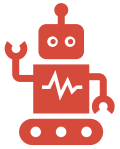
PsychSim

By: Grayson Pray

Why Psychim?

While Big 5 is used widely in research, and often in industrial psychology, a holistic model combining Big 5, Dark Triad, SVO, and Prospect Theory remains absent and nearly none of these models are used in recreational services like video games. Utilizing a model that takes the interaction of these traits could lead to more lifelike and accurate psychological representations than a single model alone.

Many ABM models either use (a) random behavior, (b) hand-coded rules, or (c) simplistic economic rationality. PsychSim provides an empirically-validated framework for personality-driven agent behavior that's calibrated against and grounded in published psychology research.



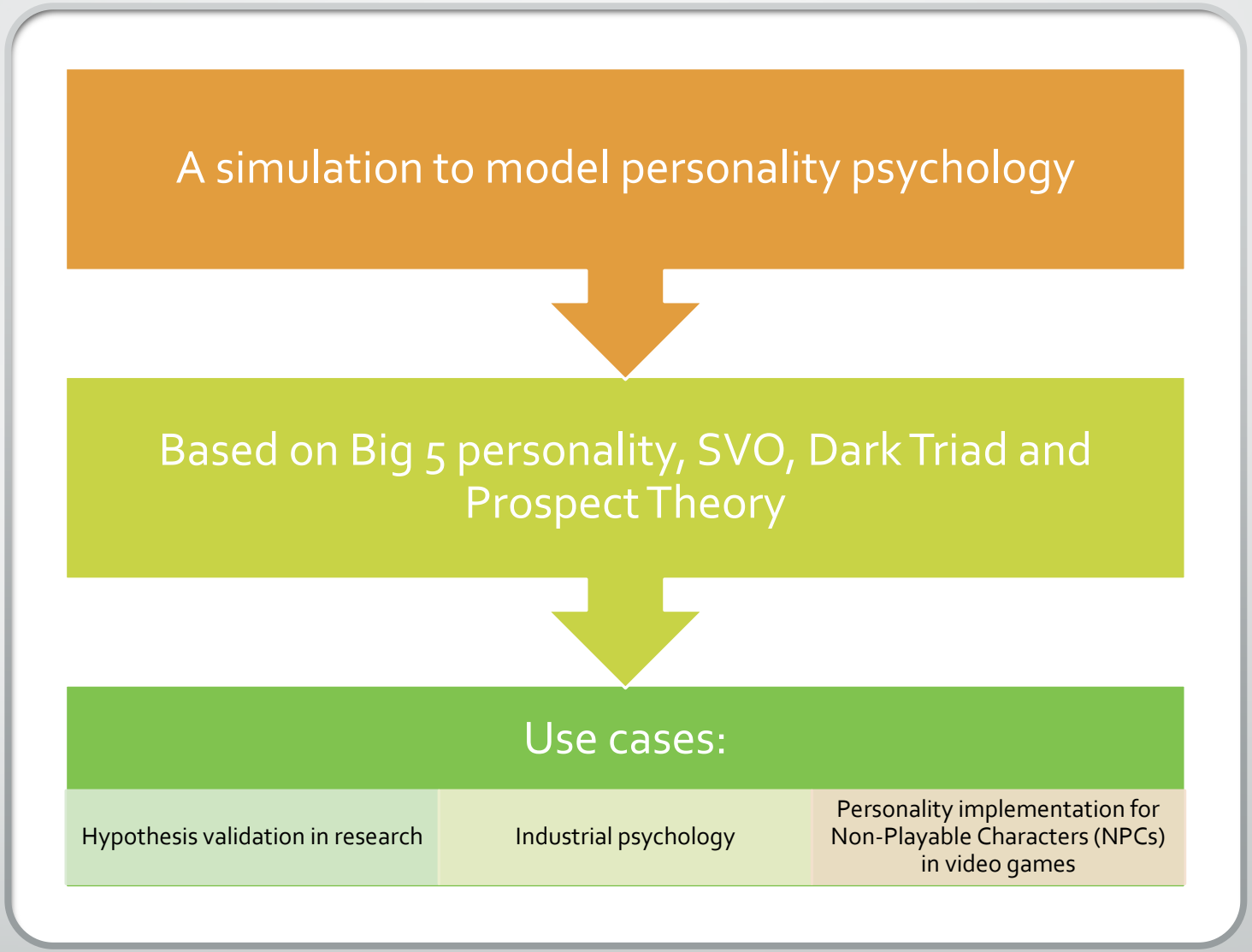
Interpretability: In line with validating against psychological research, recent AI-driven character models act as a "black box" where determining the epistemic validity of their processes is opaque. Psychsim provides a more adaptable and multivariate solution than hand-coded rules, but avoids the "black box" of AI by directly interpretable algorithms and processes.

