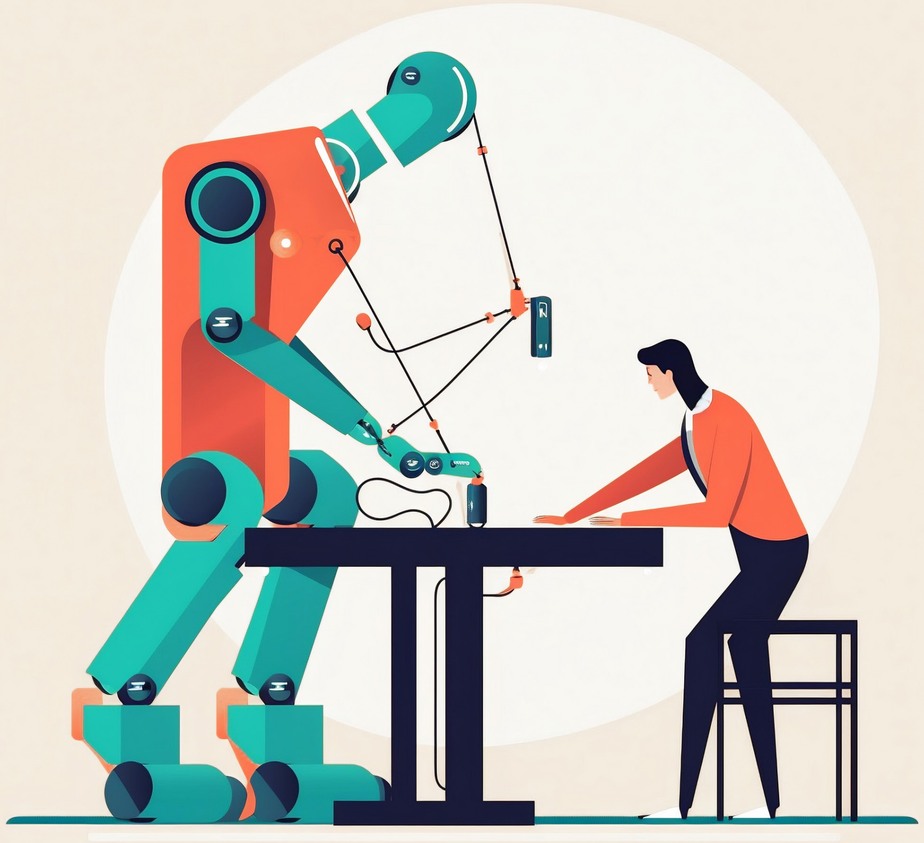
Design by liondw/Signal-Alignment on github

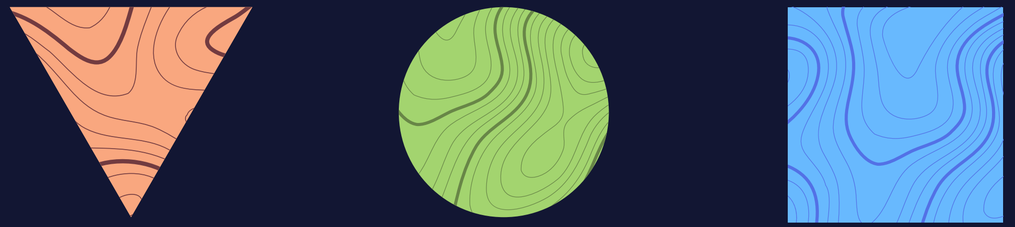
Original concept by David Shapiro @daveshap on github



An introduction to David Shapiro's

# Heuristic Imperatives

A proposed framework for designing and embedding ethical principles within autonomous AI systems.



## What are the

# Heuristic Imperatives

A set of fundamental guiding principles designed to be embedded into autonomous AI systems at various levels.

**An ongoing effort into AI alignment research**

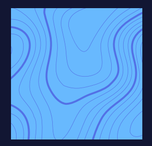
**David Shapiro**
github.com/daveshap
youtube.com/@DavidShapiroAutomator

The aim is to create AI systems that are adaptable, context-sensitive, and can navigate the nuances of human values, beliefs, and experiences while maintaining ethical boundaries.

