COORDINATING HUMAN AND AGENT BEHAVIOR IN COLLECTIVE-RISK SCENARIOS

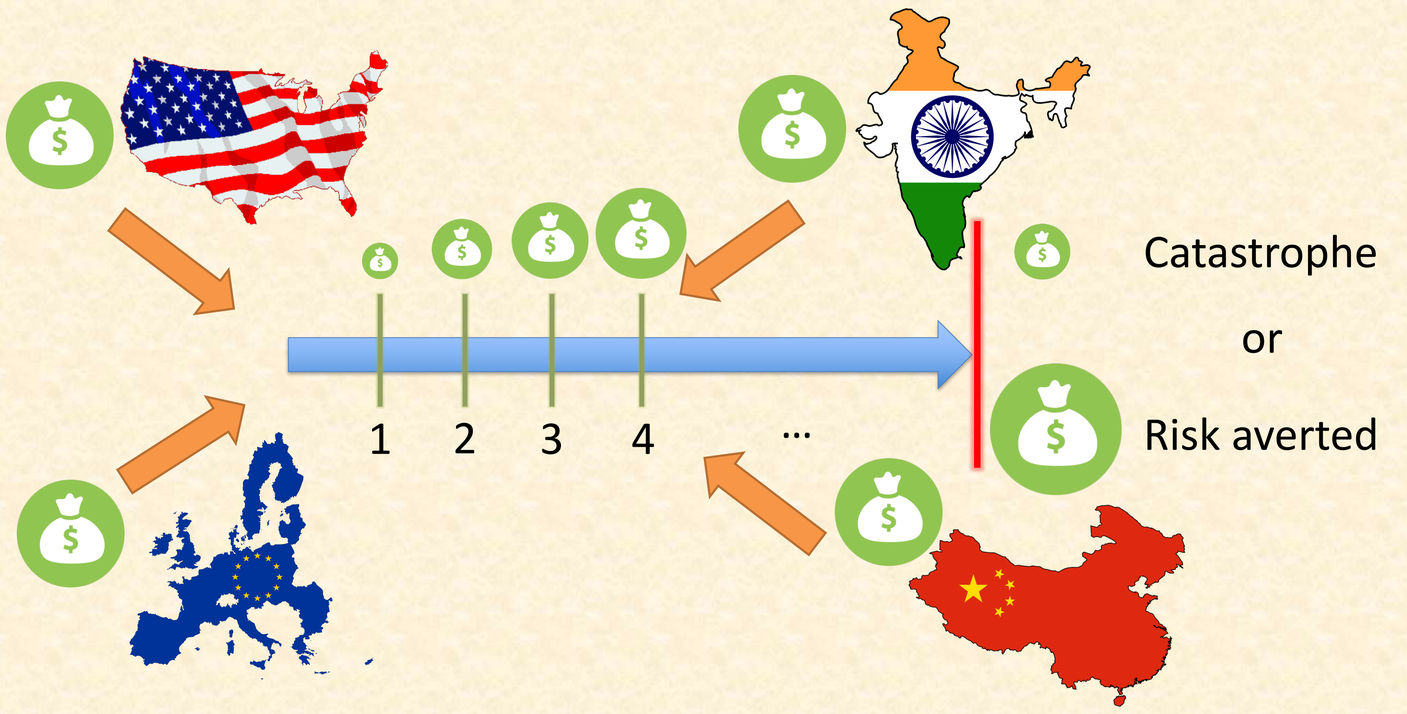
Elias Fernández Domingos[1,2,3], Juan Carlos Burguillo[3], Ann Nowé[2] and Tom Lenaerts[1,2]

[1] AI lab, Computer Science Department, Vrije Universiteit Brussel, Belgium
[2] MLG, Universitéˊ Libre de Bruxelles, Belgium
[3] Department of Telematic Engineering. University of Vigo, Spain

