Differentiating 'Human in the Loop' Decision Strategies

Sarah Walsh and Dr. Karen Feigh

Cognitive Engineering Center

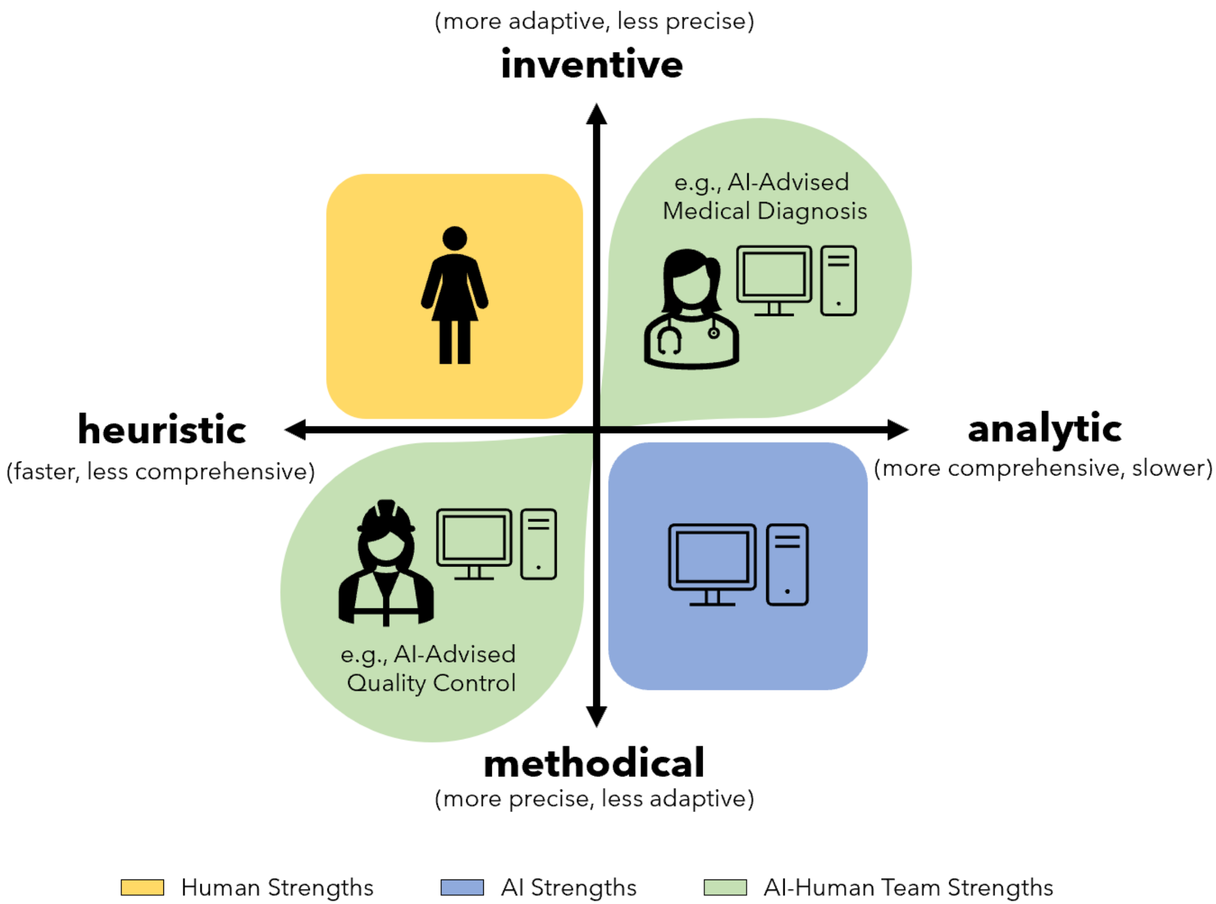
Georgia Institute of Technology

Research Questions

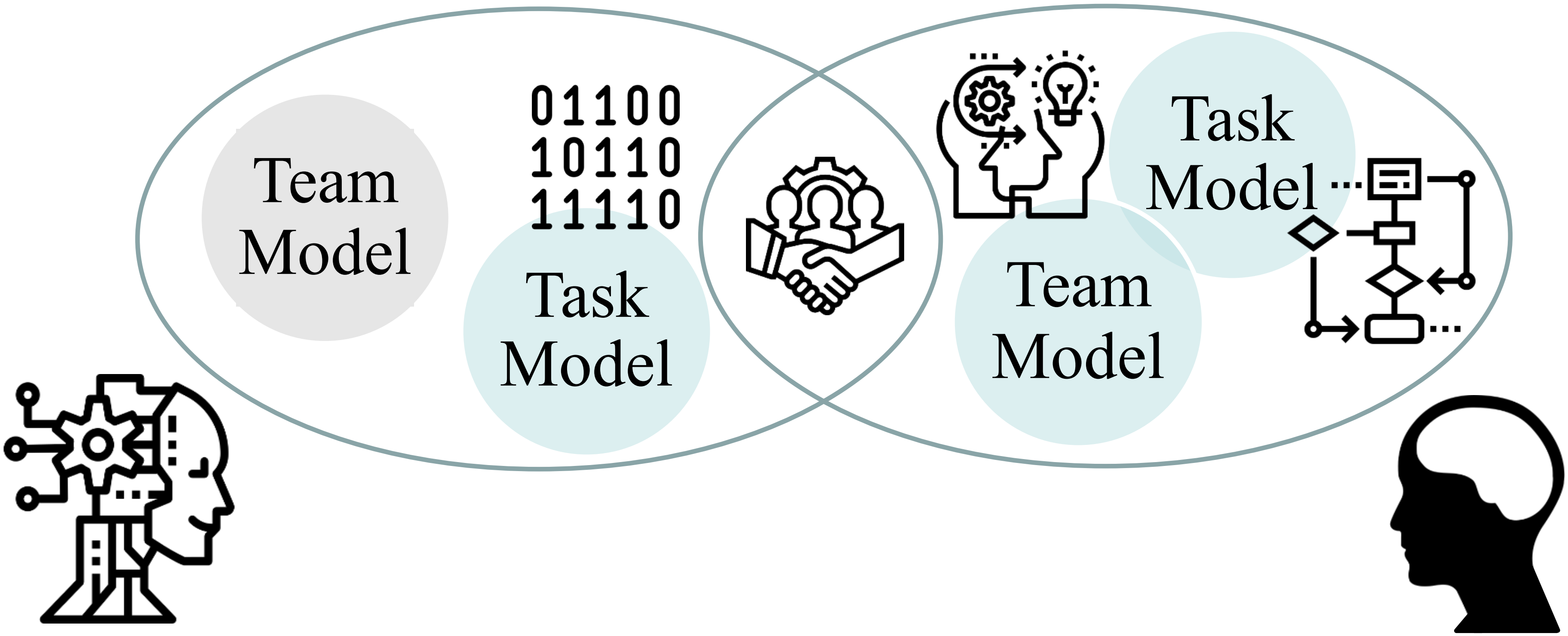
1. Can we infer decision strategies from dynamic behavioral data?
2. Can we detect when people diverge in their decision making approach?
3. Can we classify these inferred decision strategies based solely on behavioral data?