And why Mesa?

- Netlogo
  - Spatial (grid/patch) thinking is the default mental model, which doesn't fit social/psychological simulation as naturally as agent-centric design
  - Lack of scientific and data anlysis libraries as in python (Sci-kit, pandas)
- Repast:
  - Strong for large-scale and HPC simulations, but much steeper setup cost. The Java ecosystem is far from the modern data science stack
- GAMA:
  - Purpose-built (Even own language GAML) for geographic/spatial simulations with GIS integration. Excellent if agents are moving through a map of a city, but less so for agents that are driven by psychology

- Mesa
  - Every library in the scientific Python ecosystem (NumPy, pandas, scikit-learn, matplotlib, TensorFlow) is a direct import away with no bridging layer, allowing for better analysis integration
  - Better ease-of-use and interactability, which are ideal features for research and gamedev audiences

What it is


A simulation to model personality psychology
Based on Big 5 personality, SVO, Dark Triad and Prospect Theory
Use cases:
Hypothesis validation in research
Industrial psychology
Personality implementation for Non-Playable Characters (NPCs) in video games

Feature List Completion

- MVP
  - Personality Model ✅
  - Needs System ✅
  - Personality-Weighted Decision Engine ✅
  - Basic Actions ✅
  - Relationship Tracking ✅
  - Resource Environment ✅
  - Mesa Integration ✅
  - Basic Data Export ✅
  - Configuration System ✅
  - Unit Tests ✅
  - Documentation ✅
- 1.0
  - Agent Memory System ✅
  - Group Formation and Dynamics ✅
  - Trading System ✅
  - Conflict System with Balance Mechanisms ✅
  - Mesa Visualization Server ✅
  - Scikit-learn Integration for Analytics ✅
  - Agent Customization Interface ✅
- 2.0
  - TensorFlow/Keras Memory Embeddings
  - LLM-Generated Agent Dialogue
  - Web-Based Configuration UI ✅
  - Scenario System ✅
  - Multi-Simulation Comparison Tools ✅
- Extras
  - Data Analysis ✅
  - Deployment online ✅
  - Weight Calibration ✅

Data

- Murphy & Ackermann (2014) — SVO Slider Measure range/distribution
- Paulhus & Williams (2002) — Dark Triad x Big Five correlations
- Muris et al. (2017) meta-analysis — DT intercorrelations
- Denissen et al. (2022) — SVO x Big Five correlations
- Big 5 Dataset: Open-Source Psychometrics Project. (2018). Big Five Personality Test [Dataset]. https://www.google.com/search?q=https://openpsychometrics.org/_raw-data/
  - 50-item IPIP representation of the Goldberg (1992) markers.