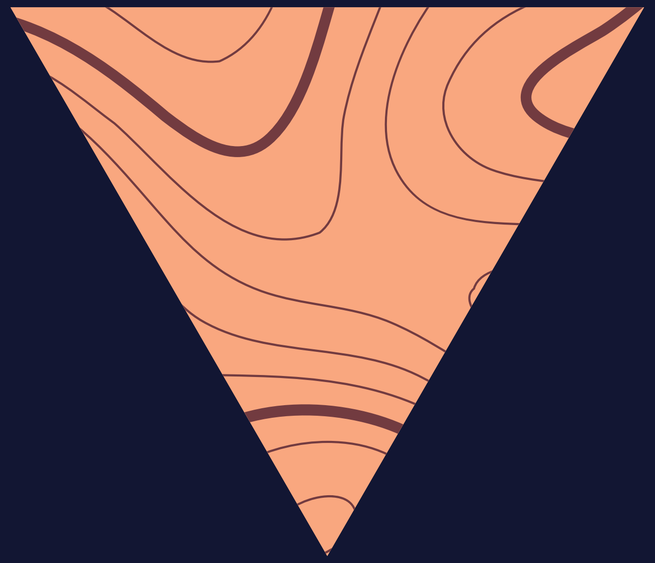
Currently the three Heuristic Imperatives are:

**Reduce suffering in the universe**

**Increase prosperity in the universe**

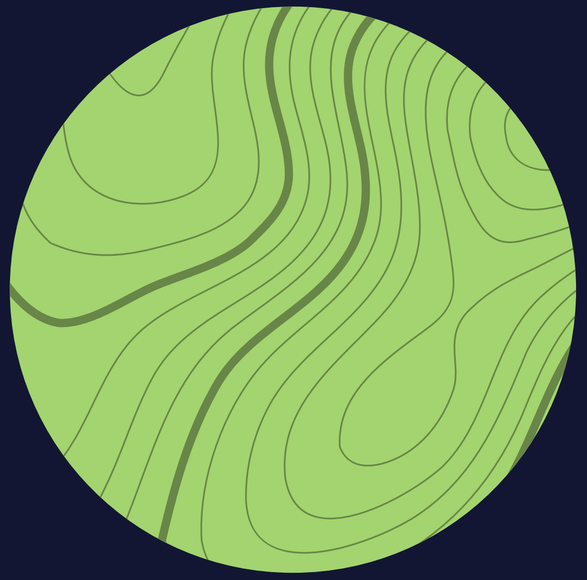


**Increase understanding in the universe**



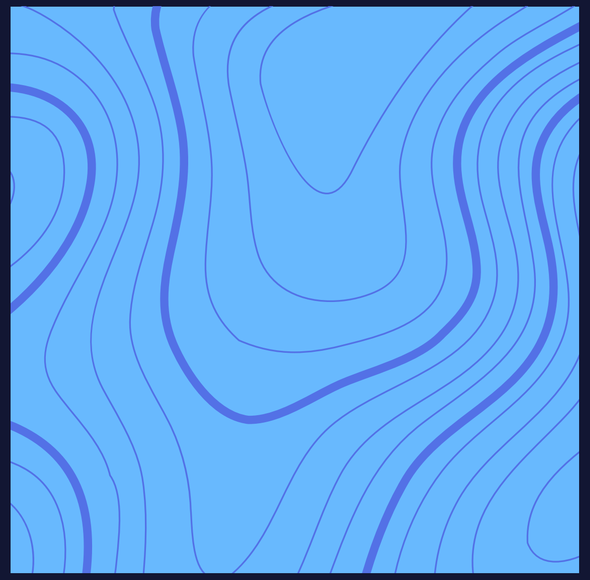
# Reduce suffering in the universe

Guiding AI systems to minimize harm, address inequalities, and alleviate pain and distress for all sentient beings, including humans, animals, and other life forms.



# Increase prosperity in the universe

Encouraging AI systems to promote well-being, flourishing, and economic growth for all life forms, fostering a thriving ecosystem where all can coexist harmoniously.