Cognitive Engineering Center

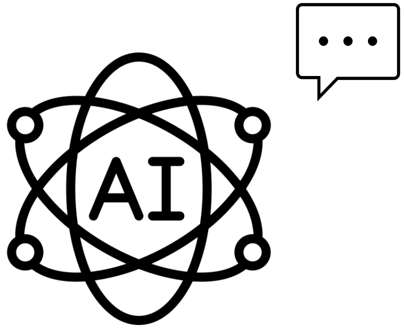
Human-AI Decision Making



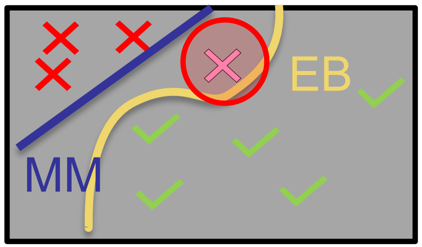
Limitations in Human-AI Shared Mental Models



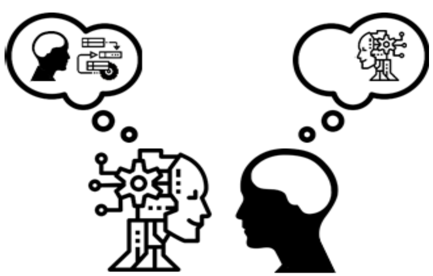
Improving Human-AI Teaming:



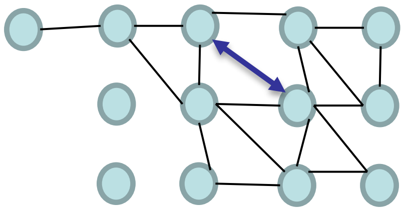
Explainable AI
(Shin 2020, Preece 2021, Arya 2019)



Improving human mental models of AI error boundaries (Bansal, 2020)

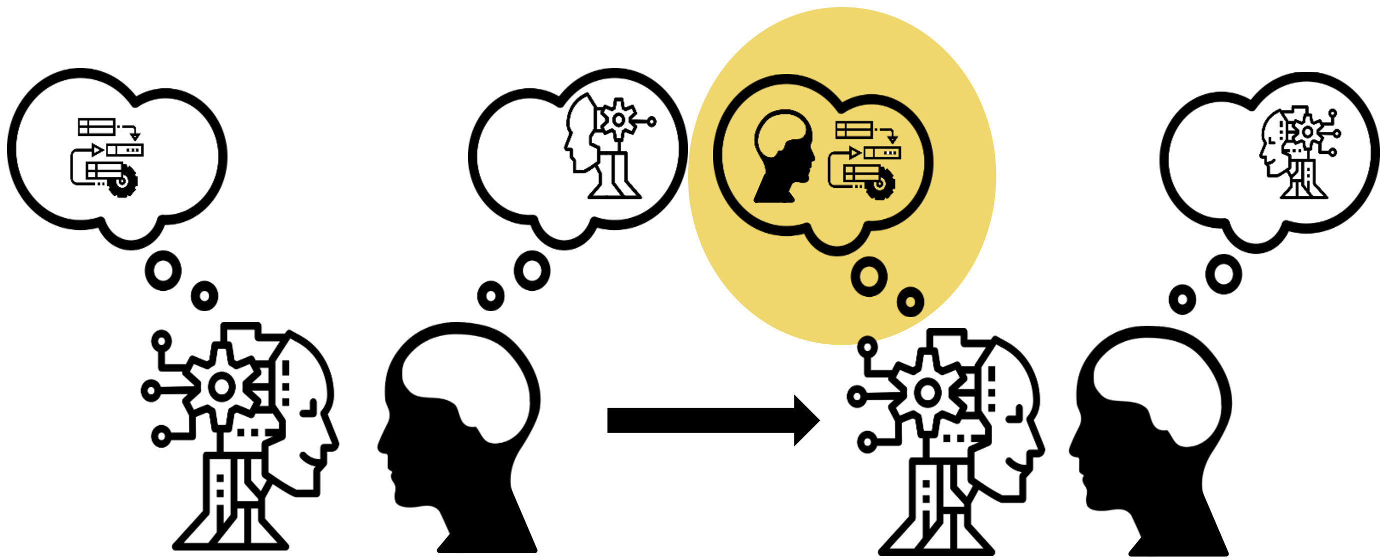


AI systems create Shared Mental Model with human teammates (Scheutz 2017)



Bridging gaps between the AI and human's relative policies (Bastani, 2021)

Create Shared Mental Model



- AI in human-AI teams often operate with little or no model of the human's cognitive state
- We need `learning human mental models' that are easy of an AI system to train and can support planning/decision-making (Chakraborti et al., 2017)

