **explore and exploit** strategy to improve **serendipity** for users



# What is Serendipity?

Serendipity is the lucky chance upon stumbling upon something that makes you happy and engaged



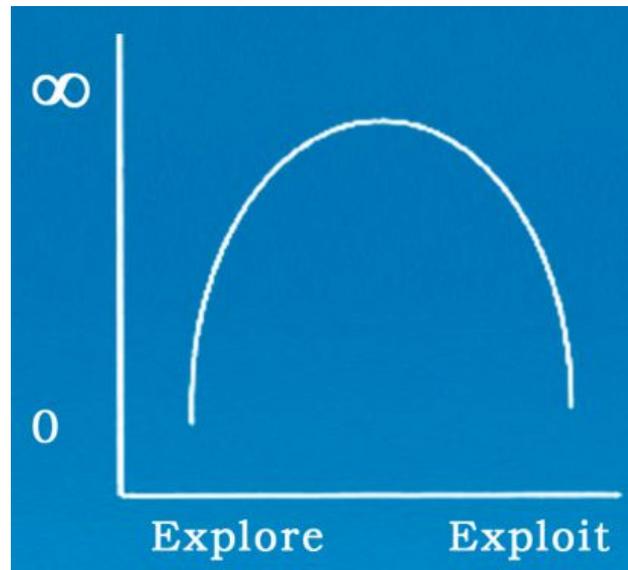
# Explore and Exploit

- **Explore**

Introduce new options to the user(search bars)

- **Exploit**

Give users what they know they love(similar videos, similar artists, etc)



Follow the link in the graph to modify its data and then paste the new one here. [For more info, click here](#)

# One Armed Bandits Literature

An example of the explore and exploit literature is the one-armed-bandits concept

- Stick with a specific slot machine(exploit)
- Use a new slot machine(explore)



# Limitations of Algorithms

Recommendation algorithms should **not** only feature an explore **OR** exploit recommendation algorithm

- Too much exploration can overwhelm the user and make them lose engagements
- Too much exploitation can lead to boredom



# Are algorithms allowing us to make decisions autonomously?



**No**

Certain algorithms hurt user attention spans over time causing a lack of psychological reasoning leading to creating decisions without autonomy



**Yes**

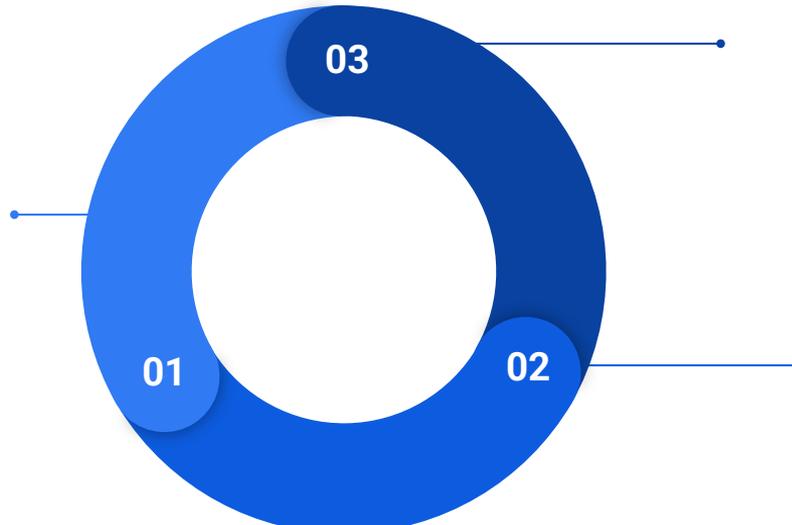
At the end, it is the end user that decides on choosing the content they want, you have to make the call to action

# Autonomy Dilemmas:

From nudging to Structural control

Autonomy Challenges exist on a spectrum

Soft behavioral nudge



Structural life path control

Preference shaping

# Netflix autoplay

Soft manipulation → Minimal risk  
recommendation system

The features:

- Automatically plays next episode
- Reduces friction and active decision making
- Increases engagement time



NETFLIX



NETFLIX

# Tinder & the “One Best Choice” Dilemma

Medium level autonomy influence

How it works:

- Algorithm recommend one “optimal” partner
- Reduces exploration of alternatives
- Maximizes statistical compatibility



# LinkedIn & Structural Autonomy Limits

High-impact autonomy shaping

## The System

- Job opportunities based on data profile
- Limits experimentation outside digital identity
- Shapes long-term professional trajectory



2021- now

## Autonomy Question

Can you choose freely if options are pre-filtered?