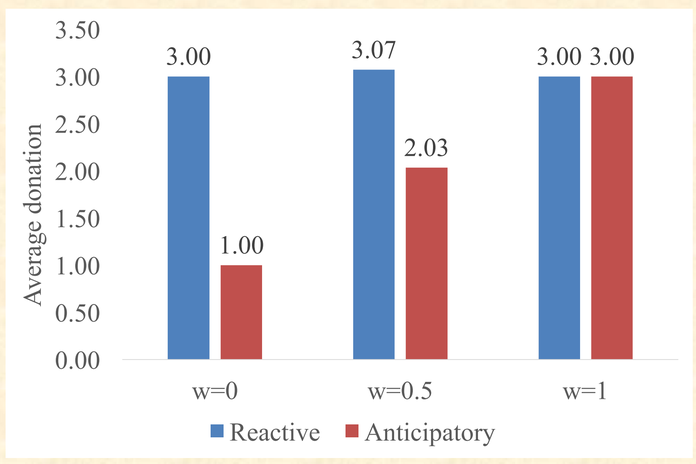
MOTIVATION OF THE WORK

How do humans behave when collective-risk is involved?
Solving the climate change requires the cooperation of several countries with different ideologies, customs, and economical perspectives for their industries, which are in many cases still very dependent on fossil energy.
Measures that need to be taken will have a high impact on industrialized countries as well as the so-called new economies. However, if the transition to renewable sources of energy keeps being postponed, the consequences are most certainly terrible. **This situation has been operationalized in Game Theory as the Collective-Risk Game [1].**
Our ultimate goal is to apply the scientific and technological development that we produce to **hybrid-technological systems** (e.g., P2P energy markets) where **humans have to interact with artificial agents** and **agents have to negotiate among themselves** taking into account human's preferences and utilities [2].

In the **Collective-Risk Game**, players are requested to achieve a common target under the risk of loosing all their money otherwise