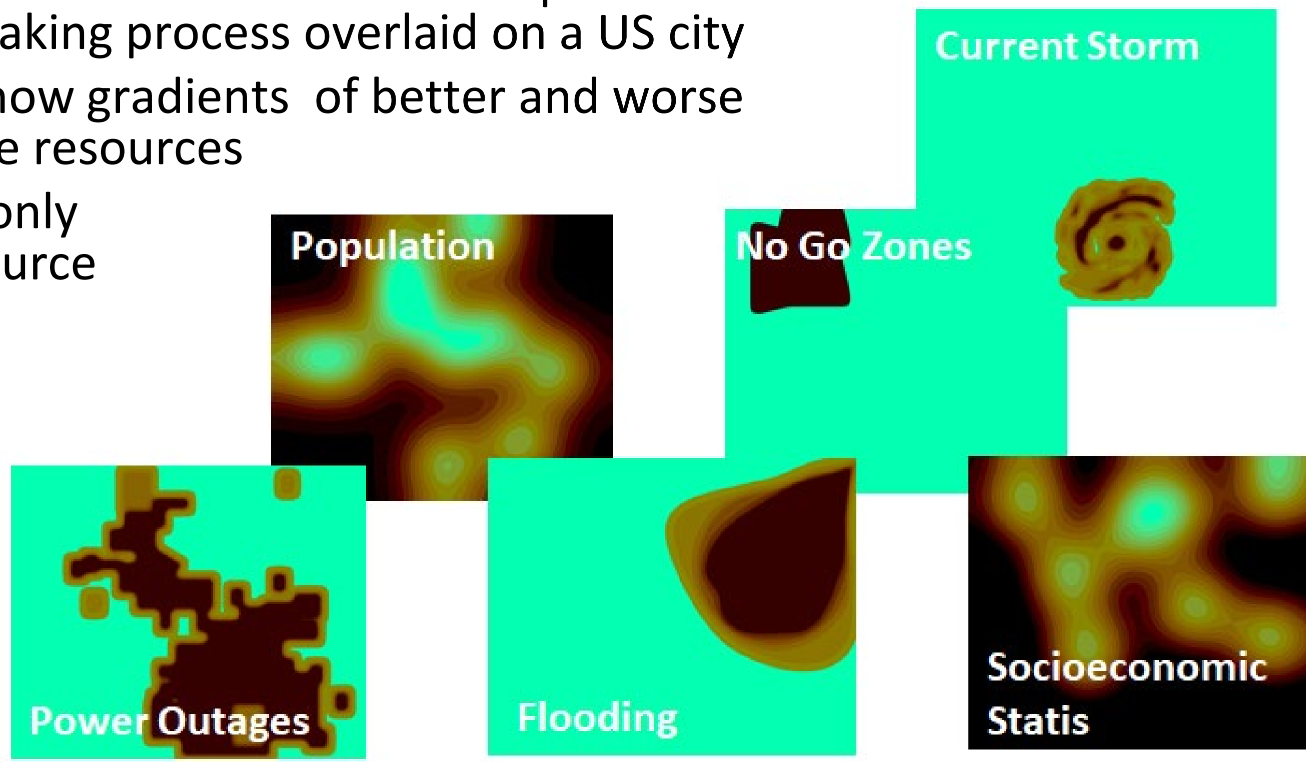
Approach to Learning Human Mental Models of Decision Making

1. Design an experiment to capture real-world decision making
2. Capture and classify decision strategies

Cognitive Engineering Center

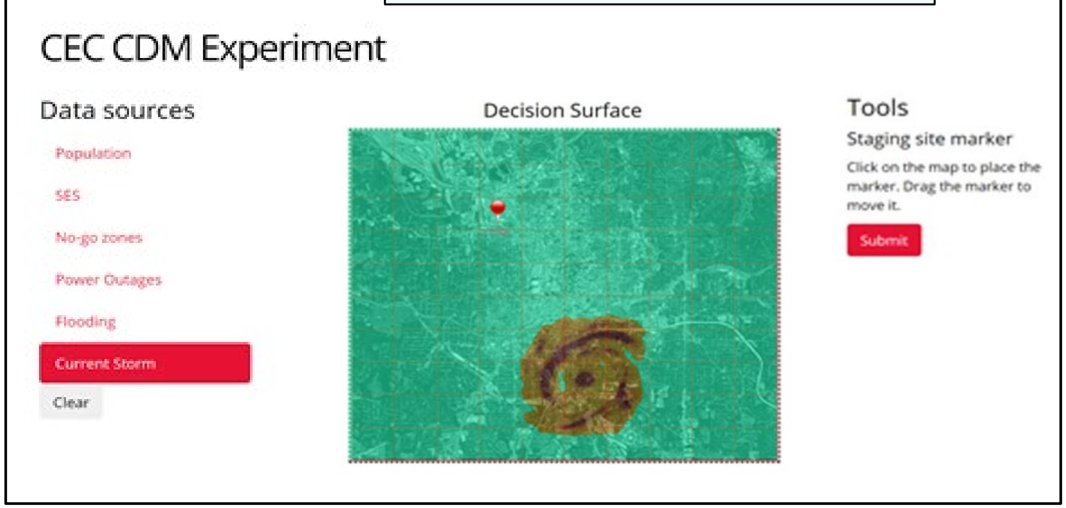
Experiment with Geospatial, Sequential Task

- Participants will be assuming the role of a disaster relief planner making decisions about how to allocate resources prior to and during a storm
- The participants will have several heat maps that will aid in the decision making process overlaid on a US city
- The heat maps show gradients of better and worse locations to place resources
- Participants can only observe one resource at a time