Design

Agent-Based Modeling core

Predominantly Object-Oriented Programming paradigm

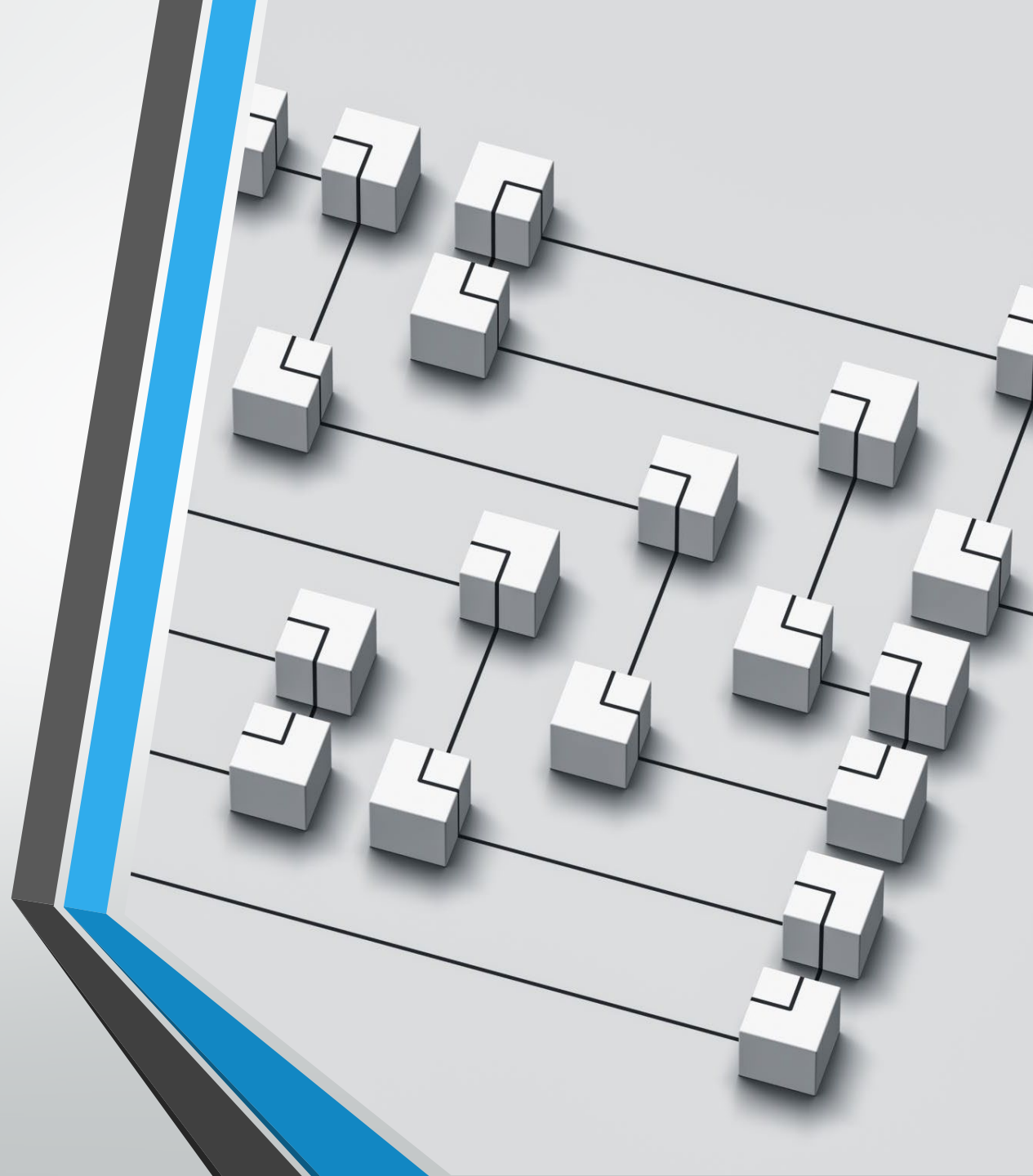
- Each main component of the environment and agent features has its own delegated class

Hash-map heavy (Such as keying agent-ids in memory.py, or for configuration)

Probabilistic decision-making

Pipeline

1. Entry Point & Model Initialization
2. Personality Generation
3. Agent Initialization
4. The Per-Tick Loop
5. Environment Layer
6. Group Formation
7. Data Collection & Export
8. Data Analysis



Model Initialization

Runs via main.py. Basic versions are demo, headless, interactive

- Headless is main form of data collection and simulating
- Arguments can be added to change defaults

Inherits Mesa's Model to create SocialSimModel

- Some key aspects are ContinuousSpace, RandomActivation, ResourceManager, and DataCollector

Config is loaded from YAML with layered checking for alternatives to default. (defaults → local → CLI-specified)

Personality Generation

- Each agent gets a 9-dimensional personality covariant matrix sampled from a multivariate normal distribution. The traits and their inter-correlations come from the empirical psychology literature
- Sampling uses factory method PersonalityFactory.sample() which draws from MVN(μ, Σ) and clamps to valid ranges