# Manipulation vs Autonomy Spectrum

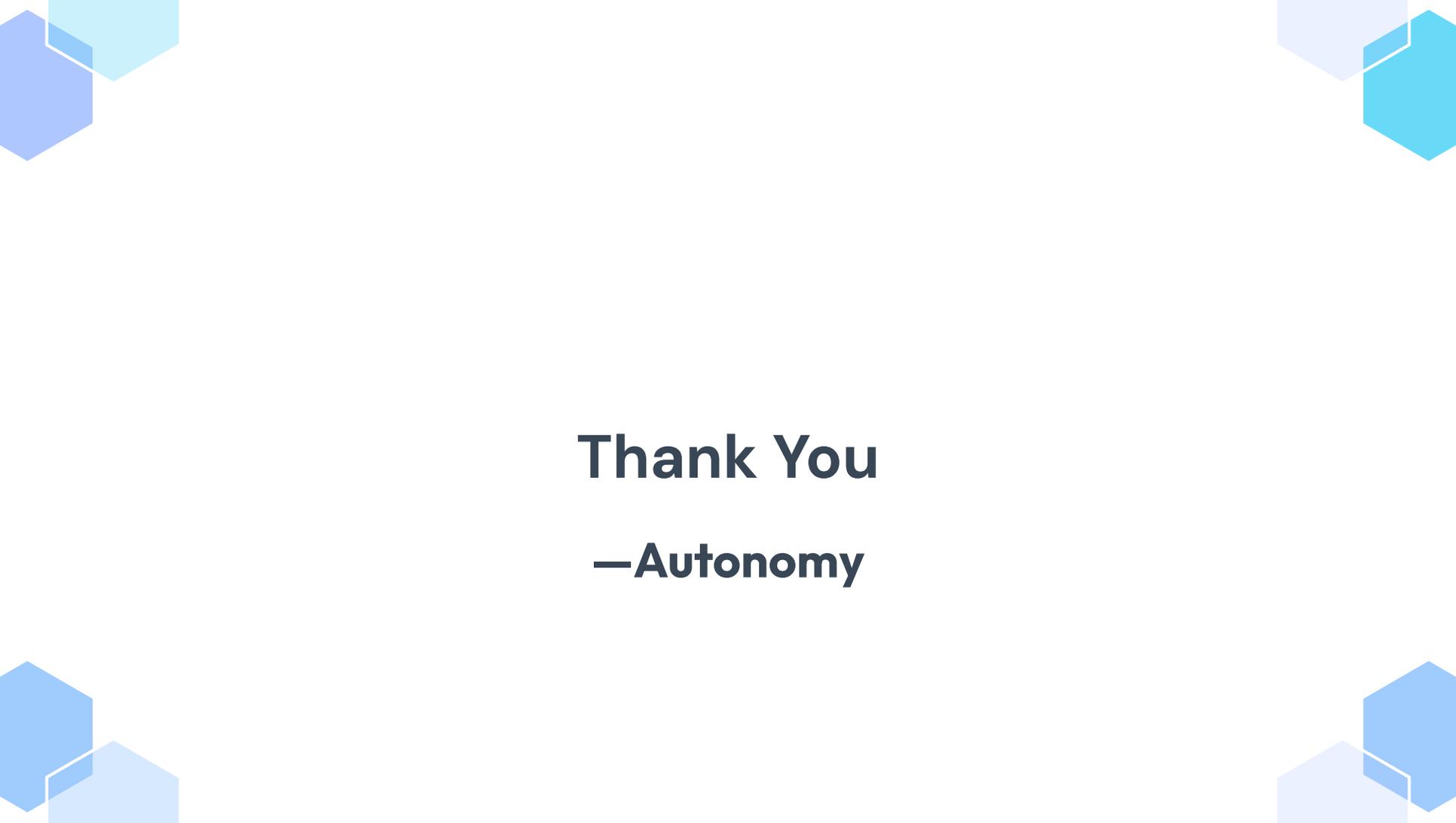
Where ethical design makes the difference

## Design that preserves autonomy:

- Transparency about influence
- Easy opt-out mechanisms
- Alignment with user goals

## Design that undermines autonomy:

- Hidden manipulation tactics
- Hard-to-resist defaults
- Prioritizing profit over agency

The slide features a white background with decorative hexagonal shapes in the corners. The top-left and bottom-right corners each contain two overlapping hexagons: one in a medium blue and one in a light blue. The top-right and bottom-left corners each contain two overlapping hexagons: one in a light blue and one in a medium blue.

**Thank You**  
**—Autonomy**