Experiment in a Sequential Environment

Experiment Interface

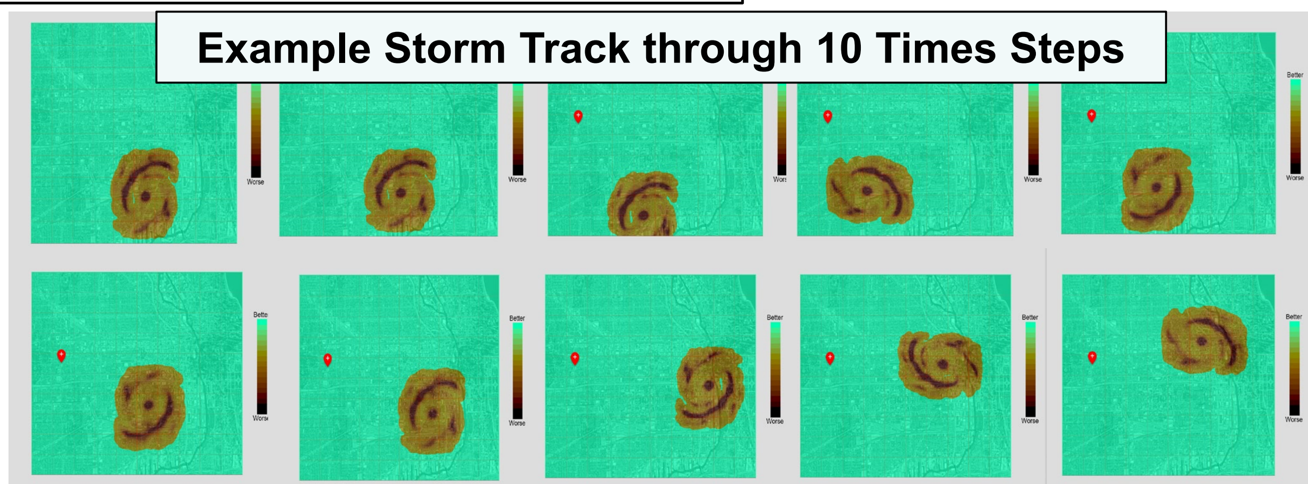


Storm tracks change through time. Each time a storm track changes the participant is asked to update their decision (resource location)

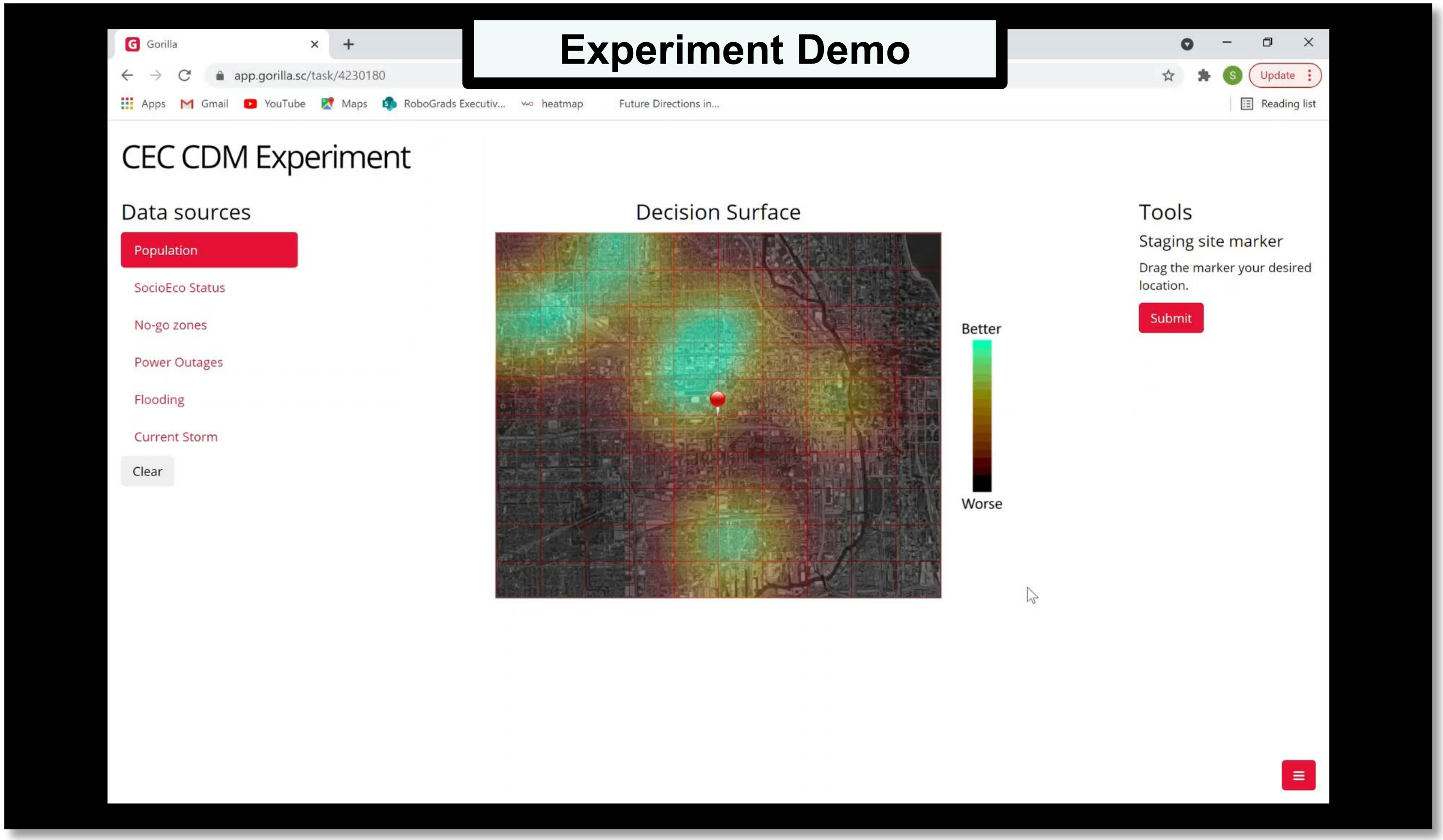
<u>Experiment Features</u>

- 10 timesteps / decision events
- 6 data sources (3-dynamic and 3-static)
- 1 resource placements