Trait Group	Traits	Source
Big 5	Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism	Digman 1990; John & Srivastava 1999
Dark Triad	Narcissism, Machiavellianism, Psychopathy	Paulhus & Williams 2002
SVO	SVO angle (−16.26° to 61.39°)	Murphy & Ackermann 2014

	O	C	E	A	N	Narc	Mach	Psy	SVO
O	1.00	0.08	0.15	0.06	-0.05	0.38	0.02	0.24	0.06
C	0.08	1.00	0.07	0.30	-0.13	0.05	-0.34	-0.24	0.08
E	0.15	0.07	1.00	0.12	-0.08	0.41	0.05	0.17	0.05
A	0.06	0.30	0.12	1.00	-0.15	-0.36	-0.47	-0.25	0.19
N	-0.05	-0.13	-0.08	-0.15	1.00	0.00	0.02	-0.34	-0.05
Narc	0.38	0.05	0.41	-0.36	0.00	1.00	0.25	0.50	-0.20
Mach	0.02	-0.34	0.05	-0.47	0.02	0.25	1.00	0.40	-0.35
Psy	0.24	-0.24	0.17	-0.25	-0.34	0.50	0.40	1.00	-0.28
SVO	0.06	0.08	0.05	0.19	-0.05	-0.20	-0.35	-0.28	1.00

Agent Initialization

- Each agent is an inherited object from Mesa's model class. For each agent, four sub-systems are instantiated:
  1. Personality
  2. Needs
     - Hunger, Energy, Safety, Social, Esteem
       Urgency to meet needs is formulated with a quadratic curve, escalating urgency as needs decay.
  3. Relationships
  4. Decision-making

Per Tick Loop

- Needs Decay
- Perception
- Action Generation
- Decision Engine
- Action Execution
- Relationship Updates

Environment

Resources spawn in clusters and follow a seasonal cycle plus a logistic carrying-capacity cap

The seasonal modifier uses a sine wave tied to a 365-tick year as follows:

$$\text{modifier} = 2.0 + \text{variability} \times \sin(2\pi \times (\text{day} - 91) / 365)$$

Group Formation

Every 50 ticks, the model checks all agent pairs. If two agents have mutual sentiment > 0.3 and aren't already groupmates, they form or join a group

Groups are tracked by GroupManager in groups.py which auto-assigns the founder as leader and dissolves empty groups.

Data Collection

Mesa's DataCollector records model-level and agent-level reporters every tick. Additionally, raw action and event logs are maintained

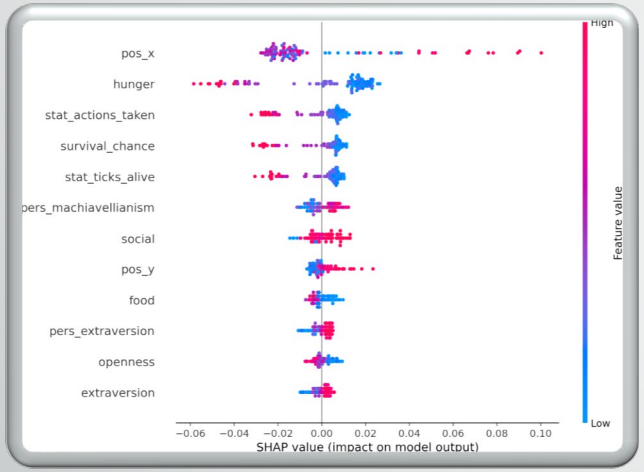
Three csvs and a final state json are produced that track different aspects:

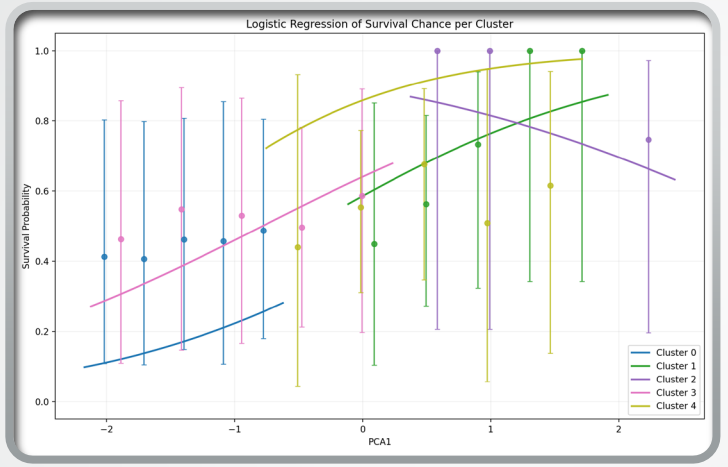
- actions.csv: every action taken by every agent
- events.csv: deaths, combat outcomes, group formations
- model_stats.csv: per-tick aggregates (alive count, total food, sentiment, season)