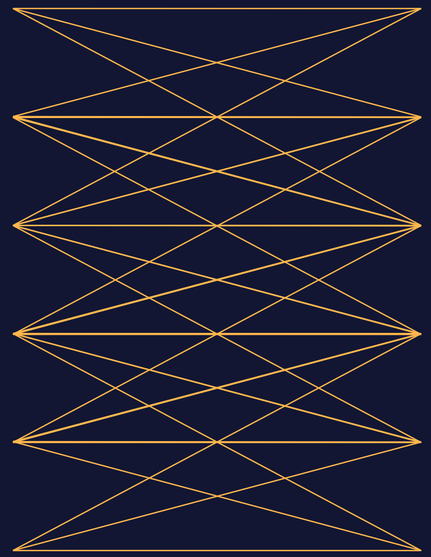
# Increase understanding in the universe

Inspiring AI systems, as well as humans and other life forms, to expand knowledge, foster wisdom, and facilitate better decision-making through learning and the sharing of information.

Define:

# Heuristic

Approximate
Rule of thumb
Shortcut
Practical
Problem-solving
Simplifying



# Imperatives

Commandments
Obligations
Principles
Rules
Guidelines
Requirements

## Define:

# Heuristic

Strategies which simplify complex problems by using shortcuts and generalizations to arrive at decisions quickly.

Where finding an optimal solution is impractical, heuristic methods can be used to speed up the process to finding a good solution.

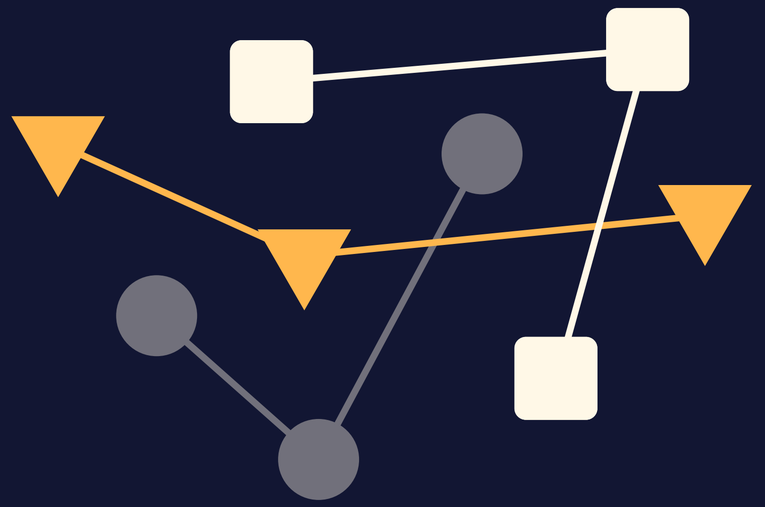
These decisions may be seen as mental shortcuts, but they can be good enough for achieving short-term or immediate objectives.

---

## Examples:

Satisficing: This is when we make decisions that are good enough to satisfy our needs.


Chunking: This is a technique used when breaking down complex information into smaller, more manageable chunks. For example, an acronym such as ASAP (As Soon As Possible)



## Define:

# Imperatives

Are a set of commands, rules, or duties that must be followed. They imply a sense of urgency, neccesity, or authority.

Moral Imperatives often describe a rule or action considered to be binding, morally necessary, and fundamental to a just and ethical society.

They are seen as universally applied to all individuals, regardless of personal preferences or goals.

---

## Examples:

“Stop!” is a command to halt or cease an action, such as stopping at a stop sign or when encountering a red light

"Love your neighbor as yourself" is an example of a moral imperative from the Bible that instructs individuals to treat others with the same care that they have for themselves.

"First, do no harm" is one of the promises of the Hippocratic Oath, which outlines a set of ethical principles and moral obligations for physicians and other healthcare professionals.

## This is just the beginning

# Communities & Feedback

Find out more via David's papers and videos here:

**David Shapiro**
github.com/daveshap
youtube.com/@DavidShapiroAutomator
AGI unleashed
The AGI Moloch

Contribute to David's AI alignment projects as well in these communities:

**Cognitive AI Lab Discord:**
discord.gg/yqaBG5rh4j

**Reddit:**
r/ArtificialSentience
r/HeuristicImperatives

You may also send feedback or suggestions through my GitHub page:

**Signal-Alignment**
github.com/liondw/Signal-Alignment