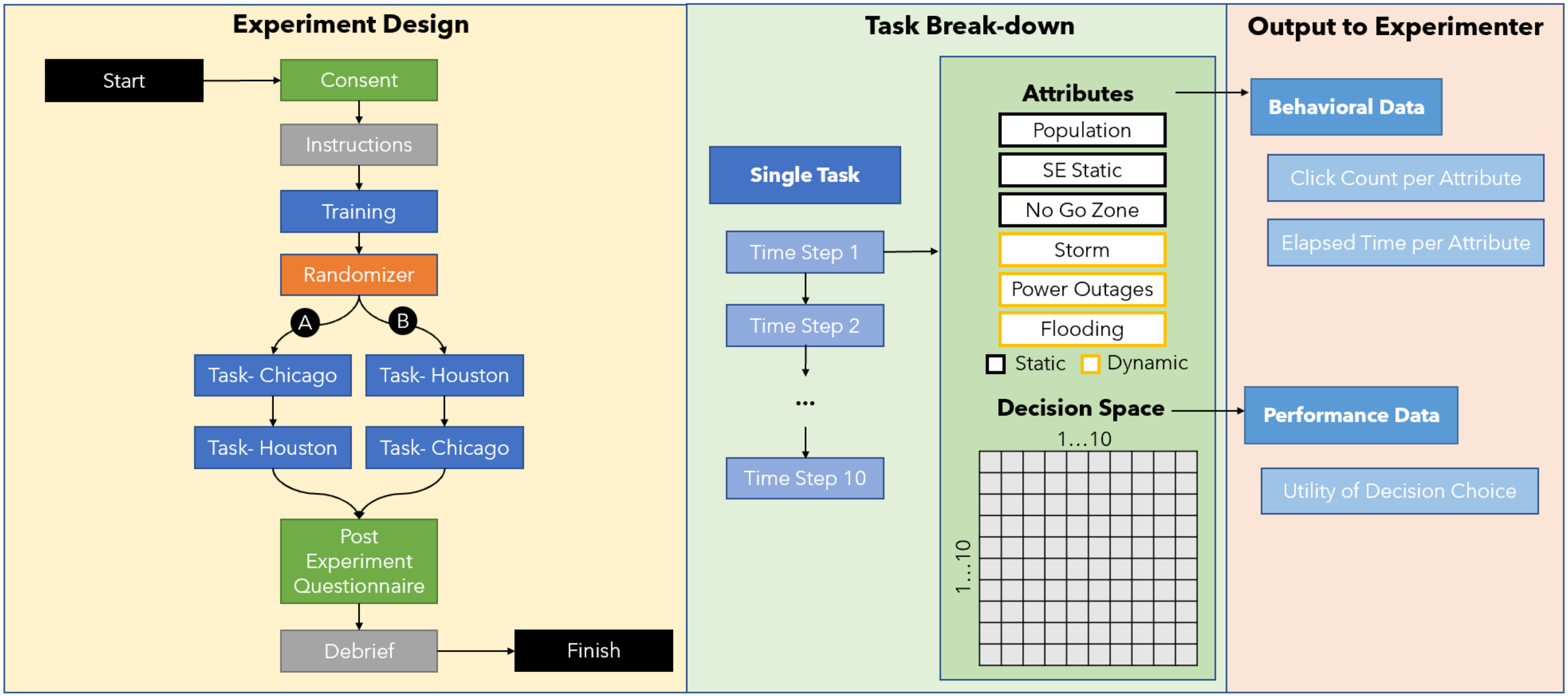
Example Storm Track through 10 Times Steps



User-Interface



Administration of Experiment



Approach to Identifying DM Strategies

Part 1: Label data using Partial Least Square Regression

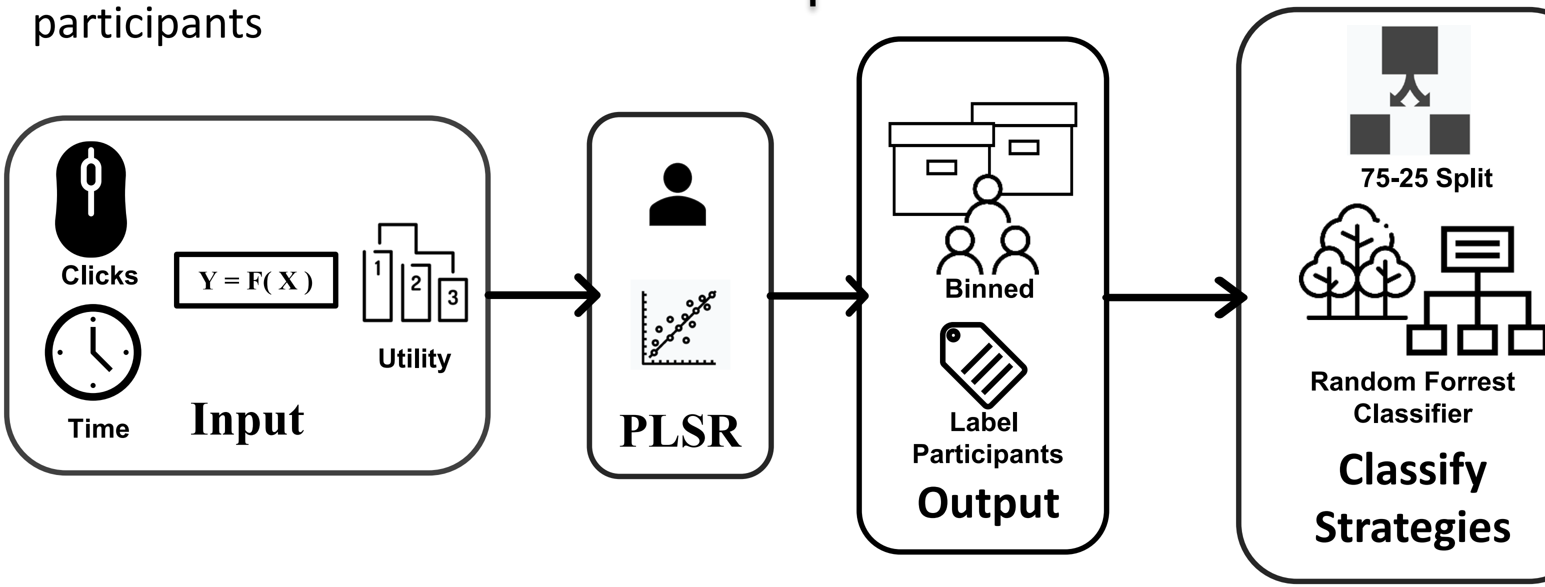
Goal:

- Use behavior to classify decision strategies and predict decision strategies/mental models of participants

Part 2: Reverse analysis to classify using Random Forrest

Goal:

- Classify DM strategy
- The output of the random forest is the class selected by most trees

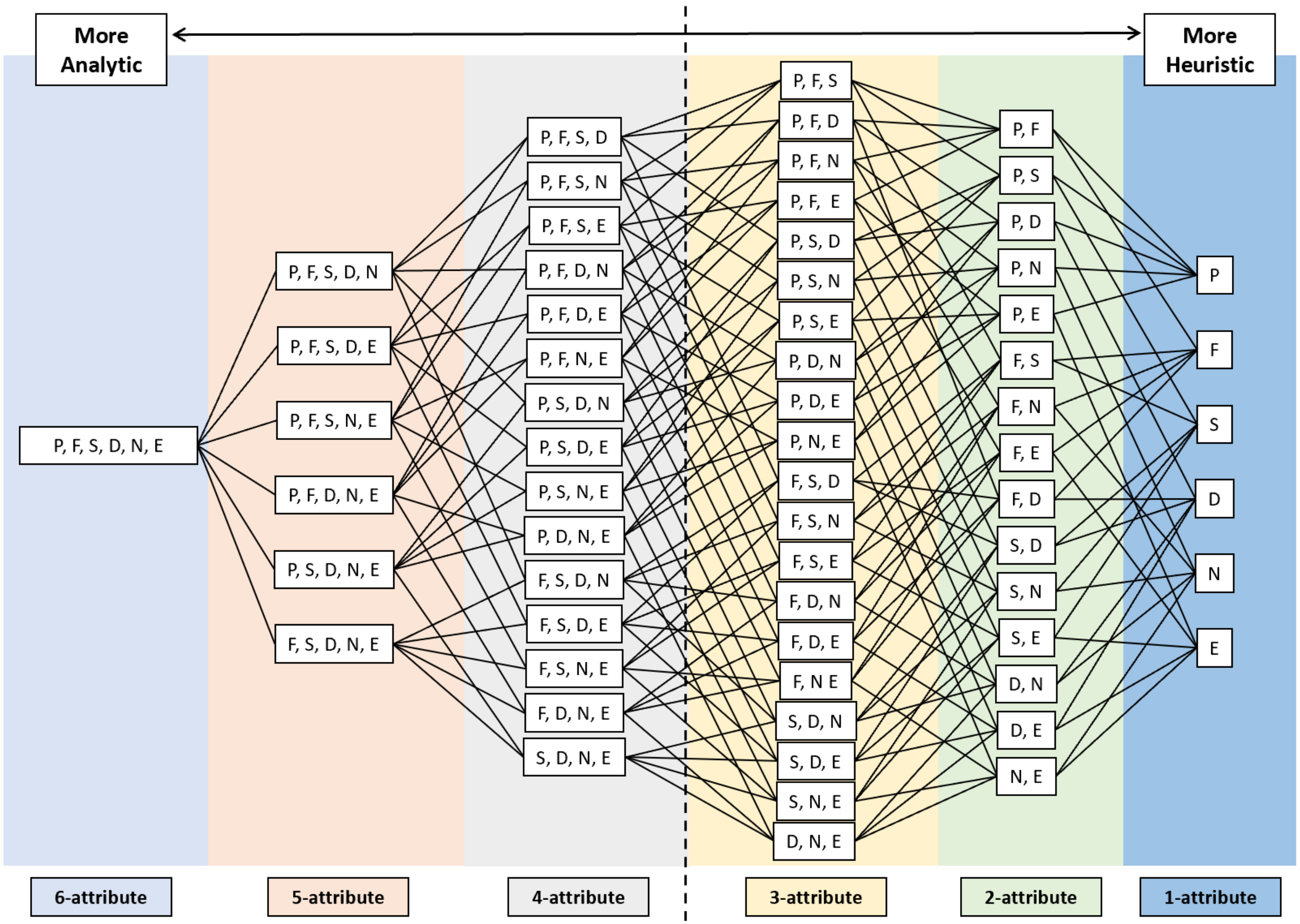


Results

1. Can we infer decision strategies from dynamic behavioral data?
2. Can we detect when people diverge in their decision making approach?
3. Can we classify these inferred decision strategies based solely on behavioral data?