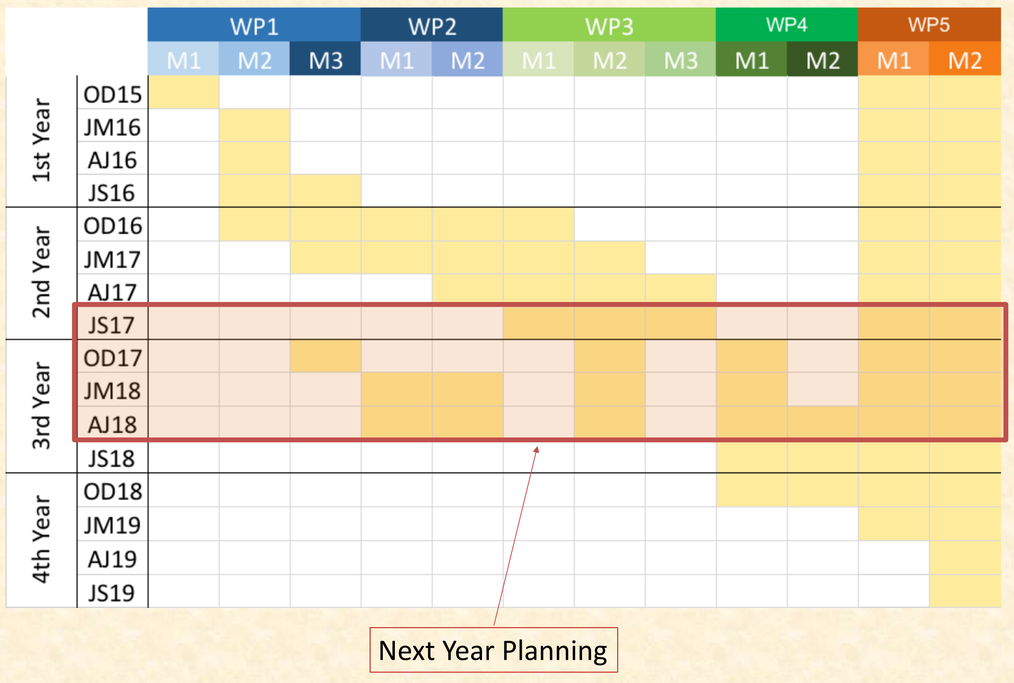


ANTICIPATORY AGENTS ARE ABLE TO ACCOUNT FOR CHANGES IN THE CONTEXT OF GAMES



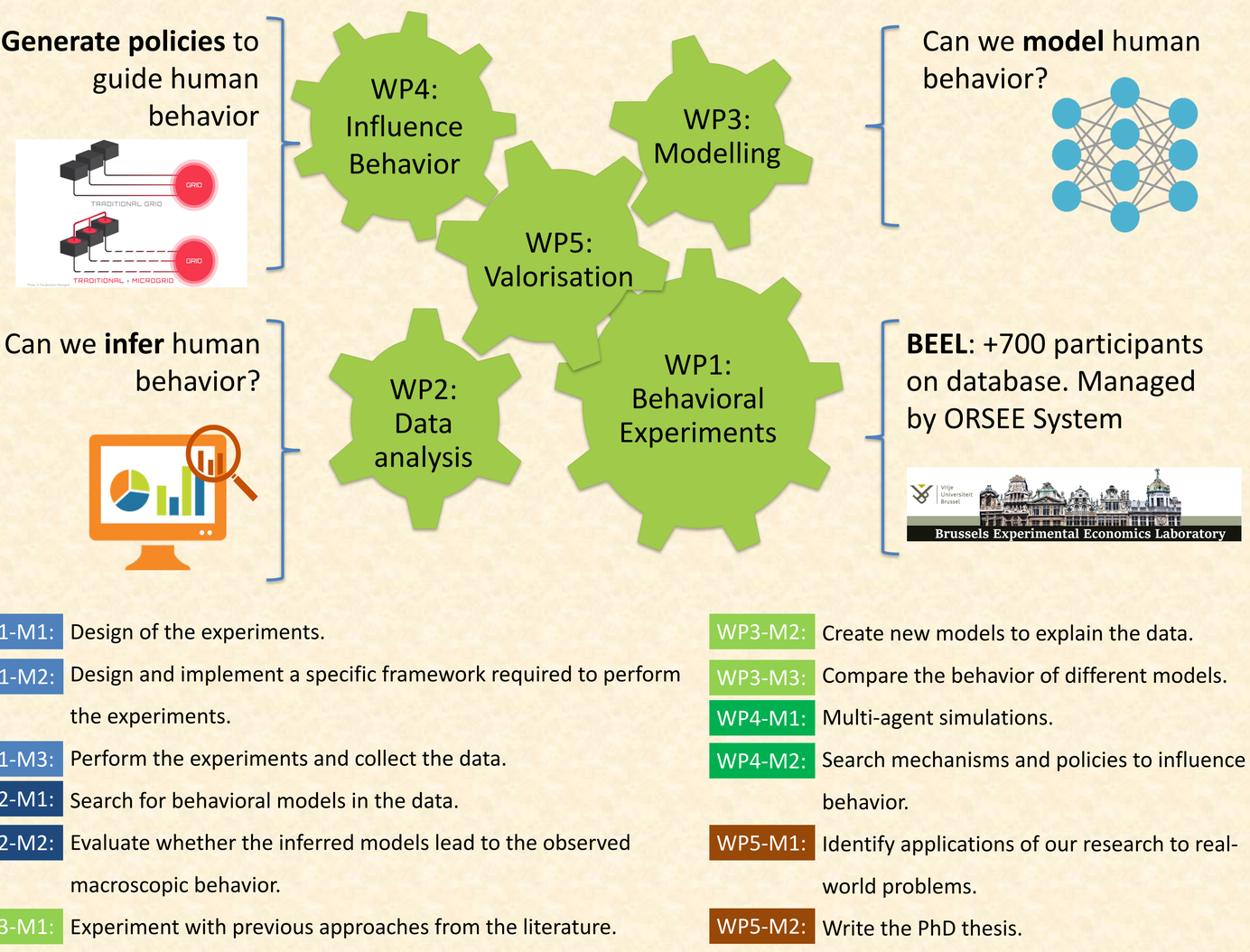
We were able to show that anticipatory models represent better the data from behavioral experiments than reactive ones [3]. Moreover, **anticipatory agents** developed the ability to predict and identify changes in the context (w in the figure) of the game, which allowed them to optimize their behavior to each situation. This is very relevant as it is a characteristic present in the participants of our experiments and that other models failed to capture. Such results led the path to our current work, where we analyze the learning, evolutionary and collective dynamics of different types of agents. **We will soon submit our progress in this matter in two parts, to a conference and a journal.**

RESEARCH PLAN & NEXT YEAR PLANNING



	What is known?	What is missing?
Experiments	Perception of Risk ↑ Cooperation Summary statistics	Can we **infer the behavior directly from data**? How can we **influence participants**?
AI models	Analytical models with hypothetical behaviors of limited complexity.	Create **models of human behavior** that take into account their inherent **cognitive capacities**

THESIS OBJECTIVES



- **WP1-M1:** Design of the experiments.
- **WP1-M2:** Design and implement a specific framework required to perform the experiments.
- **WP1-M3:** Perform the experiments and collect the data.
- **WP2-M1:** Search for behavioral models in the data.
- **WP2-M2:** Evaluate whether the inferred models lead to the observed macroscopic behavior.
- **WP3-M1:** Experiment with previous approaches from the literature.
- **WP3-M2:** Create new models to explain the data.
- **WP3-M3:** Compare the behavior of different models.
- **WP4-M1:** Multi-agent simulations.
- **WP4-M2:** Search mechanisms and policies to influence behavior.
- **WP5-M1:** Identify applications of our research to real-world problems.
- **WP5-M2:** Write the PhD thesis.

[1] Milinski, M.; Sommerfeld, R. D.; Krambeck, H.-J.; Reed, F. a.; and Marotzke, J. 2008. The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences of the United States of America* 105(7):2291–2294.
[2] **Fernández Domingos, E.; Burguillo, J.C.; Nowé, A.; and Lenaerts, T.** 2017. Coordinating Human and Agent Behavior in Collective-Risk Scenarios. In *Proceedings of the thirty-first AAAI conference*, AAAI.
[3] **Fernández Domingos, E.; Burguillo, J. C.; and Lenaerts, T.** 2017. Reactive versus anticipative decision making in a novel gift-giving game. In *Proceedings of the thirty-first AAAI conference*, AAAI.

VUB ULB UNIVERSIDADE DE VIGO fwo Research Foundation Flanders Opening new horizons

Check here for more information about me and the project