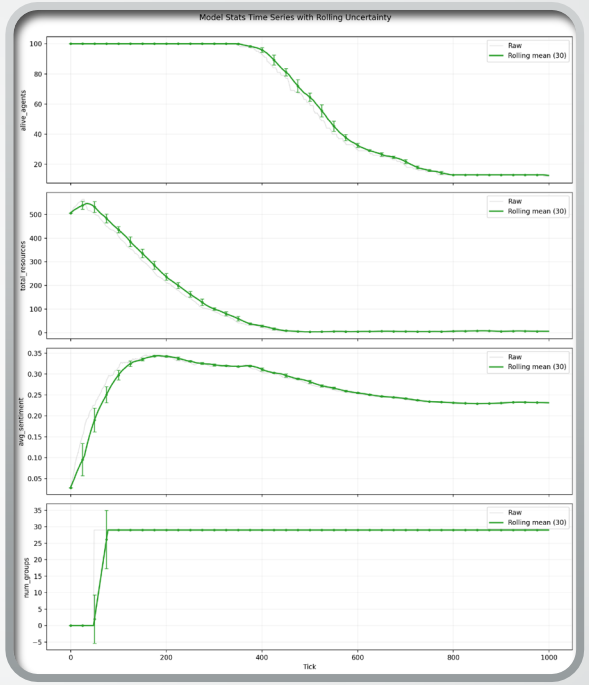
Data Analysis

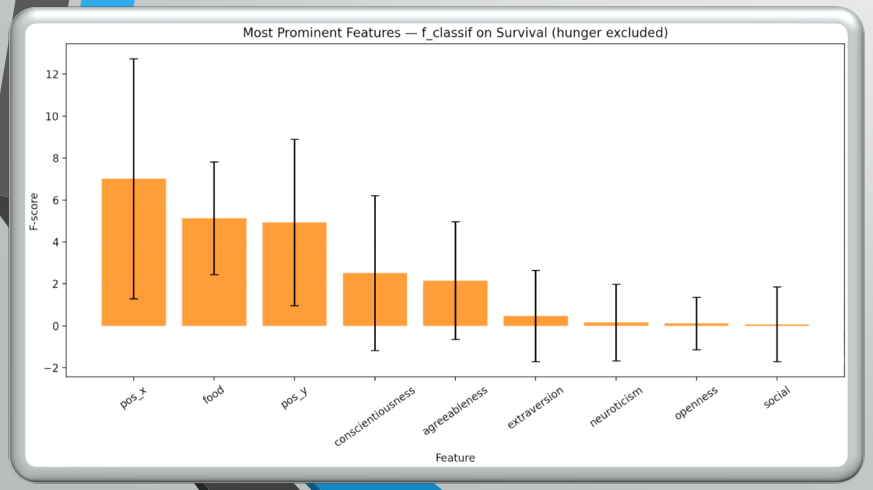
Feature selection of most prominent feature	PCA Clustering of personalities	Random Forest sentiment importance via SHAP
Time-series via matplotlib	Logistic Regression on survival chance	Pearson correlation of traits