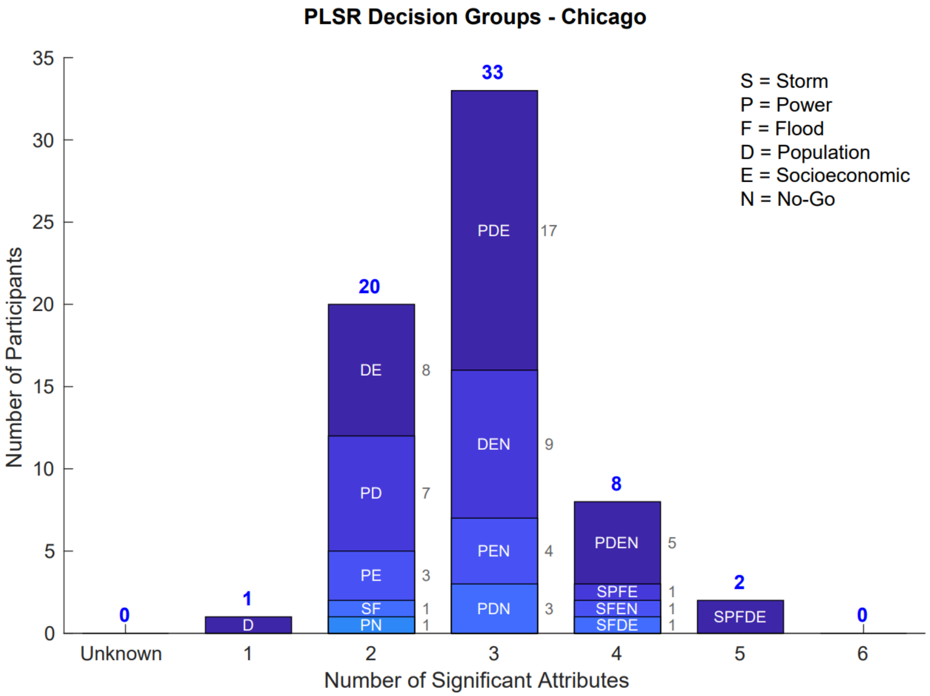
COGNITIVE ENGINEERING CENTER

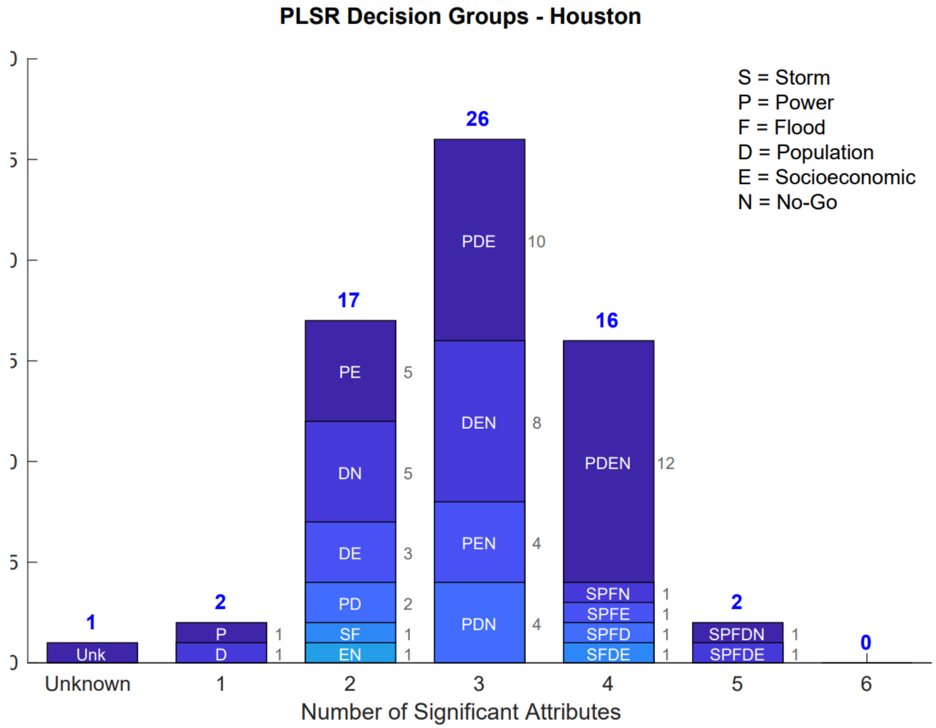
Diagram of Possible Decision Strategies



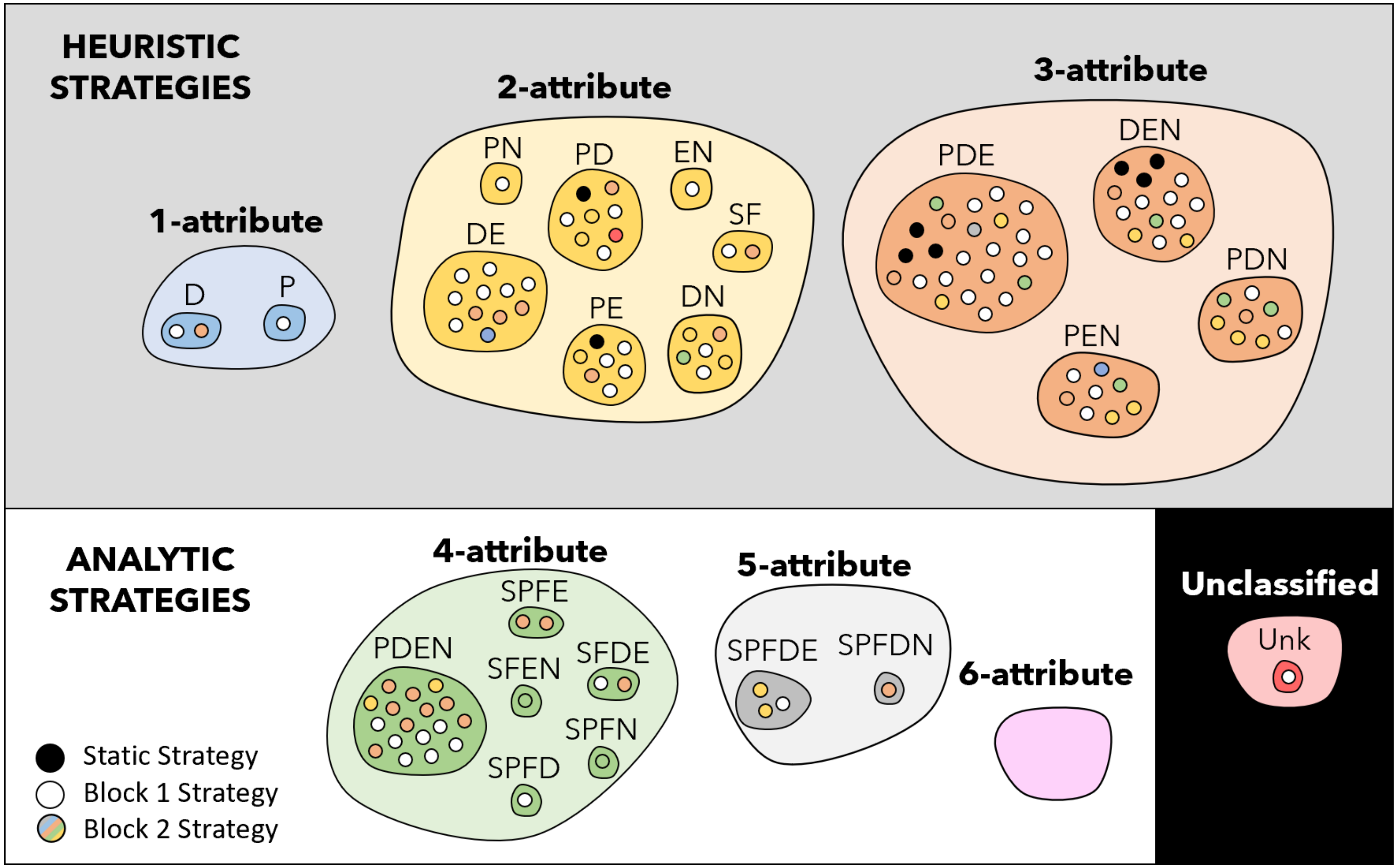
Attribute	Abbrev.
Power Outages	P
Population **D**ensity	D
Storm	S
No Go Zones	N
Flooding	F
Socio**E**conomic Status	E

Results: Can we infer decision strategies from dynamic behavioral data?





Results: How stable are people's decision strategies?

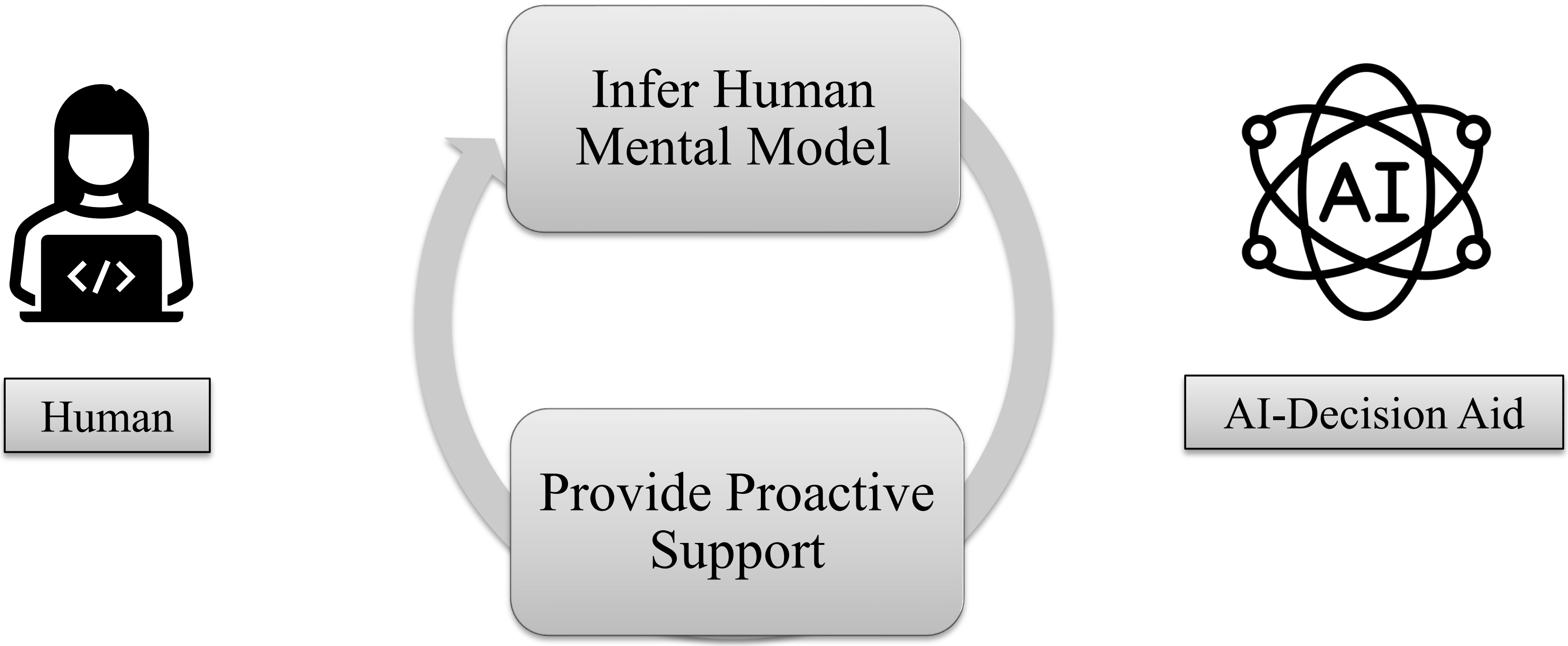


Results: Can we classify individuals into these inferred decision strategies based solely on observable behavioral data?

Chicago			
Number of Trees:			700
No. of splitting vars.:			10
OOB estimate of error rate:			7.8%
Confusion Matrix:			
	Analytic	Heuristic	Class Error
Analytic	373	27	6.7%
Heuristic	38	394	8.8%

Houston			
Number of Trees:			700
No. of splitting vars.:			10
OOB estimate of error rate:			19.1%
Confusion Matrix:			
	Analytic	Heuristic	Class Error
Analytic	319	56	14.9%
Heuristic	85	273	24.3%

Implications and Next Steps



Differentiating ‘Human in the Loop’ Decision Processes

Sarah Walsh and Dr. Karen Feigh

sewalsh@gatech.edu karen.feigh@gatech.edu

COGNITIVE ENGINEERING CENTER

GEORGIA INSTITUTE OF TECHNOLOGY