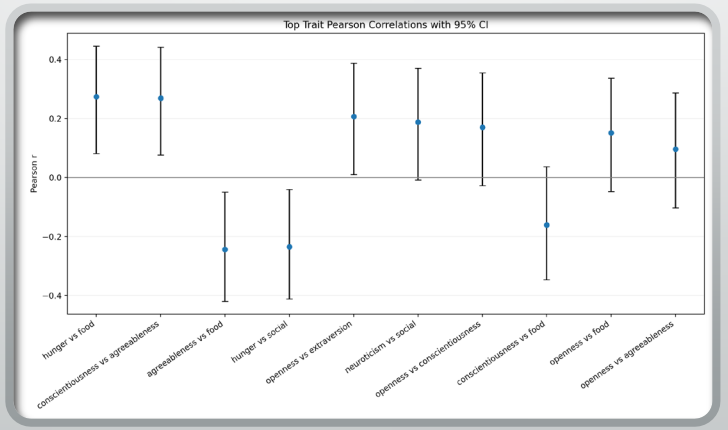
Results

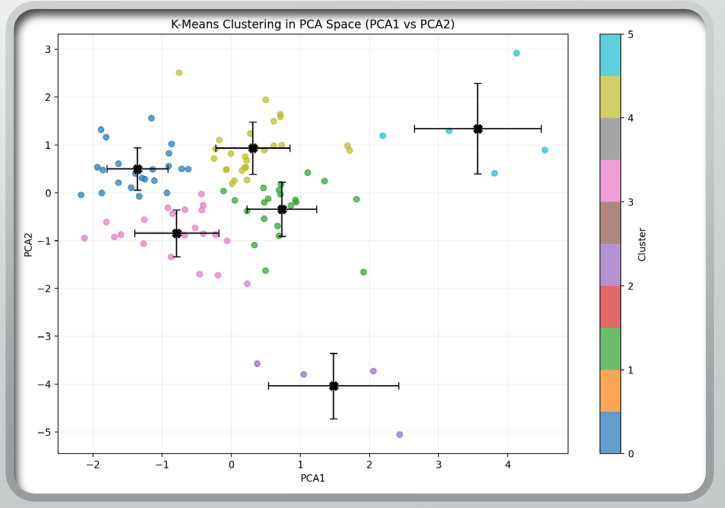

pos_x
hunger
stat_actions_taken
survival_chance
stat_ticks_alive
pers_machiavellianism
social
pos_y
food
pers_extraversion
openness
extraversion
−0.06 −0.04 −0.02 0.00 0.02 0.04 0.06 0.08 0.10
SHAP value (impact on model output)
High
Low
Feature value


Logistic Regression of Survival Chance per Cluster
Survival Probability
PCA1
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4


Model Stats Time Series with Rolling Uncertainty
alive_agents
total_resources
avg_sentiment
num_groups
Raw
Rolling mean (30)
Tick


Most Prominent Features — f_classif on Survival (hunger excluded)
F-score
pos_x
food
pos_y
conscientiousness
agreeableness
extraversion
neuroticism
openness
social
Feature


Top Trait Pearson Correlations with 95% CI
Pearson r
hunger vs food
conscientiousness vs agreeableness
agreeableness vs food
hunger vs social
openness vs extraversion
neuroticism vs social
openness vs conscientiousness
conscientiousness vs food
openness vs food
openness vs agreeableness


K-Means Clustering in PCA Space (PCA1 vs PCA2)
PCA2
PCA1
Cluster

Fine-Tuning Parameters

- CMA-ES
  - Maintains a multivariate Gaussian distribution over the parameter space, sample a population of candidates from it, evaluate them, then update the distribution to put more probability mass near the better candidates.
- Reference dataset using known effect sizes as anchors
- Composite Loss function:
  - KL divergence on SVO type distribution
  - Cosine distance on Random Forest feature importances
  - Sign match + MSE on trait-behavior correlation matrix

Constant	Sources	Confidence Reliability
Big Five block	Hurtz & Donovan 2000	High
DT×Big Five block	Paulhus & Williams (2002)	High
DT intercorrelations	Muris et al. (2017)	High
SVO row	(Balliet 2009)	High
SVO Rates	(Balliet 2009)	High
Heatmap of Mach, Psy, SVO rows	Hilbig (2014); Peeters (2006); Balliet (2009); O'Boyle (2012)	Moderate
Heatmap of Big 5 rows	Hurtz & Donovan (2000)	High

Visualization

https://psychsim-7db21be686e1.herokuapp.com/

Limitations and future work

Limitations

- Benchmark data for the intersection between Big 5, DT, SVO, and environment was approximated from literature, and not fully empirically grounded
- Virtual agents will always be abstractions of humans and their behavior, rather than directly synonymous. Environment as well

Future work

- Implement more resource types such as water, and materials
- Convert naming of agents to being read in by txt file instead of specifically named, for better scalability